

1 **Discrete mechanistic pathways underlying genetic predisposition**
2 **to atrial fibrillation are associated with different intermediate**
3 **cardiac phenotypes and risk of cardioembolic stroke**

4

5 Gajendragadkar PR¹, Von Ende A¹, Murgia F¹, Offer A¹, Camm CF¹, Wijesurendra RS¹,
6 Casadei B^{2*}, Hopewell JC^{1*}

7

8 ¹Clinical Trial Service Unit and Epidemiological Studies Unit, Nuffield Department of Population
9 Health, University of Oxford;

10 ²Division of Cardiovascular Medicine, Radcliffe Department of Medicine, University of Oxford

11 *Senior authors contributed equally

12

13 **Address for correspondence**

14 Professor Jemma C Hopewell,
15 CTSU - Nuffield Department of Population Health,
16 Oxford BHF Centre of Research Excellence,
17 Big Data Institute,
18 Old Road Campus, Roosevelt Drive,
19 Oxford, OX3 7LF
20 Jemma.Hopewell@ndph.ox.ac.uk

21

22 **Keywords:** Atrial fibrillation; ischaemic stroke; genetics; big data

1 ABSTRACT

2 Background

3 Genome-wide association studies (GWAS) have clustered candidate genes associated with
4 atrial fibrillation (AF) into biological pathways reflecting different pathophysiological mechanisms.
5 We investigated whether these pathways associate with distinct intermediate phenotypes and confer
6 differing risks of cardioembolic stroke.

8 Methods

9 Three distinct subsets of AF-associated genetic variants, each representing a different
10 mechanistic pathway, i.e., the cardiac muscle function and integrity pathway (15 variants), the cardiac
11 developmental pathway (25 variants), and the cardiac ion channels pathway (12 variants), were
12 identified from a previous AF GWAS. Using genetic epidemiological methods and large-scale
13 datasets such as UK Biobank, deCODE, and GIGASTROKE, we investigated the associations of
14 these pathways with AF-related cardiac intermediate phenotypes, which included ECG parameters
15 (~16,500 ECGs), left atrial and ventricular size and function (~36,000 cardiac MRI scans), and
16 relevant plasma biomarkers (NT-proBNP, ~70,000 samples; high-sensitivity troponin I and T,
17 ~87,000 samples), as well as with subtypes of ischaemic stroke (~11,000 cases).

19 Results

20 Genetic variants representing distinct AF-related mechanistic pathways had significantly
21 different effects on several AF-related phenotypes. In particular, the muscle pathway was associated
22 with a longer PR interval (P for heterogeneity between pathways [$P_{\text{het}} = 1 \times 10^{-10}$]), lower LA
23 emptying fraction ($P_{\text{het}} = 5 \times 10^{-5}$), and higher NT-proBNP ($P_{\text{het}} = 2 \times 10^{-3}$) per log-odds higher risk of
24 AF compared to the developmental and ion channel pathways. By contrast, the ion channel pathway
25 was associated with a lower risk of cardioembolic stroke ($P_{\text{het}} = 0.04$ in European, and 7×10^{-3} in
26 multi-ancestry populations) compared to the other pathways.

1 **Conclusions**

2 Genetic variants representing specific mechanistic pathways for AF are associated with distinct
3 intermediate cardiac phenotypes and a different risk of cardioembolic stroke. These findings provide a
4 better understanding of the aetiological heterogeneity underlying the development of AF and its
5 downstream impact on disease and may offer a route to more targeted treatment strategies.

1 **INTRODUCTION**

2 Atrial fibrillation (AF) is the electrical end-point of radically different pathophysiological
3 pathways^{1,2}. Highlighting this complexity, linkage and family studies have associated AF with rare
4 variants in genes encoding not just ion channels^{3,4}, but transcription factors⁵, hormones⁶, and more
5 recently, cardiac structural proteins⁷⁻¹¹. In line with these findings, pathway and functional enrichment
6 analyses from a large genome-wide association study (GWAS) have suggested that many of the genes
7 annotated to AF-associated loci can be grouped into mechanistic pathways reflecting cardiac and
8 skeletal muscle function and integrity, developmental events, hormone signalling, angiogenesis and
9 myocardial ion channels¹².

10 In individuals with AF, a number of biomarkers and intermediate phenotypes have shown
11 associations with subsequent AF-related outcomes, such as stroke¹³⁻¹⁵. These findings may provide
12 opportunities for improving the management and risk stratification of patients with AF, as well as for
13 repurposing drugs that are already available for treating other diseases¹². However, there remains
14 uncertainty in our understanding of whether AF-associated mechanistic pathways are associated with
15 distinct cardiac phenotypes and differing risks of severe complications.

16 Using clusters of genetic variants representing different AF-associated mechanistic pathways¹²,
17 we investigated associations between genetically-predicted AF pathways and AF-related intermediate
18 phenotypes, encompassing cardiac electrical, structural, and functional characteristics (ECG
19 parameters¹⁶⁻¹⁸, left atrial [LA] size and function¹⁹⁻²¹, left ventricular [LV] function^{22, 23}, N-terminal
20 pro B-type natriuretic peptide [NT-proBNP]^{24, 25}, and high-sensitivity cardiac troponins²⁵⁻²⁷), and
21 examined their relevance to risk of stroke.

1 The AF GWAS meta-analysis¹² also identified variants annotated to genes likely to be involved
2 in calcium handling (4 variants), angiogenesis (1 variant) and hormone signalling (4 variants). These
3 potential pathways were not included in the present study as they were relatively non-specific and
4 comprised multiple heterogeneous biological pathways. For example, the hormone signalling pathway
5 identified genes associated with oestrogen receptors, thyroid receptors and insulin-like growth factor
6 receptors. Further details are provided in **Supplementary Tables 1 - 4**.

7

8 **Study populations and outcome definitions**

9 Details of the datasets used in this study are listed in **Supplementary Table 5**.

10

11 *Atrial fibrillation*

12 Genetic associations with AF were generated directly in UK Biobank^{28,29} in 339,189 unrelated
13 white British participants whose genetic samples passed bioinformatic quality control using logistic
14 regression with adjustments for sex, genotype array, and 7 principal components of population
15 structure. We identified 30,631 cases of AF or flutter, defined using electronic healthcare records with
16 details of definitions in **Supplementary Table 6**. Further details of phenotype codes utilised to define
17 co-morbidities are listed in **Supplementary Table 7**.

18

19 *ECG parameters*

20 As with AF, genetic associations with ECG parameters were generated directly in UK Biobank
21 using logistic regression with adjustments for sex, genotype array, and 7 principal components of
22 population structure. Resting 12-lead ECGs and automated measurements (using Cardiosoft v6
23 program, GE) were available in 16,508 participants, and analyses of these data were conducted after
24 excluding individuals with AF on ECG, or AF identified as described above.

25

1 *Cardiac MRI parameters*

2 Genetic associations with cardiac MRI imaging data including LA structure and function³⁰
3 (35,658 individuals of European ancestry) and LV structure and function³¹ (36,041 individuals of
4 European ancestry), both regardless of history of AF, were based on existing summary data derived
5 from UK Biobank. The UK Biobank cardiac MRI protocol has been reported elsewhere³². In brief, LA
6 parameters were derived from automated processing algorithms; volumes were indexed to Du Bois
7 formula calculated body surface area and cardiac volume-time curves created to derive active and
8 passive LA emptying fractions³⁰. LV parameters were provided by UK Biobank, corrected for known
9 bias from the automated algorithm and indexed to Mosteller formula calculated body surface area³¹.

10

11 *NT-proBNP*

12 Genetic associations with NT-proBNP were based on summary data including 33,043
13 individuals from the UK Biobank Pharma Proteomics Project (UKB-PPP)³³ and 35,559 Icelandic
14 individuals from the deCODE + Icelandic cancer project³⁴. These associations were further validated
15 in summary data including 3,301 healthy blood donors from the INTERVAL study^{35,36} and a GWAS
16 of 3,394 individuals with cardiovascular risk factors from the IMPROVE study^{37,38}. Assays for NT-
17 proBNP included Olink (UKB-PPP, IMPROVE) and SomaScan (deCODE + ICP, INTERVAL).

18

19 *High-sensitivity troponin I and T (hsTnI, hsTnT)*

20 Genetic associations with hsTnI were based on summary data including 48,115 individuals
21 from the Trøndelag Health Study (HUNT) and Generation Scotland Family Health Study (GS:FHS)³⁹
22 and in 14,336 individuals of diverse ancestry (89% European, 11% African) across a number of
23 cohorts (including ARIC) but excluding individuals with baseline coronary heart disease and heart
24 failure as described by Yang *et al*⁴⁰. Genetic associations with hsTnT were based on summary data
25 including 24,617 individuals (76% European, 15% African, 7% Hispanic, 3% Asian) across a number
26 of cohorts (including ARIC, CHS, MESA) but excluding individuals with baseline coronary disease
27 and heart failure⁴⁰.

1

2 *Ischaemic stroke*

3 Genetic associations with stroke outcomes were based on summary data from the GIGASTROKE
4 GWAS meta-analysis⁴¹. Contributing studies defined the aetiology of stroke as per the TOAST (Trial
5 of Org in 10172 Acute Stroke Treatment) classification system⁴². In the present study, we examined
6 associations with cardioembolic stroke (10,804 European ancestry; 12,790 multi-ancestry) as a
7 downstream consequence of AF, and small-vessel stroke as a negative control due to its different risk
8 factor profile (6,811 European; 13,620 multi-ancestry).

9

10 **Statistical analysis**

11 *Primary analyses*

12 Ratio estimates for each individual variant were determined by dividing the gene-outcome
13 association (for each outcome of interest as described above) by the association with AF (taken from
14 the published GWAS for AF¹² after exclusion of UK Biobank participants to reduce the potential for
15 bias). For each specific pathway, the individual variant ratio estimates were then combined using
16 fixed-effects (assuming a single common effect) inverse variance weighted methods, with resulting
17 estimates corresponding to a unit increase in log-odds risk of AF (approximately a 2.7-fold higher
18 risk).

19 Heterogeneity between estimates from different pathways was assessed using Cochran's Q
20 statistic and Wald tests to examine pair-wise differences between individual pathways. Nominal
21 significance was defined as a *P*-value <0.05, with more robust statistical significance defined as a *P*-
22 value <0.01 to account for within-study multiple testing.

23

24 *Sensitivity analyses*

25 We conducted sensitivity analyses to assess the robustness of the results to underlying
26 methodological assumptions (based on a Mendelian randomisation [MR] instrumental variable
27 framework) including: inverse variance weighted random effects (allowing for heterogeneity between

1 instrumental variables); weighted median (in which up to 50% of the information is permitted to be
2 from invalid instrumental variables); weighted mode (which allows the majority of information to be
3 invalid provided the most common MR effect estimate is a consistent estimate of the true MR effect);
4 and MR-Egger (in which all genetic variants are permitted to be invalid instrumental variables,
5 provided that the pleiotropic and risk factor effects of the variants are independently distributed)^{43,44}.
6 MR estimates from each pathway were also estimated after exclusion of outlier variants identified
7 using the MR-PRESSO outlier detection method⁴⁵. ‘Leave-one-out’ analyses were performed to test
8 whether the effects of the pathways on cardioembolic stroke were sensitive to individual variants.
9 Additional post-hoc analyses were also undertaken to estimate the direct effects of the pathways on
10 outcomes showing heterogeneity between pathways (i.e. cardioembolic stroke), independent of their
11 associations with ECG parameters, LA parameters, and markers of LV dysfunction. To further
12 confirm that the scores were unrelated to conventional risk factors, within UK Biobank, participant-
13 level weighted AF risk scores were calculated for each of the pathways and the characteristics of the
14 top and bottom quintiles of score compared.

15 All analyses were performed using SAS (v 9.4, SAS Institute) and R (v 4.3.1, The R
16 Foundation for Statistical Computing).

17

18 **Standard protocol approvals, registrations, and patient consents**

19 UK Biobank data were used from the UK Biobank Resource under Application Number 14568.
20 All procedures and data collection in UK Biobank were approved by the UK Biobank Research Ethics
21 Committee (reference number 11/NW/0274) with participants providing full written informed consent
22 for participation and subsequent use of their data for approved applications. Publicly available non-
23 identifiable summary statistics were used in this study for which the multiple contributory studies to
24 the various datasets had their own consent procedures.

1 RESULTS

2 **Comparable risks of AF unrelated to demographics or co-morbidities across mechanistic** 3 **pathways of genetic predisposition to AF**

4 The estimated associations of AF of the three discrete pathways ('muscle'; 15 genetic variants,
5 'developmental'; 25 variants, and 'ion channel'; 12 variants) are based on the individual genetic
6 variants' associations in the AF GWAS meta-analysis (excluding UK Biobank)¹². As a validation
7 step, we estimated the effects of these pathways on AF in the UK Biobank population. As anticipated,
8 the pathways conferred comparable risks of AF (P for heterogeneity 0.72; **Supplementary Figure 2**).

9 To investigate if the different pathways were associated with risk factors for AF, individual
10 weighted genetic AF risk scores for each of the pathways were calculated in UK Biobank. Participants
11 were split into quintiles of AF risk score and characteristics of the top versus bottom 20% compared
12 (**Supplementary Table 8**). As expected, other than risk of AF, no meaningful differences in
13 demographics or co-morbidities were noted.

15 **Differing associations with cardiac intermediate phenotypes across mechanistic pathways for** 16 **same genetic predisposition to AF**

17 *Surface ECG parameters*

18 In 16,508 participants from UK Biobank with no history of AF, the genetically-predicted AF
19 pathways were associated with different effects on P-wave duration, and PR interval (P for
20 heterogeneity 0.03, and 1×10^{-10} respectively; **Figure 1**). In particular, the developmental and ion
21 channel pathways were associated with a shorter P-wave duration, but the muscle pathway was not
22 (**Figure 1**). Conversely, the muscle pathway was associated with a *longer* PR interval (by 5.58 ms,
23 95% CI: 2.59 to 8.58, $P = 3 \times 10^{-4}$ per log-odds higher genetically-predicted risk of AF), whilst the
24 ion channel pathway was associated with a *shorter* PR interval (by 8.64 ms, 95% CI: -11.53 to -5.75,
25 $P = 4 \times 10^{-9}$ per log-odds higher genetically-predicted risk of AF; **Figure 1**). This pattern remained the
26 same after adjusting the ECG PR interval for the P-wave duration, (P for heterogeneity 3×10^{-10} ;
27 **Supplementary Figure 3**).

1 There were small but significant differences between the associations with QRS duration for
2 the three pathways (P for heterogeneity 2×10^{-5} ; **Figure 1**), but not with corrected QT interval (P for
3 heterogeneity 0.18; **Figure 1**). There were differences between the associations of the pathways with
4 heart rate (P for heterogeneity 0.01; **Figure 1**), with a log-odds higher genetically-predicted risk of
5 AF via the ion channel pathway being associated with a lower heart rate (-1.54 bpm 95% CI: -2.70 to
6 -0.39, $P = 9 \times 10^{-3}$), whereas other pathways showed no association.

7 By contrast, overall AF genetic susceptibility (i.e., as conferred by the 142-variant set) was not
8 associated with differences in PR interval duration ($P = 0.33$; **Supplementary Figure 4**) but was
9 associated with a shorter P-wave (-1.25 ms; 95% CI: -1.79 to -0.71; $P = 6 \times 10^{-6}$) and QRS duration (-
10 0.95 ms; 95% CI: -1.37 to -0.53, $P = 1 \times 10^{-5}$) per unit log-odds higher risk of AF.

11 These findings indicate that mechanistic pathways of genetic predisposition to AF have
12 different associations with ECG parameters (most strikingly for the PR interval), which may not be
13 apparent when overall genetic AF susceptibility is considered.

14

15 *Left atrial size and function*

16 We then investigated the association of the pathways with LA structure and function using
17 GWAS summary data from 35,658 UK Biobank participants who underwent a cardiac MRI scan³⁰.
18 Whilst each AF pathway was similarly associated with larger maximum LA volumes (P for
19 heterogeneity 0.33; **Figure 2A**), the magnitude of the associations with minimum LA volume differed
20 across the pathways (P for heterogeneity 0.02; **Figure 2A**). Specifically, the muscle pathway was
21 associated with a larger minimum LA volume compared with the developmental ($P = 7 \times 10^{-3}$) and
22 with the ion channel pathway ($P = 0.04$). The 142-variant set for overall AF susceptibility was
23 associated with larger maximal and minimum LA volumes (**Supplementary Figure 5A**).

24 When assessing LA function, we used previously defined passive and active phase LA
25 emptying fraction measurements from cardiac MRI³⁰. There were differences in the effects observed
26 between different pathways for passive phase LA emptying fraction (P for heterogeneity 5×10^{-5}).
27 Specifically, a higher genetically-predicted risk of AF *via* the muscle pathway was associated with a

1 lower passive phase LA emptying fraction (-1.18%; 95% CI: -1.75 to -0.61, $P = 5 \times 10^{-5}$), but no
2 association was observed with the other pathways (**Figure 2B**). The comprehensive 142-variant set
3 for AF susceptibility was not associated with passive phase LA emptying fraction (**Supplementary**
4 **Figure 5B**).

5 By contrast, each of the three mechanistic pathways and the 142-variant set were associated
6 with a lower active phase LA emptying fraction (**Figure 2B, Supplementary Figure 5B**), and whilst
7 the effect of the muscle pathway was the most extreme (as observed for the passive phase of LA
8 emptying fraction), there was no significant difference between the pathways (P for heterogeneity
9 0.06; **Figure 2B**).

10 In summary, different mechanistic pathways of genetic predisposition to AF had different
11 associations with LA structure and function. Amongst these, the muscle pathway was associated with
12 larger LA size and more impaired LA function compared with the other pathways.

13

14 *Left ventricular function*

15 Left ventricular dysfunction and AF share a complex bi-directional relationship. To investigate
16 whether biologically-informed pathways underpinning AF risk were differentially linked to imaging
17 biomarkers of LV function, we examined their associations with LV ejection fraction (LVEF) using
18 GWAS summary data from 36,041 UK Biobank participants who underwent a cardiac MRI scan³¹.
19 None of the individual pathways were associated with LVEF, and there were no significant
20 differences between pathways (**Supplementary Table 9**). The muscle pathway was associated with a
21 0.65 ml/m² lower indexed LV stroke volume (95% CI: -1.18 to -0.12; $P = 0.02$), but associations did
22 not differ significantly between the pathways (P for heterogeneity 0.14; **Supplementary Table 9**).
23 The 142-variant set was associated with a 0.17% lower LV ejection fraction (95% CI: -0.23 to -0.004;
24 $P = 0.04$) and a 0.20 ml/m² lower indexed LV stroke volume (95% CI: -0.35 to -0.04; $P = 0.01$).

25

1 *NT-proBNP*

2 We also examined associations with NT-proBNP using summary statistics from a cohort of
3 33,043 individuals from the UKB-PPP³³ (mean age 59 years and similar to the whole UK Biobank
4 cohort in terms of co-morbidities) and a cohort of 35,559 Icelandic individuals from the deCODE +
5 Icelandic cancer projects³⁴ (mean age 55 years with a variety of co-morbidities).

6 The biological pathways showed different associations with NT-proBNP (P for heterogeneity 2
7 $\times 10^{-3}$ and 2×10^{-5} for UKB-PPP and deCODE + ICP respectively; **Figure 3**). In these two large
8 cohorts, only the muscle pathway was associated with higher levels of NT-proBNP (0.13 standard
9 deviation (SD) unit higher; 95% CI: 0.06 to 0.20, $P = 5 \times 10^{-4}$ in UKB-PPP and 0.24 SD unit higher;
10 95% CI: 0.16 to 0.32, $P = 4 \times 10^{-9}$ in deCODE + ICP), with no associations noted for the other
11 pathways (**Figure 3**).

12 To assess the sensitivity of these findings to population characteristics, we examined the
13 associations in a population of 3,301 healthy blood donors from the INTERVAL study³⁵ (mean age 43
14 years) and in 3,394 older adults with risk factors for cardiovascular disease from the IMPROVE
15 study³⁷ (mean age 65 years). Consistent with UKB-PPP and deCODE + ICP, only the muscle pathway
16 was associated with a higher level of NT-proBNP, with significant heterogeneity between pathways
17 seen in both cohorts ($P = 4 \times 10^{-3}$ and $P = 0.04$ respectively; **Supplementary Figure 6**). The 142-
18 variant set for overall risk of AF was also associated with higher levels of NT-proBNP, but only in the
19 UKB-PPP and deCODE + ICP cohorts (**Supplementary Figure 7**).

20

21 *High sensitivity Troponins*

22 To evaluate whether changes in LA function and biomarkers of LV dysfunction were linked to
23 myocardial injury, we investigated associations with hsTnI (multi-ancestry population of 48,115
24 individuals³⁹ and in a separate multi-ancestry population of 14,336 individuals free from prevalent
25 coronary heart disease or heart failure⁴⁰) and hsTnT (24,617 multi-ancestry individuals free from
26 prevalent coronary heart disease or heart failure⁴⁰) using GWAS summary data.

1 We found no differences between the biological pathways in terms of their associations with
2 hsTnI or hsTnT (P for heterogeneity 0.14 and 0.36 for the hsTnI cohorts, and 0.85 for the hsTnT
3 cohort; **Supplementary Figure 8**).

4 The 142-variant set was not associated with higher hsTnI levels in the general population but
5 was associated with a slightly higher level of both TnI and TnT in multi-ethnic populations free from
6 prevalent coronary heart disease or heart failure (**Supplementary Figure 9**).

7 Overall, these data provide further evidence that different mechanisms underlying the genetic
8 risk of AF are associated with distinct cardiac phenotypes and suggest that the muscle pathway is
9 associated with biomarkers of LV dysfunction, but not myocardial injury, in addition to left atrial
10 dysfunction.

11

12 **Genetic predisposition to AF conferred by mechanistic pathways is associated with different** 13 **risks of cardioembolic, but not small vessel ischaemic stroke**

14 We next investigated whether different pathways of genetically-predicted risk for AF may be
15 associated with a different risk of cardioembolic stroke, for which AF is a significant risk factor, but
16 not with small-vessel stroke (as a negative control) due to its differing risk factor profile. We used
17 summary statistics from the GIGASTROKE meta-analysis employing the TOAST classification
18 system⁴² to categorise acute stroke subtypes.

19 In analyses limited to individuals of European ancestry (comparable to the population in which
20 the AF GWAS was undertaken¹²), the three pathways were differentially associated with risk of
21 cardioembolic stroke (P for heterogeneity 0.04; **Figure 4**). Specifically, the ion channel pathway was
22 associated with a lower risk of cardioembolic stroke, most profoundly when compared with the
23 developmental pathway ($P = 0.01$). In the larger multi-ancestry GIGASTROKE meta-analysis, we
24 were similarly able to replicate the differences in the risk of cardioembolic stroke across the different
25 pathways (P for heterogeneity 7×10^{-3}), again with a lower risk conferred by the ion channel pathway
26 (OR for cardioembolic stroke, 1.66; 95% CI 1.43, 1.93 per log odds higher genetically predicted risk
27 of AF), particularly when compared with the developmental pathway ($P = 1 \times 10^{-3}$; **Figure 4**). By
28 contrast, none of the individual pathways was associated with small-vessel ischaemic stroke, with no

1 significant heterogeneity among the estimates (P for heterogeneity 0.20; **Supplementary Figure 10**).
2 Consistent with previous MR studies of AF and stroke^{46,47}, the 142-variant set for overall AF
3 susceptibility was associated with a higher risk of cardioembolic stroke, but not with small-vessel
4 ischaemic stroke (**Supplementary Figure 11**).

5 Overall, these findings suggest that different mechanisms underlying the genetic risk of AF are
6 not only associated with distinct intermediate phenotypes but also with different risks of
7 cardioembolic stroke.

8

9 **Sensitivity Analyses**

10 *Investigations for pleiotropic effects within the pathways*

11 Sensitivity analyses conducted for each phenotype did not suggest pleiotropy, and indicated
12 effect estimates were robust to methodological assumptions. Further details are available in the
13 **Supplementary Results** section and **Supplementary Table 10**. Funnel plots to visualise the effects
14 of each variant on risk of cardioembolic stroke within each AF score are shown in **Supplementary**
15 **Figures 12 - 14** and ‘leave-one-out’ analyses suggested results were robust to the exclusion of
16 individual genetic variants (**Supplementary Figures 15 - 17**).

17

18 *Potential effects of pathway-associated intermediate phenotypes on risk of cardioembolic stroke*

19 We performed analyses for cardioembolic stroke adjusted for the effects of the variants on the
20 intermediate phenotypes, grouped by ECG parameters, markers of LA size and function, and then NT-
21 proBNP and LV ejection fraction. Estimates were comparable before and after adjustment
22 (**Supplementary Tables 11 - 13**), suggesting that associations between the variants and intermediate
23 phenotypes examined in this study are unlikely to fully explain the differences in association with
24 cardioembolic stroke.

25

26

1 DISCUSSION

2 Our findings indicate that distinct biological pathways underlying the genetically-predicted risk
3 of AF are differentially associated with atrial and ventricular phenotypes, and, importantly, with the
4 risk of cardioembolic stroke. This study not only validates the heterogeneous nature of the disease
5 processes predisposing to AF, but highlights the potential for improved risk stratification and more
6 targeted therapeutic strategies.

7 **Biological pathways for AF**

8 A variety of biological pathways underpinning the development of AF have been identified
9 using clinical, laboratory and animal methods^{1,2}. Commensurate with early understanding of AF as a
10 primarily ‘electrical’ disease, initial linkage analyses identified ion-channel genes as potentially
11 relevant in the pathogenesis of AF^{3,48}. More recently, large GWAS have identified many common
12 genetic variants underlying AF susceptibility and have used an array of *in-silico* approaches to
13 suggest their relevance to distinct biological pathways¹². Within this framework however, combining
14 information across the genome without biological considerations could obscure important mechanistic
15 distinctions, prompting us to investigate the relevance of discrete biologically-informed genetic
16 pathways associated with AF risk.

17 Our findings suggest that different AF-related mechanisms have different effects on
18 biomarkers, intermediate phenotypes, and risk of cardioembolic stroke, which may partially explain
19 the heterogeneity seen in observational studies linking ECG¹⁶⁻¹⁸, cardiac imaging parameters¹⁹⁻²², and
20 natriuretic peptides^{49,50} with risk of AF.

21 **Atrial and ventricular substrates**

22 Changes in atrial and ventricular substrates in the context of AF are complex and have been
23 well described^{1,51}, with a number of intermediate phenotypes affecting both atria and ventricles being
24 relevant to AF (e.g. P-wave duration or QT interval^{16,17}, LA volume⁵² or emptying fraction⁵³, and LV
25 dysfunction²²). We initially investigated whether genetic variation associated with risk of AF and
26

1 representing three distinct pathways (ion-channel, muscle structure and function, and developmental
2 pathways) had different effects on intermediate phenotypes in individuals in sinus rhythm.

3 Out of the surface ECG parameters, most strikingly, the muscle and the ion-channel pathways
4 had directionally opposite associations with the PR interval. Strong associations between the
5 pathways with P-wave duration were not seen in our study and thus the heterogeneity seen for
6 associations with the PR interval persisted even after adjusting for P-wave duration. This may suggest
7 that variations in AV node transit time rather than measures of atrial depolarisation reflect different
8 biological pathways to AF and related outcomes. However, ECG parameters in UK Biobank are based
9 on automated ECG recordings and measuring P-waves using automated ECG algorithms has
10 considerable challenges given their low amplitude and variability across different ECG leads^{54, 55}, so it
11 is possible that the PR interval associations simply more accurately reflect atrial electrical parameters
12 or include atrial repolarisation which the P-wave duration may not.

13 Observational data have suggested that both short and long P-wave durations and PR intervals
14 are associated with a higher risk of AF^{17, 18}, and Mendelian randomisation studies have suggested
15 shorter atrial ECG parameters are associated with a higher risk of AF, commensurate with the effects
16 of the ion-channel pathway⁵⁶. Our findings are consistent with data linking an atrial cardiomyopathic
17 substrate to slower atrial conduction⁵⁷⁻⁵⁹, whereas shorter PR intervals may reflect shortening of atrial
18 action potential duration (as seen in those with AF⁶⁰) and more rapid conduction. Further investigation
19 would be needed to establish whether these insights also provide information on the mechanisms
20 underpinning AF development (e.g. re-entrant versus multiple wavelets, or trigger-dependent versus
21 substrate dependent).

22 Our study showed differing associations of the AF-associated biological pathways with NT-
23 proBNP, with the muscle pathway being associated with higher plasma levels across all cohorts,
24 including in seemingly healthy, younger subjects. We also observed associations between the muscle
25 and ion-channel pathways and QRS interval duration on ECG, and that the muscle pathway was
26 associated with a lower indexed LV stroke volume. The muscle pathway also had the strongest
27 associations with lower left atrial function, particularly in the passive emptying phase, which may
28 indicate associations with LV diastolic dysfunction more generally.

1 Associations between NT-proBNP and AF have been reported widely in the literature alongside
2 a consistently increased risk of stroke^{24,25}. Our previous work also indicated that people with “lone”
3 AF continue to show evidence of subtle ventricular dysfunction and impaired myocardial energetics
4 on cardiac magnetic resonance imaging/spectroscopy, despite normalisation of rhythm ~6 months
5 post-ablation, consistent with the presence of a subclinical cardiomyopathy underpinning AF²².

6 Our analyses did not support any consistent difference in the effect of AF-associated pathways
7 on hsTnI/T levels in a general population, or separately in a cohort of individuals free from coronary
8 artery disease or heart failure. In subjects with AF, hsTnI/T levels have been shown to be relevant to
9 risk-stratification for future cardiovascular outcomes²⁵⁻²⁷. Whilst the muscle pathway in particular was
10 associated with markers of LA and LV dysfunction, this does not necessarily reflect the presence of
11 myocardial injury and cell death. Other findings suggest that this relationship may be accounted for, at
12 least in part, by concomitant coronary disease⁶¹.

13

14 **Different risks of stroke based on different pathways to AF**

15 Our study suggests that genetic variants representing different pathways associated with risk of
16 AF can confer differing risks of cardioembolic stroke. There is evidence suggesting that structural or
17 electrical changes within the atria may lead to thromboembolism and AF independently^{52, 62-64}.
18 Furthermore, missense mutations in genes associated with changes in atrial structure confer an
19 increased risk of both AF and stroke^{8,9,65}. This has given rise to the hypothesis that AF and stroke
20 may be the distinct endpoints of a cardiomyopathic substrate^{15, 66-69}. In our study, the same risk of AF
21 underpinned by different biological pathways was associated with differences in the risk of
22 cardioembolic stroke.

23 Whilst we demonstrated associations between the ion channel pathway and ECG parameters, as
24 well as the muscle pathway and some of the known intermediate phenotypes for atrial and ventricular
25 cardiomyopathy, the developmental pathway for AF was associated with a higher risk of stroke in the
26 absence of an ‘obvious’ atrial phenotype as assessed by conventional biomarkers. Although the
27 function of the specific variants included in the developmental pathway (*PITX2*, *NKX2-5*, *GATA 4*
28 and *TBX5*) is only partially known⁴, these genes may impact cardiac development in early life, with

1 evidence that missense mutations in *TBX5* can lead to profound atrial transcriptional alteration and
2 AF⁷⁰. Investigation of more sophisticated biomarkers may yield more insights into understanding how
3 this pathway impacts AF and stroke risk.

4 Our study provides insights into disease aetiology that may allow better risk profiling of
5 patients and help inform treatment pathways. Further studies are needed to understand if individuals
6 with AF who have risk profile patterns associated with different biological mechanisms may derive
7 more benefit from specific interventions.

8

9 **Limitations**

10 To define our biological pathways, we used the current largest GWAS for AF and its multi-
11 modal *in-silico* gene and pathway annotations, and selected variants meeting GWAS significance¹².
12 Pathway annotation is complex and based on existing knowledge of functions of genes and thus
13 limited by new scientific discovery. Whilst the associations observed with different intermediate
14 phenotypes validates the different pathways represented to some extent, we did not comprehensively
15 test distinct biological pathways as a whole, but rather specific variants within these pathways known
16 to be associated with AF.

17 The GWAS for AF used for genetic variant selection were conducted in participants with
18 European ancestry, and outcomes were also chiefly examined in studies of European ancestry, so
19 further studies would be needed to assess the generalisability of our results across more diverse
20 ethnicities.

21

22 **Conclusions**

23 Genetic variants representing distinct biological pathways for AF susceptibility have differing
24 effects on intermediate cardiac electrical, structural and functional phenotypes, extending to the
25 ventricle, and risk of cardioembolic stroke. These findings provide a better understanding of the
26 aetiological heterogeneity driving the development of AF and may offer a route to genetically-
27 informed treatment strategies.

1 **ACKNOWLEDGEMENTS AND SOURCES OF FUNDING**

2 This work was supported by the British Heart Foundation (grant numbers FS/17/17/32438
3 [PRG, BC], FS/20/15/34920 [CFC, JCH], RG/16/12/3245 and CH/12/3/29609 [BC], and
4 FS/14/55/30806 and CH/1996001/9454 [JCH]). The authors also acknowledge support from: the
5 European Union (grant agreement 633196 - CATCH ME); BHF Centre for Research Excellence,
6 Oxford, National Institute of Health Research (NIHR) Oxford Biomedical Research Centre, and
7 Nuffield Department of Population Health, University of Oxford, UK. British Heart Foundation.

8 The funders had no role in study design, data collection or analysis, preparation of the
9 manuscript, or decision to publish.

10

11 **DISCLOSURES**

12 PRG, AVE, CFC, FM, RSW, and JCH work at the Clinical Trial Service Unit &
13 Epidemiological Studies Unit, Nuffield Department of Population Health, which receives research
14 grants from industry that are governed by University of Oxford contracts that protect its
15 independence, and has a staff policy of not taking personal payments from industry; further details
16 can be found at <https://www.ndph.ox.ac.uk/about/independence-of-research>. BC is supported by a
17 British Heart Foundation (BHF) personal chair; her research is funded by two BHF Programme
18 Grants, the BHF Centre for Research Excellence and the NIHR Oxford Biomedical Research Centre.
19 She also receives in-kind research support from iRhythm and Roche Diagnostics. JCH is supported by
20 the British Heart Foundation; her research is funded by the BHF Centre for Research Excellence
21 Oxford and the NIHR Oxford Biomedical Research Centre, and by grants from industry held in
22 accordance with the policy detailed above, and the Nuffield Department of Population Health.

23

24 **DATA AVAILABILITY STATEMENT**

25 Data are available from UK Biobank (<http://biobank.ndph.ox.ac.uk/showcase/>) in accordance
26 with their published data access procedures. All other data are publicly available or were obtained

- 1 from corresponding authors of studies as detailed in the **Methods** section and **Supplementary Table**
- 2 **5.**

REFERENCES

1. Wijesurendra RS, Casadei B. Atrial fibrillation: effects beyond the atrium? *Cardiovasc Res.* 2015;105:238-47.
2. Nattel S, Heijman J, Zhou L, Dobrev D. Molecular Basis of Atrial Fibrillation Pathophysiology and Therapy: A Translational Perspective. *Circ Res.* 2020;127:51-72.
3. Chen YH, Xu SJ, Bendahhou S, Wang XL, Wang Y, Xu WY, Jin HW, Sun H, Su XY, Zhuang QN et al. KCNQ1 gain-of-function mutation in familial atrial fibrillation. *Science.* 2003;299:251-4.
4. Roselli C, Rienstra M, Ellinor PT. Genetics of Atrial Fibrillation in 2020: GWAS, Genome Sequencing, Polygenic Risk, and Beyond. *Circ Res.* 2020;127:21-33.
5. Postma AV, van de Meerakker JB, Mathijssen IB, Barnett P, Christoffels VM, Ilgun A, Lam J, Wilde AA, Lekanne Deprez RH, Moorman AF. A gain-of-function TBX5 mutation is associated with atypical Holt-Oram syndrome and paroxysmal atrial fibrillation. *Circ Res.* 2008;102:1433-42.
6. Hodgson-Zingman DM, Karst ML, Zingman LV, Heublein DM, Darbar D, Herron KJ, Ballew JD, de Andrade M, Burnett JC, Jr., Olson TM. Atrial natriuretic peptide frameshift mutation in familial atrial fibrillation. *N Engl J Med.* 2008;359:158-65.
7. Holm H, Gudbjartsson DF, Sulem P, Masson G, Helgadóttir HT, Zanon C, Magnusson OT, Helgason A, Saemundsdóttir J, Gylfason A et al. A rare variant in MYH6 is associated with high risk of sick sinus syndrome. *Nat Genet.* 2011;43:316-20.
8. Gudbjartsson DF, Holm H, Sulem P, Masson G, Oddsson A, Magnusson OT, Saemundsdóttir J, Helgadóttir HT, Helgason H, Johannsdóttir H et al. A frameshift deletion in the sarcomere gene MYL4 causes early-onset familial atrial fibrillation. *Eur Heart J.* 2017;38:27-34.
9. Thorólfsson RB, Sveinbjörnsson G, Sulem P, Helgadóttir A, Gretarsdóttir S, Benonisdóttir S, Magnúsdóttir A, Davidsson OB, Rajamani S, Roden DM et al. A Missense Variant in PLEC Increases Risk of Atrial Fibrillation. *J Am Coll Cardiol.* 2017;70:2157-2168.
10. Ahlberg G, Refsgaard L, Lundegaard PR, Andreasen L, Ranthe MF, Linscheid N, Nielsen JB, Melbye M, Haunso S, Sajadieh A et al. Rare truncating variants in the sarcomeric protein titin associate with familial and early-onset atrial fibrillation. *Nat Commun.* 2018;9:4316.
11. Choi SH, Weng LC, Roselli C, Lin H, Haggerty CM, Shoemaker MB, Barnard J, Arking DE, Chasman DI, Albert CM et al. Association Between Titin Loss-of-Function Variants and Early-Onset Atrial Fibrillation. *JAMA.* 2018;320:2354-2364.

12. Nielsen JB, Thorolfsdottir RB, Fritsche LG, Zhou W, Skov MW, Graham SE, Herron TJ, McCarthy S, Schmidt EM, Sveinbjornsson G et al. Biobank-driven genomic discovery yields new insight into atrial fibrillation biology. *Nat Genet.* 2018;50:1234-1239.
13. Hijazi Z, Oldgren J, Siegbahn A, Granger CB, Wallentin L. Biomarkers in atrial fibrillation: a clinical review. *Eur Heart J.* 2013;34:1475-80.
14. Cameron A, Cheng HK, Lee RP, Doherty D, Hall M, Khashayar P, Lip GYH, Quinn T, Abdul-Rahim A, Dawson J. Biomarkers for Atrial Fibrillation Detection After Stroke: Systematic Review and Meta-analysis. *Neurology.* 2021;97:e1775-e1789.
15. Kamel H, Okin PM, Elkind MS, Iadecola C. Atrial Fibrillation and Mechanisms of Stroke: Time for a New Model. *Stroke.* 2016;47:895-900.
16. Nielsen JB, Graff C, Pietersen A, Lind B, Struijk JJ, Olesen MS, Haunso S, Gerds TA, Svendsen JH, Kober L et al. J-shaped association between QTc interval duration and the risk of atrial fibrillation: results from the Copenhagen ECG study. *J Am Coll Cardiol.* 2013;61:2557-64.
17. Nielsen JB, Kuhl JT, Pietersen A, Graff C, Lind B, Struijk JJ, Olesen MS, Sinner MF, Bachmann TN, Haunso S et al. P-wave duration and the risk of atrial fibrillation: Results from the Copenhagen ECG Study. *Heart Rhythm.* 2015;12:1887-95.
18. Nielsen JB, Pietersen A, Graff C, Lind B, Struijk JJ, Olesen MS, Haunso S, Gerds TA, Ellinor PT, Kober L et al. Risk of atrial fibrillation as a function of the electrocardiographic PR interval: results from the Copenhagen ECG Study. *Heart Rhythm.* 2013;10:1249-56.
19. Laukkanen JA, Kurl S, Eranen J, Huttunen M, Salonen JT. Left atrium size and the risk of cardiovascular death in middle-aged men. *Arch Intern Med.* 2005;165:1788-93.
20. Njoku A, Kannabhiran M, Arora R, Reddy P, Gopinathannair R, Lakkireddy D, Dominic P. Left atrial volume predicts atrial fibrillation recurrence after radiofrequency ablation: a meta-analysis. *Europace.* 2018;20:33-42.
21. Reddy YNV, Obokata M, Verbrugge FH, Lin G, Borlaug BA. Atrial Dysfunction in Patients With Heart Failure With Preserved Ejection Fraction and Atrial Fibrillation. *J Am Coll Cardiol.* 2020;76:1051-1064.
22. Wijesurendra RS, Liu A, Eichhorn C, Ariga R, Levelt E, Clarke WT, Rodgers CT, Karamitsos TD, Bashir Y, Ginks M et al. Lone Atrial Fibrillation Is Associated With Impaired Left Ventricular Energetics That Persists Despite Successful Catheter Ablation. *Circulation.* 2016;134:1068-1081.

23. Dries DL, Exner DV, Gersh BJ, Domanski MJ, Waclawiw MA, Stevenson LW. Atrial fibrillation is associated with an increased risk for mortality and heart failure progression in patients with asymptomatic and symptomatic left ventricular systolic dysfunction: a retrospective analysis of the SOLVD trials. *Studies of Left Ventricular Dysfunction. J Am Coll Cardiol.* 1998;32:695-703.
24. Di Castelnuovo A, Veronesi G, Costanzo S, Zeller T, Schnabel RB, de Curtis A, Salomaa V, Borchini R, Ferrario M, Giampaoli S et al. NT-proBNP (N-Terminal Pro-B-Type Natriuretic Peptide) and the Risk of Stroke. *Stroke.* 2019;50:610-617.
25. Hijazi Z, Oldgren J, Andersson U, Connolly SJ, Ezekowitz MD, Hohnloser SH, Reilly PA, Vinereanu D, Siegbahn A, Yusuf S et al. Cardiac biomarkers are associated with an increased risk of stroke and death in patients with atrial fibrillation: a Randomized Evaluation of Long-term Anticoagulation Therapy (RE-LY) substudy. *Circulation.* 2012;125:1605-16.
26. Hijazi Z, Wallentin L, Siegbahn A, Andersson U, Alexander JH, Atar D, Gersh BJ, Hanna M, Harjola VP, Horowitz JD et al. High-sensitivity troponin T and risk stratification in patients with atrial fibrillation during treatment with apixaban or warfarin. *J Am Coll Cardiol.* 2014;63:52-61.
27. Costabel JP, Burgos LM, Trivi M. The Significance Of Troponin Elevation In Atrial Fibrillation. *J Atr Fibrillation.* 2017;9:1530.
28. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D, Delaneau O, O'Connell J et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature.* 2018;562:203-209.
29. Littlejohns TJ, Sudlow C, Allen NE, Collins R. UK Biobank: opportunities for cardiovascular research. *Eur Heart J.* 2019;40:1158-1166.
30. Ahlberg G, Andreassen L, Ghouse J, Bertelsen L, Bundgaard H, Haunso S, Svendsen JH, Olesen MS. Genome-wide association study identifies 18 novel loci associated with left atrial volume and function. *Eur Heart J.* 2021;42:4523-4534.
31. 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.
32. 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.

33. Sun BB, Chiou J, Traylor M, Benner C, Hsu YH, Richardson TG, Surendran P, Mahajan A, Robins C, Vasquez-Grinnell SG et al. Plasma proteomic associations with genetics and health in the UK Biobank. *Nature*. 2023;622:329-338.
34. Ferkingstad E, Sulem P, Atlason BA, Sveinbjornsson G, Magnusson MI, Styrismisdottir EL, Gunnarsdottir K, Helgason A, Oddsson A, Halldorsson BV et al. Large-scale integration of the plasma proteome with genetics and disease. *Nat Genet*. 2021;53:1712-1721.
35. Sun BB, Maranville JC, Peters JE, Stacey D, Staley JR, Blackshaw J, Burgess S, Jiang T, Paige E, Surendran P et al. Genomic atlas of the human plasma proteome. *Nature*. 2018;558:73-79.
36. Di Angelantonio E, Thompson SG, Kaptoge S, Moore C, Walker M, Armitage J, Ouwehand WH, Roberts DJ, Danesh J. Efficiency and safety of varying the frequency of whole blood donation (INTERVAL): a randomised trial of 45 000 donors. *Lancet*. 2017;390:2360-2371.
37. Folkersen L, Fauman E, Sabater-Lleal M, Strawbridge RJ, Franberg M, Sennblad B, Baldassarre D, Veglia F, Humphries SE, Rauramaa R et al. Mapping of 79 loci for 83 plasma protein biomarkers in cardiovascular disease. *PLoS Genet*. 2017;13:e1006706.
38. Baldassarre D, Nyssonen K, Rauramaa R, de Faire U, Hamsten A, Smit AJ, Mannarino E, Humphries SE, Giral P, Grossi E et al. Cross-sectional analysis of baseline data to identify the major determinants of carotid intima-media thickness in a European population: the IMPROVE study. *Eur Heart J*. 2010;31:614-22.
39. Moksnes MR, Rosjo H, Richmond A, Lyngbakken MN, Graham SE, Hansen AF, Wolford BN, Gagliano Taliun SA, LeFaive J, Rasheed H et al. Genome-wide association study of cardiac troponin I in the general population. *Hum Mol Genet*. 2021;30:2027-2039.
40. Yang Y, Bartz TM, Brown MR, Guo X, Zilhao NR, Trompet S, Weiss S, Yao J, Brody JA, Defilippi CR et al. Identification of Functional Genetic Determinants of Cardiac Troponin T and I in a Multiethnic Population and Causal Associations With Atrial Fibrillation. *Circ Genom Precis Med*. 2021;14:e003460.
41. Mishra A, Malik R, Hachiya T, Jurgenson T, Namba S, Posner DC, Kamanu FK, Koido M, Le Grand Q, Shi M et al. Stroke genetics informs drug discovery and risk prediction across ancestries. *Nature*. 2022;611:115-123.
42. Adams HP, Jr., Bendixen BH, Kappelle LJ, Biller J, Love BB, Gordon DL, Marsh EE, 3rd. Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment. *Stroke*. 1993;24:35-41.

43. Burgess S, Bowden J, Fall T, Ingelsson E, Thompson SG. Sensitivity Analyses for Robust Causal Inference from Mendelian Randomization Analyses with Multiple Genetic Variants. *Epidemiology*. 2017;28:30-42.
44. Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int J Epidemiol*. 2017;46:1985-1998.
45. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat Genet*. 2018;50:693-698.
46. Lubitz SA, Parsons OE, Anderson CD, Benjamin EJ, Malik R, Weng LC, Dichgans M, Sudlow CL, Rothwell PM, Rosand J et al. Atrial Fibrillation Genetic Risk and Ischemic Stroke Mechanisms. *Stroke*. 2017;48:1451-1456.
47. Frerich S, Malik R, Georgakis MK, Sinner MF, Kittner SJ, Mitchell BD, Dichgans M. Cardiac Risk Factors for Stroke: A Comprehensive Mendelian Randomization Study. *Stroke*. 2021:STROKEAHA121036306.
48. Chen LY, Ballew JD, Herron KJ, Rodeheffer RJ, Olson TM. A common polymorphism in SCN5A is associated with lone atrial fibrillation. *Clin Pharmacol Ther*. 2007;81:35-41.
49. Seegers J, Zabel M, Gruter T, Ammermann A, Weber-Kruger M, Edelmann F, Gelbrich G, Binder L, Herrmann-Lingen C, Groschel K et al. Natriuretic peptides for the detection of paroxysmal atrial fibrillation. *Open Heart*. 2015;2:e000182.
50. Werhahn SM, Becker C, Mende M, Haarmann H, Nolte K, Laufs U, Zeynalova S, Loffler M, Dargès N, Husser D et al. NT-proBNP as a marker for atrial fibrillation and heart failure in four observational outpatient trials. *ESC Heart Fail*. 2022;9:100-109.
51. Coats AJS, Heymans S, Farmakis D, Anker SD, Backs J, Bauersachs J, de Boer RA, Celutkiene J, Cleland JGF, Dobrev D et al. Atrial disease and heart failure: the common soil hypothesis proposed by the Heart Failure Association of the European Society of Cardiology. *Eur Heart J*. 2021.
52. Benjamin EJ, D'Agostino RB, Belanger AJ, Wolf PA, Levy D. Left atrial size and the risk of stroke and death. The Framingham Heart Study. *Circulation*. 1995;92:835-41.
53. Lim DJ, Ambale-Ventakesh B, Ostovaneh MR, Zghaib T, Ashikaga H, Wu C, Watson KE, Hughes T, Shea S, Heckbert SR et al. Change in left atrial function predicts incident atrial fibrillation: the Multi-Ethnic Study of Atherosclerosis. *Eur Heart J Cardiovasc Imaging*. 2019;20:979-987.

54. Schlapfer J, Wellens HJ. Computer-Interpreted Electrocardiograms: Benefits and Limitations. *J Am Coll Cardiol*. 2017;70:1183-1192.
55. Saclova L, Nemcova A, Smisek R, Smital L, Vitek M, Ronzhina M. Reliable P wave detection in pathological ECG signals. *Scientific Reports*. 2022;12:6589.
56. Gajendragadkar PR, Von Ende A, Ibrahim M, Valdes-Marquez E, Camm CF, Murgia F, Stiby A, Casadei B, Hopewell JC. Assessment of the causal relevance of ECG parameters for risk of atrial fibrillation: A mendelian randomisation study. *PLoS Med*. 2021;18:e1003572.
57. Fukumoto K, Habibi M, Ipek EG, Zahid S, Khurram IM, Zimmerman SL, Zipunnikov V, Spragg D, Ashikaga H, Trayanova N et al. Association of Left Atrial Local Conduction Velocity With Late Gadolinium Enhancement on Cardiac Magnetic Resonance in Patients With Atrial Fibrillation. *Circ Arrhythm Electrophysiol*. 2016;9:e002897.
58. Honarbakhsh S, Schilling RJ, Orini M, Providencia R, Keating E, Finlay M, Sporton S, Chow A, Earley MJ, Lambiase PD et al. Structural remodeling and conduction velocity dynamics in the human left atrium: Relationship with reentrant mechanisms sustaining atrial fibrillation. *Heart Rhythm*. 2019;16:18-25.
59. Heil E, Gerds-Li JH, Keznickl-Pulst J, Furundzija-Cabraja V, Hohendanner F, Boldt LH, Stawowy P, Schoeppenthau D. Left atrial conduction times and regional velocities in persistent atrial fibrillation patients with and without fibrotic atrial cardiomyopathy. *Heart Vessels*. 2023;38:1277-1287.
60. Reilly SN, Liu X, Carnicer R, Recalde A, Muszkiewicz A, Jayaram R, Carena MC, Wijesurendra R, Stefanini M, Surdo NC et al. Up-regulation of miR-31 in human atrial fibrillation begets the arrhythmia by depleting dystrophin and neuronal nitric oxide synthase. *Sci Transl Med*. 2016;8:340ra74.
61. Kaura A, Arnold AD, Panoulas V, Glampson B, Davies J, Mulla A, Woods K, Omigie J, Shah AD, Channon KM et al. Prognostic significance of troponin level in 3121 patients presenting with atrial fibrillation (The NIHR Health Informatics Collaborative TROP-AF study). *J Am Heart Assoc*. 2020;9:e013684.
62. Spartera M, Stracquadiano A, Pessoa-Amorim G, Von Ende A, Fletcher A, Manley P, Ferreira VM, Hess AT, Hopewell JC, Neubauer S et al. The impact of atrial fibrillation and stroke risk factors on left atrial blood flow characteristics. *Eur Heart J Cardiovasc Imaging*. 2021;23:115-123.

63. Kamel H, Bartz TM, Elkind MSV, Okin PM, Thacker EL, Patton KK, Stein PK, deFilippi CR, Gottesman RF, Heckbert SR et al. Atrial Cardiopathy and the Risk of Ischemic Stroke in the CHS (Cardiovascular Health Study). *Stroke*. 2018;49:980-986.
64. Kamel H, Hunter M, Moon YP, Yaghi S, Cheung K, Di Tullio MR, Okin PM, Sacco RL, Soliman EZ, Elkind MS. Electrocardiographic Left Atrial Abnormality and Risk of Stroke: Northern Manhattan Study. *Stroke*. 2015;46:3208-12.
65. Disertori M, Quintarelli S, Grasso M, Pilotto A, Narula N, Favalli V, Canclini C, Diegoli M, Mazzola S, Marini M et al. Autosomal recessive atrial dilated cardiomyopathy with standstill evolution associated with mutation of Natriuretic Peptide Precursor A. *Circ Cardiovasc Genet*. 2013;6:27-36.
66. Kamel H, Okin PM, Longstreth WT, Jr., Elkind MS, Soliman EZ. Atrial cardiopathy: a broadened concept of left atrial thromboembolism beyond atrial fibrillation. *Future Cardiol*. 2015;11:323-31.
67. Goette A, Kalman JM, Aguinaga L, Akar J, Cabrera JA, Chen SA, Chugh SS, Corradi D, D'Avila A, Dobrev D et al. EHRA/HRS/APHRS/SOLAECE expert consensus on atrial cardiomyopathies: definition, characterization, and clinical implication. *Europace*. 2016;18:1455-1490.
68. Guichard JB, Nattel S. Atrial Cardiomyopathy: A Useful Notion in Cardiac Disease Management or a Passing Fad? *J Am Coll Cardiol*. 2017;70:756-765.
69. Sajeev JK, Kalman JM, Dewey H, Cooke JC, Teh AW. The Atrium and Embolic Stroke: Myopathy Not Atrial Fibrillation as the Requisite Determinant? *JACC Clin Electrophysiol*. 2020;6:251-261.
70. van Ouwkerk AF, Bosada FM, van Duijvenboden K, Houweling AC, Scholman KT, Wakker V, Allaart CP, Uhm JS, Mathijssen IB, Baartscheer T et al. Patient-Specific TBX5-G125R Variant Induces Profound Transcriptional Deregulation and Atrial Dysfunction. *Circulation*. 2022;145:606-619.

1 **FIGURE LEGENDS**

2 **Figure 1. Effects of genetically predicted AF biological pathways on surface ECG parameters in**
3 **UK Biobank participants**

4 Effect of genetically predicted AF biological pathways on ECG parameters (in milliseconds) or
5 heart rate (in beats per minute) in a UK Biobank cohort of 16,508 participants without a history of
6 AF. Boxes represent point estimates of effect per log-odds higher genetically predicted risk of AF,
7 with their size inversely proportional to variance and solid lines representing 95% confidence
8 intervals (CI). Effect sizes were calculated using inverse variance weighted methods, and
9 heterogeneity using Cochran Q.

10

11 **Figure 2. Effects of genetically predicted AF biological pathways on left atrial structure and**
12 **function in UK Biobank participants**

13 Effect of AF pathways on (A) indexed left atrial (LA) volumes in ml/m² and (B) LA passive
14 and active phase emptying fractions from 35,658 cardiac MRIs in UK Biobank cohort. Boxes
15 represent point estimates of effect per log odds higher genetically predicted risk of AF, with their size
16 inversely proportional to variance and solid lines representing 95% confidence intervals (CI). Effect
17 sizes were calculated using inverse variance weighted methods from summary GWAS data, and
18 heterogeneity using Cochran Q.

19

20 **Figure 3. Effects of genetically predicted AF biological pathways on natriuretic peptide levels**

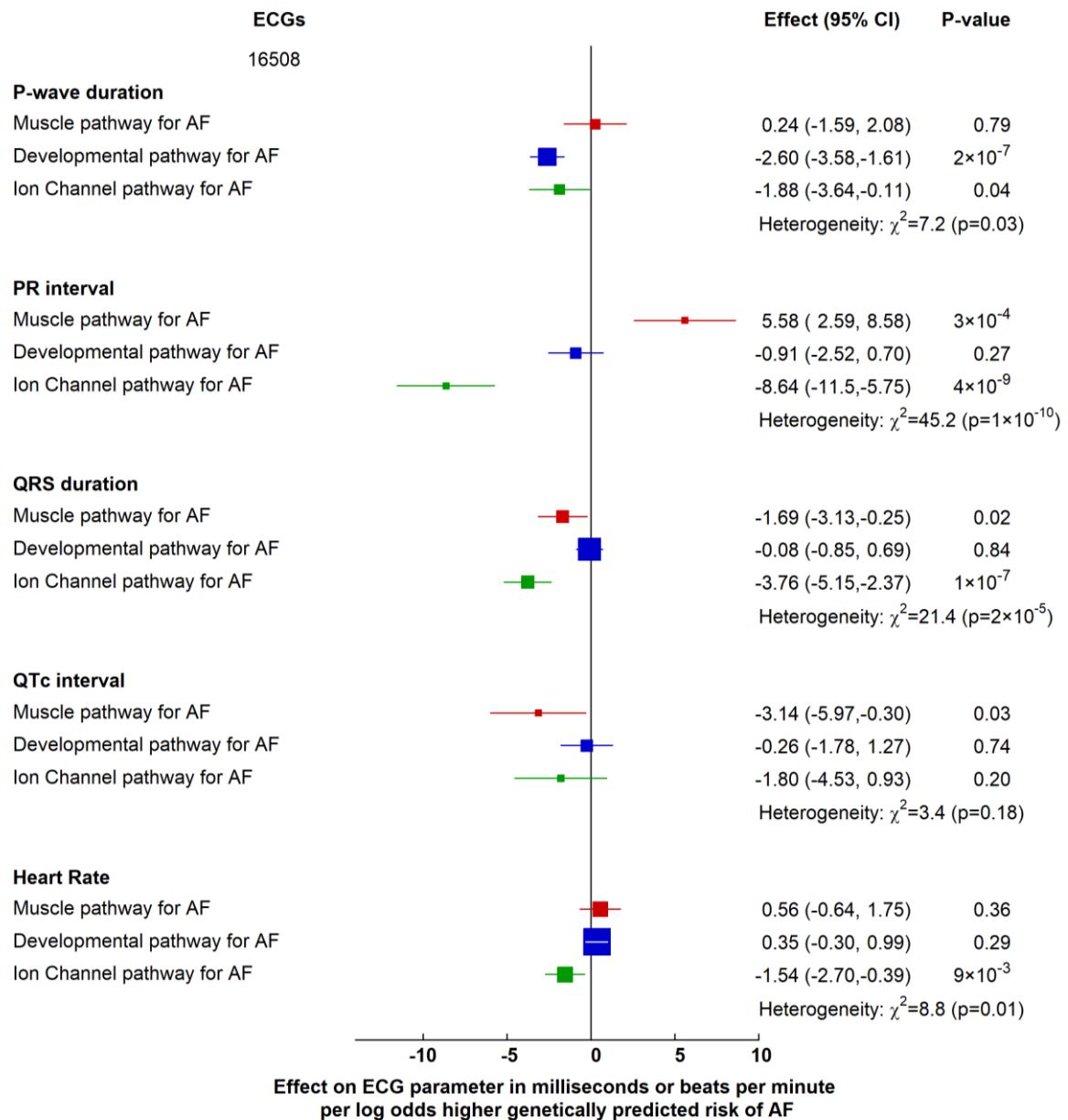
21 Effect of AF pathways on N-terminal pro B-type natriuretic peptide (NT-proBNP) in standard
22 deviation (SD) units in 33,043 British individuals from the UK Biobank Pharma Proteomics Project
23 (UKB-PPP) and, in 35,559 Icelandic individuals from the deCODE and Icelandic cancer project
24 (ICP). Boxes represent point estimates of effect per log odds higher genetically predicted risk of AF,
25 with their size inversely proportional to variance and solid lines representing 95% confidence
26 intervals (CI). Effect sizes were calculated using inverse variance weighted methods from summary
27 GWAS data and heterogeneity using Cochran Q.

1 **Figure 4. Effects of genetically predicted AF biological pathways on cardioembolic stroke**

2 Odds ratio (OR) for developing types of ischaemic stroke in the GIGASTROKE dataset. Boxes
3 represent point estimates of effect per log odds higher genetically predicted risk of AF, with their size
4 inversely proportional to variance and solid lines representing 95% confidence intervals (CI). ORs
5 were calculated using inverse variance weighted methods from summary GWAS data, and
6 heterogeneity using Cochran Q.

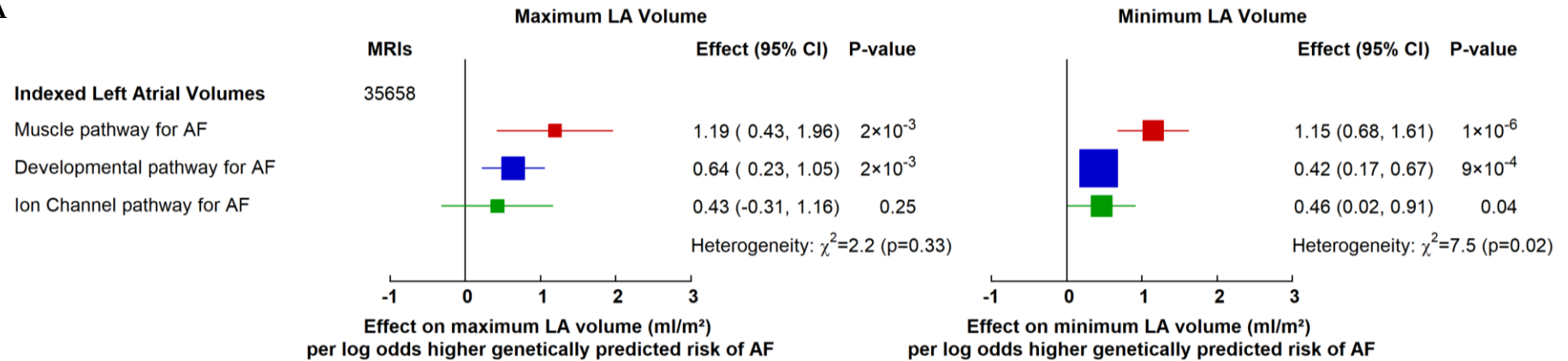
FIGURES

Figure 1. Effects of genetically predicted AF biological pathways on surface ECG parameters in UK Biobank participants



Effect of genetically predicted AF biological pathways on ECG parameters (in milliseconds) or heart rate (in beats per minute) in a UK Biobank cohort of 16,508 participants without a history of AF. Boxes represent point estimates of effect per log-odds higher genetically predicted risk of AF, with their size inversely proportional to variance and solid lines representing 95% confidence intervals (CI). Effect sizes were calculated using inverse variance weighted methods, and heterogeneity using Cochran Q.

A



B

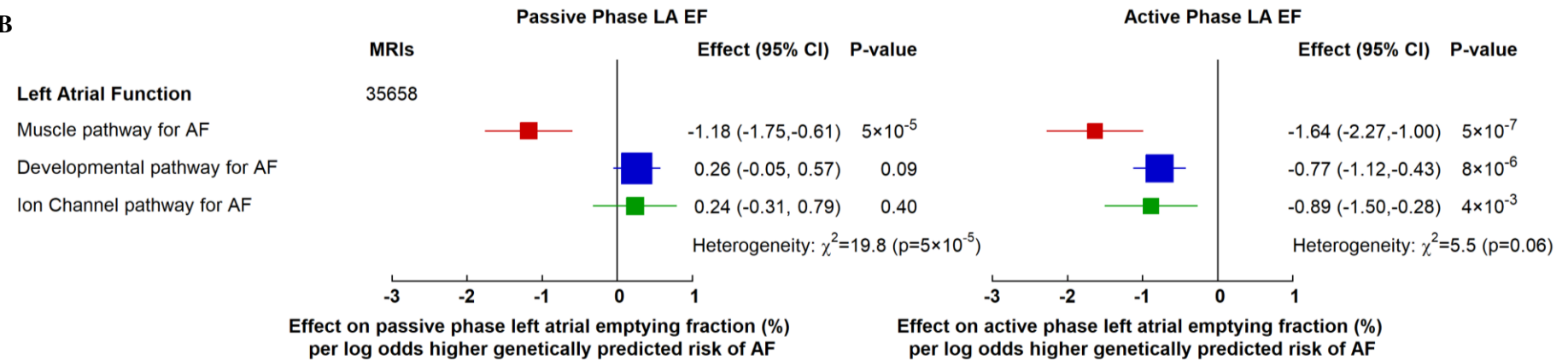
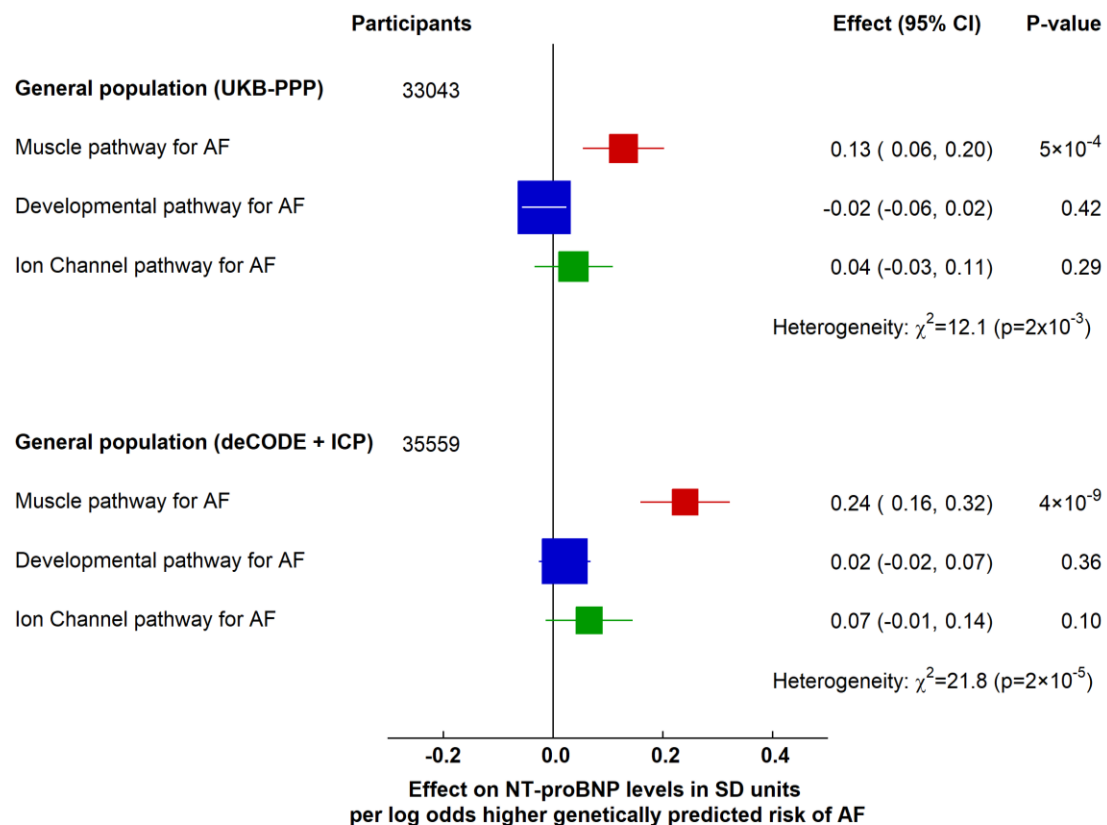


Figure 2. Effects of genetically predicted AF biological pathways on left atrial structure and function in UK Biobank participants

Effect of AF pathways on (A) indexed left atrial (LA) volumes in ml/m² and (B) LA passive and active phase emptying fractions from 35,658 cardiac MRIs in UK Biobank cohort. Boxes represent point estimates of effect per log odds higher genetically predicted risk of AF, with their size inversely proportional to variance and solid lines representing 95% confidence intervals (CI). Effect sizes were calculated using inverse variance weighted methods from summary GWAS data, and heterogeneity using Cochran Q.

Figure 3. Effects of genetically predicted AF biological pathways on natriuretic peptide levels



Effect of AF pathways on N-terminal pro B-type natriuretic peptide (NT-proBNP) in standard deviation (SD) units in 33,043 British individuals from the UK Biobank Pharma Proteomics Project (UKB-PPP) and, in 35,559 Icelandic individuals from the deCODE and Icelandic cancer project (ICP). Boxes represent point estimates of effect per log odds higher genetically predicted risk of AF, with their size inversely proportional to variance and solid lines representing 95% confidence intervals (CI). Effect sizes were calculated using inverse variance weighted methods from summary GWAS data and heterogeneity using Cochran Q.

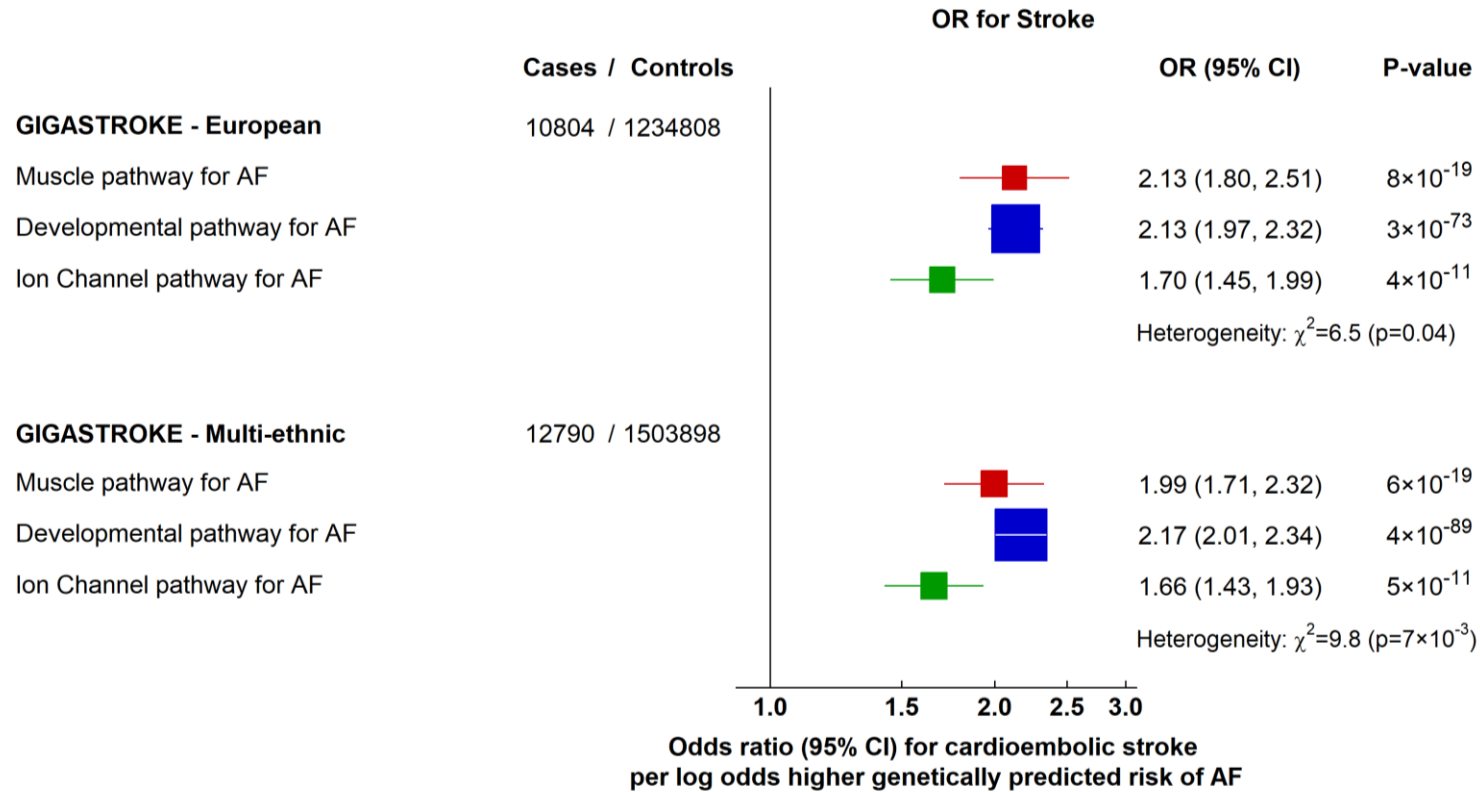


Figure 4. Effects of genetically predicted AF biological pathways on cardioembolic stroke

Odds ratio (OR) for developing types of ischaemic stroke in the GIGASTROKE dataset. Boxes represent point estimates of effect per log odds higher genetically predicted risk of AF, with their size inversely proportional to variance and solid lines representing 95% confidence intervals (CI). ORs were calculated using inverse variance weighted methods from summary GWAS data, and heterogeneity using Cochran Q.