**A genetic model of ivabradine recapitulates results from randomized clinical trials**

Marc-André Legault B.Sc., Johanna Sandoval M.Sc., Sylvie Provost M.Sc., Amina Barhdadi Ph.D., Louis-Philippe Lemieux Perreault Ph.D., Sonia Shah Ph.D., R.Thomas Lumbers Ph.D., Simon de Denus B. Pharm., Ph.D., Benoit Tyl M.D., Jean-Claude Tardif M.D., Marie-Pierre Dubé Ph.D.

# Supplementary Methods

## Additional information on UK Biobank variable selection

Phenotype data based on a touchscreen-based questionnaire followed by a verbal interview with a trained nurse was gathered. Hospitalization records are also available through linkage to the Health Episode Statistics (HES). For this project we used data from hospitalization episodes between the beginning of the HES linkage (April 1st 1997) and the last available date for the current data release (March 1st 2016). The date and cause of death were also available from death records made available through linkage to the National Health Services records for England, Scotland and Wales. We defined clinically relevant variables based on combinations of self-reported diseases, operation codes and hospitalization or death record ICD9/ICD10 codes. For the definition of most variables, self-reported diseases were included. However, for myocardial infarction we noticed that many self-reported events were unsupported by HES data even though they occurred within the time period of the HES linkage. We used the baseline resting heart rate measurement (variable #102), prioritizing the manual reading (variable #95) if available and taking the average value if many readings were available. We used age at recruitment defined in variable #21022 and sex in variable #31. Many of the self-reported myocardial infarction events co-occurred with ICD10 codes for related but distinct disorders such as I25.1 (atherosclerotic heart disease), I20.0 (unstable angina) or R07.4 (chest pain) without diagnostic codes for myocardial infarction, suggesting ischemic disease without a myocardial infarction event. For this reason, we ignored self-reported events for myocardial infarction as well as for angina, unstable angina and coronary artery disease as participants may incorrectly report them. For prospective analyses we used time from first baseline assessment centre visit (extracted from variable #53) in years. The censure date was defined as the date of death or date of end of follow up. The end of follow-up date was set to 2016-03-01 for England and Wales and to 2015-11-30 for Scotland as defined in the UK Biobank documentation (<https://biobank.ctsu.ox.ac.uk/~bbdatan/death_cancer_report_Sept16.pdf>). Individuals were assigned to countries based on the location of the UK Biobank assessment centre visited for the baseline visit.

## UK Biobank additional genetic quality controls

All UK Biobank participants were previously genotyped using two similar arrays, the UK BiLEVE Axiom Array and the UK Biobank Axiom Array and genome-wide imputation was conducted using the Haplotype Reference Consortium as the main reference panel. Additional genetic quality control was done using pyGenClean version 1.8.3 1. Variants or individuals with more than 2% missing genotypes (per sample and per variant, respectively) were filtered out. The self-reported sex and the genetic sex based on sexual chromosome was compared and individuals with discrepancies or with aneuploidies were removed from the analysis. We only considered individuals of European descent for this study as they represent the major population in the UK Biobank. We used the computed principal components from the UK Biobank and defined a region in principal components space using individuals identified as “white British ancestry” as a reference population 2. To avoid including related individuals, we used the kinship estimates from the UK Biobank and randomly selected an individual for pairs with a kinship coefficient > 0.0884 corresponding to individuals with 2nd degree relationships or less 2. The resulting post QC dataset included 413,083 individuals. There were 1,165 available imputed variants located at the *HCN4* gene region (chr15:73,612,200-73,661,605 ± 200kb padding) with a MAF above 1% and that were bi-allelic.

## External summary statistics

The CARDIoGRAMplusC4D consortium published a 1000 Genomes based meta-analysis of myocardial infarction and CAD of 60,810 CAD cases 3 and a more recent meta-analysis that adds the UK Biobank and the MIGen / CARDIOGRAM exome chip study 4. The main data release was based on the UK Biobank “soft” CAD definition that includes self-reported chronic ischemic heart disease and angina patients as well as the cases for the “hard” CAD definition of previous myocardial infarction or revascularization.

Two recent GWAS of atrial fibrillation were used. The first dataset was published by Nielsen *et al.* 5 and was based on 60,620 atrial fibrillation patients of European ancestry from six studies. The second dataset was published by Roselli *et al .*6 and included 65,446 atrial fibrillation patients predominantly of European ancestry. This study conducted a trans-ethnic association analysis that also included participants of Japanese (12.5%), African American (1.3%) and Brazilian and Hispanic (1.3%) populations.

For heart failure, we summary association statistics from the HERMES case-control consortium including 47,309 cases and 910,014 controls 7.

Finally, for stroke, we obtained summary associated statistics from the MEGASTROKE consortium who conducted GWAS for many stroke subtypes including ischemic stroke, large artery stroke, cardioembolic stroke and small vessel strokes. There were two available datasets, one based on European individuals including 40,585 stroke cases and a trans-ethnic GWAS dataset including 67,162 stroke cases8.

## Mendelian randomization

Mendelian randomization is a technique to infer the causal effect of an exposure such as heart rate on an outcome such as atrial fibrillation. Because genetic variants are randomly assigned at birth and are generally not influenced by the environment, they represent an unconfounded way of modulating the exposure that may be suitable for causal inference. Unfortunately, MR is susceptible to other methodological issues some of which can be accounted for in more sophisticated models. A commonly used model that assumes the classical instrument variable conditions is the inverse variance weighted (IVW) approach 9. In this approach, the causal estimates for multiple variants are averaged with weights corresponding to the precision of the causal estimates without accounting for the possibility of invalid genetic instruments. The MR-Egger test is a similar approach that extends the IVW by allowing an intercept term that corresponds to directional pleiotropy 10. Directional pleiotropy is detectable if the mean direct effect of genetic variants on the outcome is different from zero. The MR-Egger approach relaxes the exclusion-restriction (or IV3) assumption that requires genetic variants to be conditionally independent of the outcome given the exposure and covariates. Instead, MR-Egger requires the instrument strength independent of direct effect (InSIDE) assumption which stipulates that the effect of genetic variants on the exposure should be independent from the direct effects. MR-Egger is useful method, but the InSIDE assumption is hard to verify and is likely to fail in many biologically plausible scenarios leading to possibly biased estimates 11. Moreover, the IVW and MR-Egger are greatly influenced by outlier variants as they are based on a linear regression of individual variant effects. A more recent set of MR methods further relax these assumptions and rely on the hypothesis that the largest set of variants with homogeneous effects is likely to represent the set of valid instrument variables. The contamination mixture method uses a mixture model to assign variants to a distribution of valid causal effects and a distribution of noisy variants with a null expected causal effect and a large variance 12. Similarly, the Mendelian Randomization Pleiotropy RESidual Sum and Outlier (MR-PRESSO) iteratively eliminates variants whose effects are outliers when compared to the others until a homogeneous signal remains corresponding to the estimated causal effect 13. These methods are interesting because they are less dependent on hard to verify assumptions, they are robust even when some invalid instruments are included, and they allow variants to be individually tested for their heterogeneity and further investigated.

For MR with the heart rate GRS, we used the two-stage method which is akin to the previously described IVW method, but uses individual level data 14. Heart rate expressed in units of 10 bpm reduction was predicted based on a fitted model including the GRS as a continuous variable and covariates (age, sex and PCs). The predicted heart rates were then used in the second stage to estimate the causal effect on CAD, heart failure and atrial fibrillation using logistic regression adjusted for the same covariates. The standard errors for the causal effect were estimated using the percentile method based on 5,000 bootstrap resamples which is an empirical way of estimating standard errors without assuming their distribution. This approach does not account for violations of the instrument variable assumptions, but the estimates rely on a strong genetic instrument whose effect is closer to pharmacological effects than what is observed using individual genetic variants making extrapolations less problematic.

Analyses were performed with the “MendelianRandomization” R package (<https://cran.r-project.org/web/packages/MendelianRandomization/>) and MR-PRESSO (<https://github.com/rondolab/MR-PRESSO>).

## Bi-directional MR

Bi-directional MR is a method used to infer the direction of the causality between two traits 15. Genetic variants are ascertained for their association with the first trait and used as an instrument variable to estimate its effect on the second trait. The procedure is then repeated with instruments for the second trait (with the first trait as the outcome). The estimated causal effects can then be compared, and the direction of effect can be elucidated. Here, we were interested in the causality between atrial fibrillation and heart failure, CAD and myocardial infarction. We used 11 genome-wide significant independent variants from the HERMES consortium and estimated the causal effect of heart failure on atrial fibrillation using these variants. The previously described MR methods were used for the bi-directional analysis as well. To test the effect of atrial fibrillation on heart failure, we selected uncorrelated variants that reached genome-wide significance in the Nielsen *et al.* study 5. The selection was done using *grstools* and the 1000 Genomes Phase III Europeans as a reference panel for linkage disequilibrium. Variants with a linkage disequilibrium r2 above 0.15 were clumped together, keeping the most significant variant. The selection was stopped when no genome-wide significant variant remained leaving 152 variants to be used as genetic instruments.

The variants for the myocardial infarction genetic instrument were selected as for the atrial fibrillation instrument but using the CARDIoGRAMplusC4D summary statistics. A total of 31 variants were selected all of which were available in the atrial fibrillation summary statistics. For the CAD instrument, we used the 71 variants comprising the genetic risk score described in Verweij *et al.* 16.

## Genetic risk score for heart rate

For the construction of the heart rate Genetic Risk Score (GRS), we used the 64 genome-wide significant heart rate associated SNPs from Eppinga *et al.* 17 To ensure that there was no strand confusion due to ambiguous alleles (i.e. A/T and G/C SNPs), we compared the observed allele frequencies to the expected distribution using the 1000 Genomes Phase III Europeans as a reference panel. If the observed allele frequency fell in a 95% Clopper-Pearson confidence interval around the reference panel frequency estimate, the strand was considered to be validated and the SNP was used as-is. When variants had a minor allele frequency (MAF) above 40%, an unambiguous SNP in LD was automatically selected to avoid strand confusion as the frequencies were close to 50%. A total of 8 of the 10 ambiguous variants were validated based on allele frequency. The other 2 variants had a MAF above 40% and were replaced by tag SNPs in LD. The variant rs13165531 was replaced by rs6887889 (r2=1) and rs3951016 was replaced by rs9401060 (r2=0.91). GRS weights were adjusted by multiplying the r2 value for both SNPs. The final set of variants and their corresponding weights are shown in Supplementary Table 5. The software used to compute the genetic risk scores is publicly available at <https://github.com/legaultmarc/grstools>.

# Supplementary Tables

**Supplementary Table 1.** Summary ofivabradinecardiovascularoutcomes trials.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Study / Intervention** | **Patient population** | **Main cardiovascular exclusions** | **Primary efficacy endpoint** | **Results** |
| SHIFT 18  2.5-7.5 mg bid versus placebo | * Resting heart rate >=70 bpm * Symptomatic chronic HF for at least 4 weeks * LVEF <= 35% * Recent hospitalization for worsening HF | * HF caused by congenital heart disease or primary severe valvular disease * Recent MI, atrial fibrillation or flutter | CV death or hospitalization for worsening HF | HR 0.82 (0.75, 0.90) p<0.0001. Effect driven by hospitalization for worsening HF.  Adverse events:  Atrial fibrillation 9% in ivabradine vs 8% in placebo (p=0.012). |
| BEAUTIFUL 19  5-7.5 mg bid versus placebo | * Stable CAD (previous MI, PCI/CABG or angiographic evidence of obstruction of at least 50%) * LVEF <= 40% * Resting heart rate >= 65 bpm | * Recent revascularization or MI * Recent stroke or TIA * NYHA class IV HF * Implanted pacemaker, cardioverter or defibrillator * Valvular disease, SSS, sinoatrial block, severe hypertension | CV death or MI or hospitalization for worsening HF | Primary endpoint was not significant (p=0.94).  In a prespecified subgroup with baseline heart rate >= 70bpm, there was a reduction for ischemic endpoints, which led to conduct SIGNIFY. |
| SIGNIFY 20  5-10 mg bid (treat to target of 55-60 bpm) versus placebo | * Stable CAD * No heart failure * Resting heart rate >= 70 bpm | * Patients with LVEF <= 40% | CV death or MI | Primary endpoint and its individual components were not significant  In patients with severe angina (>= CCS II) there was an increase of the primary endpoint with ivabradine HR 1.18 (1.03, 1.35) p=0.02. Individuals in this group also found better anginal symptom improvement on ivabradine  Atrial fibrillation was 5.3% in ivabradine vs 3.8% in placebo |

CABG, coronary artery bypass graft; CAD, coronary artery disease; CV, cardiovascular; HF, heart failure; HR, hazard ratio; LVEF, left ventricular ejection fraction; MI, myocardial infarction; PCI, percutaneous coronary intervention; TIA, transient ischemic attack

**Supplementary Table 2**. Self-reported, hospitalization (ICD10) and operation (OPCS) codes used to define clinical variables based on the UK Biobank available data.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Included codes** | | |
| **Variable** | Self-reported disease (variable #20002) | ICD9/10 for HES primary or secondary hospitalization codes or primary cause of death | Operations (OPCS) |
| **Angina** | - | ICD9: 413  ICD10: I20 |  |
| **Unstable angina** | - | ICD10: I20.0 (as the primary hospitalization code only to ensure acute event) |  |
| **Myocardial infarction** | - | ICD9: 410, 412, 411.0, 429.79  ICD10: I21, I22, I23, I25.2 |  |
| **Coronary artery disease** | - | ICD9: 410-414 (except 414.1)  ICD10: I20-I25 | K40, K41, K42, K43, K44, K45, K46, K49, K50, K75 |
| **Stroke (any)** | 1583, 1081, 1086, 1491 | ICD9: 430, 431, 434, 436  ICD10: I60, I61, I63, I64 |  |
| **Stroke - Ischemic** | 1583 | ICD9: 434, 436  ICD10: I63, I64 |  |
| **Atrial fibrillation** | 1471 | ICD9: 427.3  ICD10: I48 |  |
| **Heart failure** | 1076 | ICD9: 428, 425  ICD10: I50, I42 |  |

**Supplementary Table 3**. Results from the NHGRI-EBI GWAS catalog mapped to the *HCN4* gene 21. LD with the lead independent heart-rate associated variants in the UK Biobank identified through forward stepwise conditional analysis are reported.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Variant and risk allele** | **Beta** | **P value** | **Trait** | **Study reference** | **LD with rs8038766 a** | **LD with rs3743496 a** |
| rs7173389-T | 0.539 | 1.00E-32 | Resting heart rate | 17 | 1 | 0.022 |
| rs7173389-A | 0.528 | 2.00E-09 | Resting heart rate | 22 | 1 | 0.022 |
| rs8040516-T | 0.219 | 3.00E-06 | Nickel levels | 23 | 0.001 | 0.532 |
| rs74022964-T | 0.113 | 4.00E-36 | Atrial fibrillation | 5 | 0.956 | 0.022 |
| rs478438-G | 0.019 | 2.00E-09 | Heel bone mineral density | 24 | 0.237 | 0.050 |
| rs142859932-G | 0.487 | 3.00E-06 | Post bronchodilator FEV1 | 25 | - | - |
| rs16957893-C | 1.72 | 2.00E-08 | Cold medicine-related Stevens-Johnson syndrome/toxic epidermal necrolysis (SJS/TEN) with severe ocular complications | 26 | 0.025 | 0.004 |
| rs7164883-G | 0.17 | 3.00E-17 | Atrial fibrillation | 27 | 0.978 | 0.022 |
| rs7183206-A | 0.12 | 8.00E-12 | Atrial fibrillation | 28 | 0.956 | 0.018 |
| rs2680344-A | 0.024 | 5.00E-11 | Heart rate variability traits (SDNN) | 29 | 0.575 | 0 |
| rs2680344-A | 0.024 | 3.00E-11 | Heart rate variability traits (SDNN) | 29 | 0.575 | 0 |
| rs2680344-A | 0.032 | 1.00E-10 | Heart rate variability traits (RMSSD) | 29 | 0.575 | 0 |
| rs2680344-A | 0.046 | 3.00E-06 | Heart rate variability traits (pvRSA/HF) | 29 | 0.575 | 0 |
| rs4489968-T | 0.513 | 4.00E-20 | Heart rate | 30 | 1 | 0.022 |
| rs7172038-G | 0.10 | 2.00E-27 | Atrial fibrillation | 6 | 0.993 | 0.022 |
| rs74022964-T | 0.10 | 1.00E-27 | Atrial fibrillation | 6 | 0.956 | 0.022 |
| rs11072405-A | 0.021 | 1.00E-08 | Waist-to-hip ratio adjusted for BMI | 31 | 0.007 | 0.621 |
| rs11072405-A |  | 5.00E-07 | Waist-to-hip ratio adjusted for BMI×sex×age interaction (4df test) | 31 | 0.007 | 0.621 |

**a** LD measurements (r2) are for individuals of European (EUR) descent from the 1000 Genomes Project (phase 3) and were obtained using LDlink 32.

**Supplementary Table** **4**. Variants and weights used for the computation of the heart rate GRS.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Variant** | **Chr.** | **Position** | **Reference allele** | **Risk allele** | **P value** | **Effect** |
| rs145358377 | 1 | 6272136 | G | GA | 1.94E-11 | -0.259 |
| rs272564 | 1 | 45012273 | A | C | 4.51E-21 | 0.351 |
| rs2152735 | 1 | 87893132 | G | A | 7.23E-18 | -0.306 |
| rs41317993 | 1 | 207961732 | G | A | 5.42E-31 | 0.630 |
| rs11454451 | 1 | 217722890 | C | CT | 1.29E-11 | 0.256 |
| rs1260326 | 2 | 27730940 | T | C | 4.29E-16 | -0.275 |
| rs12713404 | 2 | 60006705 | G | T | 9.33E-09 | -0.199 |
| rs564190295 | 2 | 175547672 | G | GCCGCCGCCCCC | 4.95E-10 | -0.355 |
| rs151041685 | 2 | 179725237 | G | T | 7.86E-75 | 1.061 |
| rs62172372 | 2 | 188242369 | A | G | 5.99E-16 | 0.337 |
| rs907683 | 2 | 220299541 | G | T | 1.02E-20 | -0.334 |
| rs4608502 | 2 | 228134155 | T | C | 1.85E-12 | 0.249 |
| rs13002735 | 2 | 232268884 | A | C | 1.29E-17 | -0.331 |
| rs41312411 | 3 | 38621237 | C | G | 1.34E-11 | -0.320 |
| rs3749237 | 3 | 49770032 | G | A | 3.09E-13 | 0.258 |
| rs2358740 | 3 | 53455569 | G | T | 3.58E-09 | -0.208 |
| rs1483890 | 3 | 69410725 | A | G | 2.54E-15 | 0.284 |
| rs11920570 | 3 | 122090102 | G | A | 5.18E-13 | 0.268 |
| rs7612445 | 3 | 179172979 | G | T | 2.41E-24 | -0.428 |
| rs12501032 | 4 | 23951018 | C | G | 1.83E-15 | 0.288 |
| rs6845865 | 4 | 148974602 | T | C | 2.25E-14 | -0.342 |
| rs6887889 (tag for rs13165531) | 5 | 30893205 | T | G | 3.57E-09 | -0.221 |
| rs1468333 | 5 | 137552970 | T | C | 9.53E-14 | -0.255 |
| rs4868243 | 5 | 172643118 | G | A | 4.08E-16 | -0.361 |
| rs236349 | 6 | 36820565 | A | G | 1.01E-15 | 0.281 |
| rs9401060 (tag for rs3951016) | 6 | 118561348 | A | G | 4.04E-33 | 0.473 |
| rs1320761 | 6 | 122168138 | C | T | 1.22E-64 | 0.902 |
| rs58437978 | 7 | 35258277 | T | C | 2.61E-12 | -0.240 |
| rs180239 | 7 | 93550415 | G | C | 4.54E-21 | -0.326 |
| rs17881696 | 7 | 100493359 | G | A | 1.18E-41 | 0.578 |
| rs41748 | 7 | 116446573 | T | G | 7.14E-09 | -0.193 |
| rs11563648 | 7 | 126970046 | G | C | 4.42E-10 | -0.231 |
| rs138186803 | 7 | 130965408 | AT | A | 1.27E-16 | -0.333 |
| rs73158705 | 7 | 136576100 | A | G | 2.81E-18 | 0.393 |
| rs56233017 | 8 | 144981488 | G | A | 1.09E-15 | -0.666 |
| rs10739663 | 9 | 128278739 | A | G | 9.62E-16 | -0.266 |
| rs12576326 | 11 | 44980383 | A | G | 1.20E-12 | 0.253 |
| rs174536 | 11 | 61551927 | A | C | 1.65E-30 | 0.399 |
| rs75190942 | 11 | 128764571 | C | A | 1.19E-16 | -0.496 |
| rs2283274 | 12 | 2184466 | G | C | 7.21E-20 | -0.405 |
| rs10841486 | 12 | 20472202 | T | C | 2.98E-09 | -0.238 |
| rs4963772 | 12 | 24758480 | G | A | 3.23E-53 | -0.714 |
| rs1050288 | 12 | 27955296 | C | T | 2.74E-09 | -0.213 |
| rs1994135 | 12 | 33682405 | T | C | 7.19E-34 | 0.400 |
| rs10880689 | 12 | 37930102 | A | G | 8.10E-10 | 0.208 |
| rs867400 | 12 | 64976850 | T | C | 4.58E-19 | 0.298 |
| rs12579753 | 12 | 82219376 | C | T | 4.81E-10 | -0.246 |
| rs12889267 | 14 | 21542766 | A | G | 3.61E-20 | 0.416 |
| rs422068 | 14 | 23864804 | T | C | 1.52E-100 | 0.731 |
| rs17180489 | 14 | 72885471 | G | C | 9.15E-19 | -0.490 |
| rs1549118 | 14 | 78379684 | C | T | 4.67E-08 | 0.200 |
| rs17201923 | 14 | 85796564 | A | G | 6.55E-29 | -0.41 |
| rs4900069 | 14 | 91583373 | A | C | 5.38E-09 | 0.200 |
| rs7173389 | 15 | 73663903 | A | T | 1.31E-32 | -0.539 |
| rs3915499 | 16 | 15910743 | G | A | 1.24E-17 | 0.303 |
| rs7194801 | 16 | 65286870 | T | C | 3.58E-18 | -0.291 |
| rs79121763 | 17 | 15195279 | C | T | 7.17E-14 | -0.471 |
| rs11083258 | 18 | 25766218 | A | C | 5.51E-10 | -0.276 |
| rs61735998 | 18 | 34289285 | G | T | 2.06E-14 | -0.834 |
| rs16974196 | 19 | 40833470 | G | A | 1.11E-11 | 0.244 |
| rs12721051 | 19 | 45422160 | C | G | 5.23E-11 | -0.287 |
| rs6123471 | 20 | 36840156 | T | C | 6.63E-72 | -0.595 |
| rs17265513 | 20 | 39832628 | T | C | 1.12E-08 | 0.240 |
| rs2076028 | 22 | 39150450 | G | A | 5.45E-16 | -0.295 |

Variants that were substituted by a tag SNP are identified in parenthesis and the p-value from the original GWAS is reported. For these variants, the weight is computed as the effect of the original variant weighted by the LD in Europeans (r2 × β). Chromosomal positions for GRCh37. GRS, genetic risk score.

**Supplementary Table 5.** MR estimates based on 64 heart-rate associated variants and their effect on outcomes in the UK Biobank. Reported effects are per genetically predicted s.d. decrease in heart rate (1 s.d. is 11.1 bpm in the UK Biobank).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  | **Intercept\* (MR-Egger only)** | | **Causal Estimate** | |
| **Exposure** | **Outcome** | **Method** | **Estimate (95% CI)** | **p-value** | **OR (95% CI)** | **p-value** |
| Heart rate reduction (1 s.d. or 11.1 bpm) | Atrial fibrillation | IVW |  |  | 1.25 (0.97, 1.62) | 0.083 |
| MR-Egger | 0.015 (-0.004, 0.034) | 0.13 | 0.85 (0.48, 1.49) | 0.57 |
| Contamination mixture |  |  | 1.54 (1.22, 1.79) | - |
| MR-PRESSO |  |  | 1.54 (1.34, 1.78) | 1.3E-07 |
| Heart failure | IVW |  |  | 1.02 (0.87, 1.21) | 0.77 |
| MR-Egger | 0.007 (-0.005, 0.020) | 0.60 | 0.84 (0.58, 1.21) | 0.35 |
| Contamination mixture |  |  | 1.16 (0.92, 1.43) | - |
| MR-PRESSO † |  |  | - | - |
| Coronary artery disease | IVW |  |  | 1.01 (0.89, 1.15) | 0.86 |
| MR-Egger | 0.004 (-0.006, 0.014) | 0.53 | 0.91 (0.68, 1.23) | 0.44 |
| Contamination mixture |  |  | 0.97 (0.87, 1.09) | - |
| MR-PRESSO |  |  | 1.01 (0.91, 1.12) | 0.86 |

\* The MR-Egger estimate intercepts represent directional pleiotropy and are not converted to the OR scale because they do not have an intuitive interpretation on this scale.

† The MR-PRESSO did not provide adjusted estimates as the global test did not detect significant pleiotropy (p=0.31).

**Supplementary Table 6.** MR estimates based on the effect of 64 heart-rate associated variants in external summary statistics from large GWAS consortia. Reported effects are per genetically predicted s.d. decrease in heart rate (1 s.d. is 11.1 bpm).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  | **Intercept\*  (MR-Egger only)** | | **Causal estimate** | |
| **Exposure** | **Dataset  (for the outcome)** | **Outcome** | **Method** | **Estimate  (95% CI)** | **P value** | **OR (95% CI)** | **p-value** |
| Heart rate  reduction  (1 s.d. or  11.1 bpm) | Nielsen et al.5 | Atrial fibrillation | IVW |  |  | 1.24 (1.00, 1.53) | 0.047 |
| MR-Egger | 0.167  (-0.056, 0.378) | 0.14 | 0.85 (0.49, 1.46) | 0.55 |
| Contamination mixture |  |  | 1.56 (1.40, 1.56) | - |
| MR-PRESSO |  |  | 1.36 (1.24, 1.49) | 6.3E-08 |
| HERMES 7 case/control | Heart failure | IVW |  |  | 1.03 (0.97, 1.12) | 0.34 |
| MR-Egger | -0.011  (-0.089, 0.067) | 0.79 | 1.06 (0.88, 1.29) | 0.55 |
| Contamination mixture |  |  | 1.12 (1.00, 1.12) | - |
| MR-PRESSO |  |  | - | - |
| CARDIoGRAMplusC4D + UKB SOFT + MiGen 3, 4 | Coronary artery disease | IVW |  |  | 1.07 (0.96, 1.19) | 0.27 |
| MR-Egger | 0.044  (-0.078, 0.167) | 0.45 | 0.96 (0.71, 1.29) | 0.77 |
| Contamination mixture |  |  | 1.00 (1.00, 1.12) | - |
| MR-PRESSO |  |  | 1.04 (0.95, 1.14) | 0.36 |

\* The MR-Egger estimate intercepts represent directional pleiotropy and are not converted to the OR scale because they do not have an intuitive interpretation on this scale.

**Supplementary Table 7.** Bi-directional MR estimates using summary GWAS results for IVW and MR models more robust to invalid instruments.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  | **Intercept (MR-Egger only)** | | **Causal estimate** | |
| **Exposure** | **Outcome** | **Method** | **Estimate (95% CI)** | **P-value** | **OR (95% CI)** | **P-value** |
| Atrial fibrillation  (152 variants) | Heart failure | IVW |  |  | 1.23 (1.20, 1.27) | 3.7E-52 |
| MR-Egger | 0.001 (-0.004, 0.006) | 0.77 | 1.22 (1.15, 1.30) | 8.5E-10 |
| Contamination mixture |  |  | 1.25 (1.22, 1.27) | - |
| MR-PRESSO |  |  | 1.23 (1.20, 1.26) | 3.0E-33 |
| Atrial fibrillation  (152 variants) | Coronary artery disease | IVW |  |  | 1.00 (0.98, 1.03) | 0.76 |
| MR-Egger | 0.004 (-0.001, 0.009) | 0.09 | 0.96 (0.90, 1.02) | 0.17 |
| Contamination mixture |  |  | 1.01 (0.99, 1.03) | - |
| MR-PRESSO a |  |  | - | - |
| Atrial fibrillation  (152 variants) | Myocardial infarction | IVW |  |  | 0.98 (0.95, 1.02) | 0.30 |
| MR-Egger | 0.003 (-0.003, 0.009) | 0.32 | 0.95 (0.89, 1.02) | 0.18 |
| Contamination mixture |  |  | 0.98 (0.95, 1.01) | - |
| MR-PRESSO |  |  | 0.98 (0.95, 1.01) | 0.23 |
| Heart failure  (11 variants) | Atrial Fibrillation | IVW |  |  | 1.45 (1.11, 1.90) | 0.0067 |
| MR-Egger | 0.038 (0.014, 0.063) | 0.002 | 1.21 (0.97, 1.52) | 0.094 |
| Contamination mixture |  |  | 6.82 (4.62, 9.20) | - |
| MR-PRESSO |  |  | 1.94 (1.66, 2.25) | 0.07 |
| Coronary artery disease  (68 variants) | Atrial Fibrillation | IVW |  |  | 1.15 (1.11, 1.21) | 1.7E-10 |
| MR-Egger | 0.003 (-0.005, 0.011) | 0.47 | 1.12 (1.01, 1.23) | 0.03 |
| Contamination mixture |  |  | 1.17 (1.13, 1.21) | - |
| MR-PRESSO |  |  | 1.15 (1.11, 1.19) | 2.5E-10 |
| Myocardial infarction  (31 variants) | Atrial Fibrillation | IVW |  |  | 1.11 (1.06, 1.16) | 1.3E-05 |
| MR-Egger | -0.010 (-0.021, 0.001) | 0.07 | 1.22 (1.09, 1.36) | 6.6E-04 |
| Contamination mixture |  |  | 1.15 (1.12, 1.19) | - |
| MR-PRESSO |  |  | 1.11 (1.07, 1.16) | 1.4E-05 |

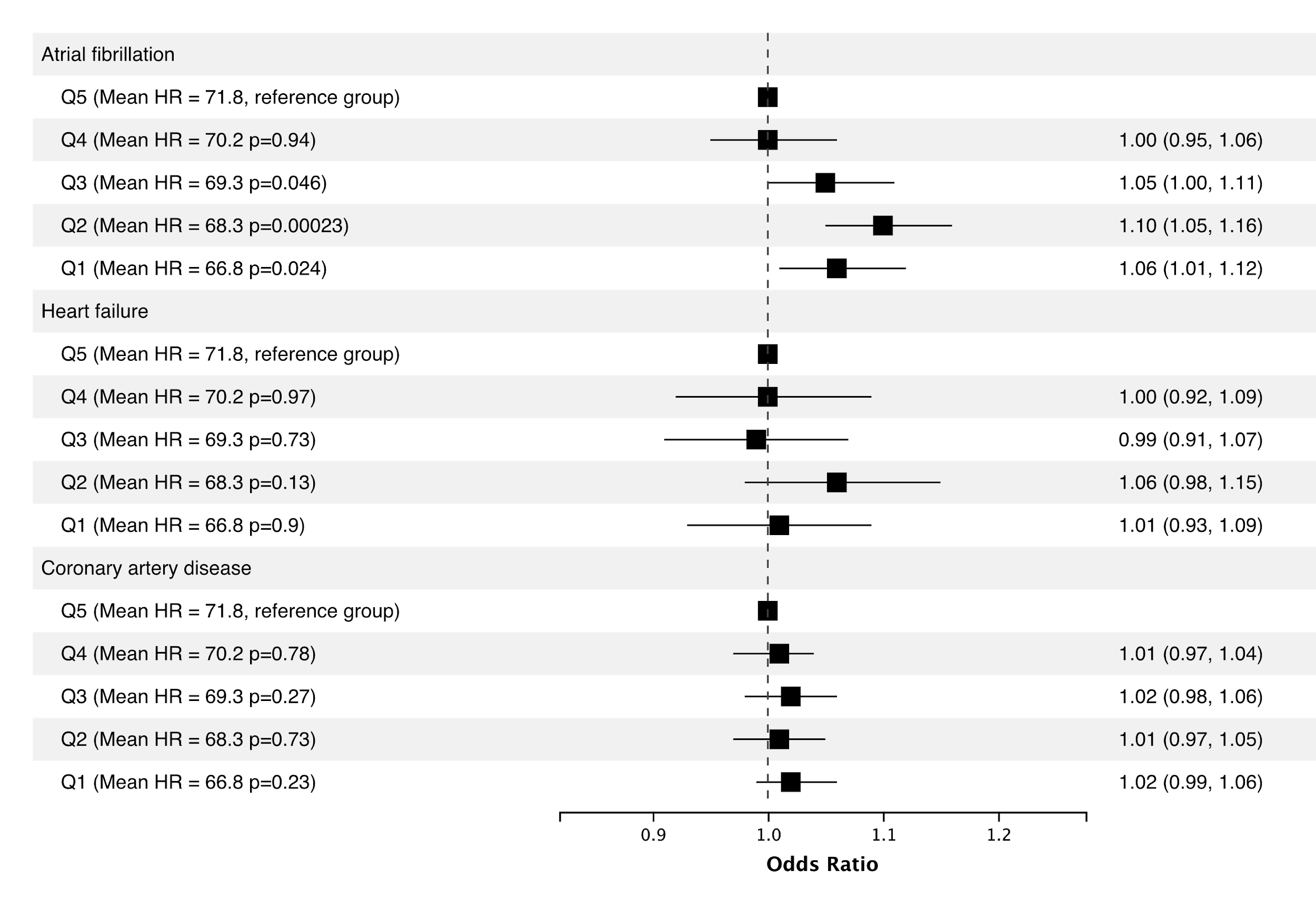
Summary statistics for atrial fibrillation taken from Nielsen et al. 5 for myocardial infarction and CAD from CARDIoGRAMplusC4D and CARDIoGRAMplusC4D + UKB SOFT + MiGen for CAD 3, 4, and for heart failure from the HERMES consortium7. The contamination mixture model does not provide P-values. IVW, inverse-variance weighted; MR, Mendelian randomization; OR, odds ratio.

a MR-PRESSO did not provide adjusted estimates as it detected no outliers.

# Supplementary Figures

|  |
| --- |
|  |
|  |
|  |

**Online Figure 1.** Results from the stepwise forward regression using 1,165 variants in the *HCN4* region tested for association with heart rate in the UK Biobank, and adjusted for age, sex and the first 10 principal components. The lead variant identified was rs8038766, β = -0.574 (95% CI -0.640, -0.509), P-value=2.76×10-66 (results shown in the top panel). For the second stage, we conditioned on the lead variant from Stage 1 (rs8038766) and repeated the analysis. The lead variant identified in Stage 2 was rs3743496, β =-0.297 (95% CI -0.349, -0.246), P-value =3.96×10-30 (results shown in the second panel). We repeated the analysis again conditioning on the lead variants from both previous stages (rs8038766 and rs3743496), but no additional variant crossed the genome-wide significance threshold (third panel). The first y-axis shows the negative log10 of P-values, the second y-axis shows the recombination rate from HapMap reference samples (black line). Genes are displayed below the x-axis from Ensembl (build37), the degree of linkage disequilibrium (r2) of each genetic variant with the lead variant.



**Supplementary Figure** **2**. Effect of heart rate genetic risk score quintiles on atrial fibrillation, heart failure and coronary artery disease in the UK biobank dataset. For every outcome, the highest heart rate group (5th quintile) is used as the reference group and the reported odds ratios are adjusted for age, sex and the first 10 principal components.

# References for Supplementary Material

1. Lemieux Perreault LP, Provost S, Legault MA, Barhdadi A ,Dube MP. pyGenClean: efficient tool for genetic data clean up before association testing. Bioinformatics. 2013;29:1704-5.

2. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping and genomic data. Nature. 2018;562:203-209.

3. Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. Nat Genet. 2015;47:1121-1130.

4. Nelson CP, Goel A, Butterworth AS, Kanoni S, Webb TR, Marouli E, et al. Association analyses based on false discovery rate implicate new loci for coronary artery disease. Nat Genet. 2017;49:1385-1391.

5. Nielsen JB, Thorolfsdottir RB, Fritsche LG, Zhou W, Skov MW, Graham SE, et al. Biobank-driven genomic discovery yields new insight into atrial fibrillation biology. Nat Genet. 2018;50:1234-1239.

6. Roselli C, Chaffin MD, Weng LC, Aeschbacher S, Ahlberg G, Albert CM, et al. Multi-ethnic genome-wide association study for atrial fibrillation. Nat Genet. 2018;50:1225-1233.

7. Shah S, Henry A, Roselli C, Lin H, Sveinbjornsson G, Fatemifar G, et al. Genome-wide association and Mendelian randomisation analysis provide insights into the pathogenesis of heart failure. Nat Commun. 2020;11:163.

8. Malik R, Chauhan G, Traylor M, Sargurupremraj M, Okada Y, Mishra A, et al. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. Nat Genet. 2018;50:524-537.

9. Burgess S, Butterworth A ,Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol. 2013;37:658-65.

10. Bowden J, Davey Smith G ,Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol. 2015;44:512-25.

11. Burgess S ,Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. Eur J Epidemiol. 2017;32:377-389.

12. Burgess S, Foley CN, Allara E, Staley JR ,Howson JMM. A robust and efficient method for Mendelian randomization with hundreds of genetic variants. Nat Commun. 2020;11:376.

13. 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.

14. Burgess S, Small DS ,Thompson SG. A review of instrumental variable estimators for Mendelian randomization. Stat Methods Med Res. 2017;26:2333-2355.

15. Brower MA, Hai Y, Jones MR, Guo X, Chen YI, Rotter JI, et al. Bidirectional Mendelian randomization to explore the causal relationships between body mass index and polycystic ovary syndrome. Hum Reprod. 2019;34:127-136.

16. Verweij N, Eppinga RN, Hagemeijer Y ,van der Harst P. Identification of 15 novel risk loci for coronary artery disease and genetic risk of recurrent events, atrial fibrillation and heart failure. Sci Rep. 2017;7:2761.

17. Eppinga RN, Hagemeijer Y, Burgess S, Hinds DA, Stefansson K, Gudbjartsson DF, et al. Identification of genomic loci associated with resting heart rate and shared genetic predictors with all-cause mortality. Nat Genet. 2016;48:1557-1563.

18. Swedberg K, Komajda M, Bohm M, Borer JS, Ford I, Dubost-Brama A, et al. Ivabradine and outcomes in chronic heart failure (SHIFT): a randomised placebo-controlled study. Lancet. 2010;376:875-85.

19. Fox K, Ford I, Steg PG, Tendera M, Ferrari R ,Investigators B. Ivabradine for patients with stable coronary artery disease and left-ventricular systolic dysfunction (BEAUTIFUL): a randomised, double-blind, placebo-controlled trial. Lancet. 2008;372:807-16.

20. Fox K, Ford I, Steg PG, Tardif JC, Tendera M, Ferrari R, et al. Ivabradine in stable coronary artery disease without clinical heart failure. N Engl J Med. 2014;371:1091-9.

21. Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic Acids Res. 2019;47:D1005-D1012.

22. Ramirez J, Duijvenboden SV, Ntalla I, Mifsud B, Warren HR, Tzanis E, et al. Thirty loci identified for heart rate response to exercise and recovery implicate autonomic nervous system. Nat Commun. 2018;9:1947.

23. Ng E, Lind PM, Lindgren C, Ingelsson E, Mahajan A, Morris A, et al. Genome-wide association study of toxic metals and trace elements reveals novel associations. Hum Mol Genet. 2015;24:4739-45.

24. Morris JA, Kemp JP, Youlten SE, Laurent L, Logan JG, Chai RC, et al. An atlas of genetic influences on osteoporosis in humans and mice. Nat Genet. 2019;51:258-266.

25. Lutz SM, Cho MH, Young K, Hersh CP, Castaldi PJ, McDonald ML, et al. A genome-wide association study identifies risk loci for spirometric measures among smokers of European and African ancestry. BMC Genet. 2015;16:138.

26. Ueta M, Sawai H, Shingaki R, Kawai Y, Sotozono C, Kojima K, et al. Genome-wide association study using the ethnicity-specific Japonica array: identification of new susceptibility loci for cold medicine-related Stevens-Johnson syndrome with severe ocular complications. J Hum Genet. 2017;62:485-489.

27. Ellinor PT, Lunetta KL, Albert CM, Glazer NL, Ritchie MD, Smith AV, et al. Meta-analysis identifies six new susceptibility loci for atrial fibrillation. Nat Genet. 2012;44:670-5.

28. Christophersen IE, Rienstra M, Roselli C, Yin X, Geelhoed B, Barnard J, et al. Large-scale analyses of common and rare variants identify 12 new loci associated with atrial fibrillation. Nat Genet. 2017;49:946-952.

29. Nolte IM, Munoz ML, Tragante V, Amare AT, Jansen R, Vaez A, et al. Genetic loci associated with heart rate variability and their effects on cardiac disease risk. Nat Commun. 2017;8:15805.

30. den Hoed M, Eijgelsheim M, Esko T, Brundel BJ, Peal DS, Evans DM, et al. Identification of heart rate-associated loci and their effects on cardiac conduction and rhythm disorders. Nat Genet. 2013;45:621-31.

31. Winkler TW, Justice AE, Graff M, Barata L, Feitosa MF, Chu S, et al. The Influence of Age and Sex on Genetic Associations with Adult Body Size and Shape: A Large-Scale Genome-Wide Interaction Study. PLoS genetics. 2015;11:e1005378.

32. Machiela MJ ,Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. Bioinformatics. 2015;31:3555-7.