

Title Page

Title

Genetic evidence for T-wave area from 12-lead electrocardiograms to monitor cardiovascular diseases in patients taking diabetes medications

Authors and affiliations

Mengling Qi¹, Haoyang Zhang², Xuehao Xiu¹, Dan He¹, David N. Cooper³, Yuanhao Yang⁵,
Huiying Zhao¹

¹Department of Medical Research Center, Sun Yat-Sen Memorial Hospital, Sun Yat-Sen University, Guangzhou 510006, China

²School of Data and Computer Science, Sun Yat-sen University, Guangzhou, China

³Institute of Medical Genetics, School of Medicine, Cardiff University, Heath Park, Cardiff CF14 4XN, UK

⁴Mater Research Institute, Translational Research Institute, Brisbane, QLD, Australia

Corresponding author

Huiying Zhao

zhaohy8@mail.sysu.edu.cn

1 **Abstract**

2 **Aims**

3 Many studies indicated use of diabetes medications can influence the electrocardiogram
4 (ECG), which remains the simplest and fastest tool for assessing cardiac functions. However,
5 few studies have explored the role of genetic factors in determining the relationship between
6 the use of diabetes medications and ECG trace characteristics (ETC).

7
8 **Methods**

9 Genome-wide association studies (GWAS) were performed for 168 ETCs extracted from the
10 12-lead ECGs of 42,340 Europeans in the UK Biobank. The genetic correlations, causal
11 relationships, and phenotypic relationships of these ETCs with medication usage, as well as
12 the risk of cardiovascular diseases (CVDs), were estimated by linkage disequilibrium score
13 regression (LDSC), Mendelian randomization (MR), and regression model, respectively.

14
15 **Results**

16 The GWAS identified 124 independent single nucleotide polymorphisms (SNPs) that were
17 study-wise and genome-wide significantly associated with at least one ETC. Regression
18 model and LDSC identified significant phenotypic and genetic correlations of T-wave area in
19 lead aVR (aVR_T-area) with usage of diabetes medications (ATC code: A10 drugs, and
20 metformin), and the risks of ischemic heart disease (IHD) and coronary atherosclerosis (CA).
21 MR analyses support a putative causal effect of the use of diabetes medications on decreasing
22 aVR_T-area, and on increasing risk of IHD and CA.

1

2 **Conclusion**

3 Patients taking diabetes medications are prone to have decreased aVR_T-area and an
4 increased risk of IHD and CA. The aVR_T-area is therefore a potential ECG marker for pre-
5 clinical prediction of IHD and CA in patients taking diabetes medications.

6

7 **Keywords**

8 Electrocardiogram; cardiovascular diseases; diabetes medication use; genome-wide
9 association study; genetic correlation; genetic causality

1 **Introduction**

2 It has been blithely assumed that glycemic control would be beneficial in reducing the risk of
3 cardiovascular disease (CVD). To this end, a large number of clinical trials have investigated
4 the potential use of diabetes medications in improving cardiac health. These studies have
5 demonstrated that many diabetes medications are not associated with a reduction in the rate
6 of heart failure(1, 2, 3, 4), and that rosiglitazone and pioglitazone are actually associated with
7 increasing the risk of heart failure(5, 6). The effects of the commonly used first line therapy,
8 metformin, on cardiovascular end points remain unclear(7).

9

10 Heart functions are usually measured by the electrocardiogram (ECG). The ECG provides an
11 electrical tracing of the heart that is recorded non-invasively by placing electrodes on the
12 chest, arms and legs. It is the simplest and fastest clinical method available to assess cardiac
13 abnormalities(8). A greater dispersion of QT intervals has been reported to be related to a
14 high dose of metformin (9) whereas the use of metformin has been found to be associated
15 with a reduced heart rate(10). However, the challenge remains as to how to identify ECG
16 markers associated with the use of diabetes medications due to our lack of knowledge of the
17 shared genetic landscape between them.

18

19 Recently, genome-wide association studies (GWAS) have been performed to identify variants
20 associated with the traditional ECG trace characteristics (ETCs) known to represent risk
21 factors for particular cardiac disorders. Importantly, the advent of the UK Biobank (UKB),
22 which collated genotype data from a large population, allowed us to perform large-scale

1 GWAS for less conventional ETCs, such as wave-related ETCs. The most recent GWAS on
2 the use of diabetes medications (11), and a recently established FinnGen Biobank, included
3 GWAS for various cardiac disorders offering us an opportunity to investigate the genetic
4 relationships between ETCs, diabetes medication use and CVDs, which is important for
5 identifying key ETCs of use in monitoring heart functions in patients using diabetes
6 medications.

7
8 In this study, GWAS were performed to identify genetic variants associated with ETCs, and
9 further detect their phenotypic and genetic relationships with diabetes medication use and
10 CVDs. The phenotypic and genetic relationships of these ETCs with medication use and
11 CVDs were investigated by regression models, LDSC(12) and seven different MR
12 approaches(13, 14, 15, 16, 17, 18, 19). This study provides genetic evidence to support the
13 use of ECG in monitoring CVDs in patients taking diabetes medications.

14

15 **Methods and Materials**

16 **Electrocardiogram morphology phenotypes and quality control**

17 The 12-lead electrocardiogram (ECG) records of 42,340 Europeans were downloaded from
18 the UK-Biobank (UKB). The ECG signal from each sample was a 10-second 12-lead signal
19 with a sampling rate of 500 Hz. The preprocessing and feature extraction of ECG can be
20 found in the supplementary materials.

21

22 **The phenotypic relationships between traits**

1 Regression models and the difference test were used to examine the phenotypic relationships
2 between ETCs, diabetes medication use, and cardiovascular diseases through medical records
3 in the UKB. The regression models were constructed using 38,953 samples from the UKB,
4 after removing samples with missing information, which were adjusted by age, sex, diabetes,
5 genotype batch and assessment centre.

6

7 **Genome-wide association studies for ETCs**

8 A genome-wide association study (GWAS) was performed for each ETC of European
9 ancestry using UKB data. European ancestry was selected by a two-stage approach as
10 described in Supplementary Material. The GWAS analysis was carried out by BOLT-
11 LMM(20), which is a Bayesian-based linear mixed model (LMM) that makes allowance for
12 the relatedness of individuals by adjusting the population structure and cryptic relatedness.
13 BOLT-LMM was corrected for age, sex, genotype batch and assessment center, and was fitted
14 by random effects of the restricted single nucleotide polymorphisms (SNPs) (with LD $r^2 <$
15 0.9). The SNPs used in BOLT-LMM were calibrated by the LD scores calculated from the
16 1000 Genomes Project Europeans reference(21). Genotyped SNPs were imputed according to
17 the Haplotype Reference Consortium reference panel(22). SNPs with minor allele frequency
18 (MAF) < 0.01 , imputation INFO score < 0.3 , P-value of Hardy-Weinberg test $< 1 \times 10^{-6}$,
19 minor allele counts < 5 , or the call rate < 0.05 were excluded, yielding a final set of 8.54M
20 SNPs.

21

22 **Causal relationship inference**

1 For the pairs of phenotypes with significant genetic correlations, Mendelian randomization
2 (MR) analysis was performed to detect their putative causal relationships through multiple
3 MR models (inverse-variance-weighted (IVW), MR-Egger, weighted mode, weighed median,
4 generalized summary-data-based Mendelian randomization (GSMR), causal analysis using
5 summary effect estimates (CAUSE), and MRlap). MRlap(23) was employed to correct the
6 sample overlap and is a method to approximate the sample overlap using cross-trait LD-score
7 regression (LDSC). It has been used in many studies for MR analysis(24, 25, 26, 27, 28). The
8 detailed descriptions of all the MR approaches are given in the Supplementary Material.

9
10 The putative causal relationship between two traits was considered to be determined if
11 consistent results were identified by more than four MR models with $FDR < 0.05$.
12 Depending on the data type of the outcome and exposure trait, the MR estimates (i.e., β_{xy})
13 were converted to liability scale using one of the methods described by Byrne et al.(29, 30).
14 The related formulae of the methods were described in the Supplementary Material.

15
16

17 **Results**

18

19 **Overview of this study**

20 Figure 1 depicts the overall workflow of this study. The study comprises three modules: (1)
21 extracting 168 ETCs from 42,340 UKB samples with 12-lead ECG data, and performing
22 genome-wide association studies (GWAS) for each ETC; (2) phenotypic relationship analysis

1 of the ETCs, cardiovascular diseases and diabetes medication use by leveraging the electronic
2 medical records of samples in UKB; (3) employing GWAS summary data of the ETCs,
3 diabetes medication use and CVDs to investigate the genetic correlations and genetic causal
4 links between them by LDSC and MR analysis. The GWAS data sources used in this study
5 were described in Supplementary Material.

6

7 **Genome-wide association study of ETCs**

8 The descriptions and short names of all the 168 ETCs are given in Supplementary Table 2.
9 Through the GWAS for these 168 ETCs, 124 independent SNPs were identified as study-wise
10 and genome-wide significantly ($P\text{-value} < 5 \times 10^{-8}/168$, LD $r^2 < 0.05$) associated with at least
11 one ETC (Figure 2, Supplementary Table 4). These 124 SNPs were systematically analyzed
12 by using both the extensive GWAS Catalog(31) and more recent phenome-wide associations
13 (PheWeb) from UKB(32) (Supplementary Table 4). Out of the 124 loci, 45 have been
14 reported to be associated with at least one type of heart function, such as electrocardiogram
15 morphology, QRS duration, left ventricular structure and function (Supplementary Table 4),
16 38 loci were found by previous studies to be associated with CVDs, such as atrial fibrillation
17 and flutter, atrioventricular block and heart failure, and 53 SNPs were significant *cis*-gene
18 expression quantitative trait loci (*cis*-eQTLs)(33)(Supplementary Table 4). These *cis*-eQTLs
19 were found to regulate genes enriched in phospholipid binding, phosphatidylinositol-5-
20 phosphate binding, transmembrane receptor protein kinase activity, and phosphatidylinositol-
21 4-phosphate binding functions ($P\text{-value} < 0.01$, Supplementary Figure 1) according to
22 clusterProfiler analysis (34). In total, 62 SNPs had not been previously reported as being

1 associated with any ETCs or heart disease (Figure 2A).

2

3 These 124 independent loci were assigned to 37 genes (Figure 2, Supplementary Table 4).

4 The SNPs in *NFIA* (7 SNPs), *KCND3* (10 SNPs), *CCDC141* genes (6 SNPs), and 12q24.21

5 region (18 SNPs) were significantly associated with more than 10 ETCs, suggesting key roles

6 for these genomic regions in characterizing ECG (Figure 2B, Supplementary Table 4). More

7 detailed information on these SNPs is provided in the Supplementary Material and

8 Supplementary Figure 2. Genes associated with each ETC, which were analyzed by gene-

9 level GWAS are given in the Supplementary Material.

10

11 **Phenotypic association among ETCs, use of diabetes medication, and CVDs.**

12 The phenotypic relationship among ETCs, diabetes medication use, and CVDs were analyzed

13 by means of difference test and regression model using ECG records of 38,953 samples (the

14 sample sizes used for the phenotypic analysis are given in Supplementary Table 7). The

15 regression analysis indicated 62 ETCs were significantly (Bonferroni adjusted P-value < 0.05)

16 correlated with use of diabetes medication (A10) and metformin (Figure 3A-B). Among them,

17 the ETCs related with R, T, ST, t_T, t_ST, and t_QT were significantly correlated with both

18 A10 and metformin use in more than 6 leads. Detailed information is given in Supplementary

19 Table 8. In addition, 6 ETCs were significantly (P-adjust < 0.05) correlated with metformin

20 use only (Figure 3B). Further analysis indicated that 125 ETCs were significantly correlated

21 with 10 CVDs (Bonferroni adjusted P-value < 0.05, Figure 3C). Additionally, coronary

22 atherosclerosis, major coronary heart disease events, cardiomyopathy, cardiovascular diseases,

1 heart failure and ischaemic heart disease were found to be significantly correlated with more
2 than 50 ETCs.

3
4 When we analysed the correlations between the use of diabetes medications and CVDs, we
5 found 8 CVDs (atrial fibrillation and flutter, angina pectoris, coronary atherosclerosis, major
6 coronary heart disease event, cardiovascular diseases, ischaemic heart disease, myocardial
7 infarction, peripheral artery disease) were significantly (Bonferroni adjusted P-value < 0.05)
8 correlated with A10 use and metformin use (Figure 3D).

9
10 **Genetic heritability and genetic correlations of ETCs, use of diabetes medications, and**
11 **cardiovascular diseases**

12 The methods for estimation of genetic heritability and genetic correlations were described in
13 Supplementary Material. The SNP-based heritability (h^2) of the 168 ETCs was estimated
14 using single-trait LDSC (Figure 4A, Supplementary Table 9). Of 168 ETCs, 143 showed an
15 FDR significance (FDR<0.05) and non-zero heritability, ranging from 0.03 to 0.19. The
16 heritability of 101 ETCs also reached the Bonferroni-corrected level of significance (P-values
17 <0.05/168; Figure 4A). The Liability-scale heritability (h^2) of A10 and metformin use were
18 0.046 (P= 5.88×10^{-52}) and 0.030 (P= 3.68×10^{-56}), respectively. The heritability of
19 cardiovascular diseases (CVDs) was estimated using 22 GWAS summary statistics of CVDs
20 from FinnGen (Supplementary Table 1). The Liability-scale h^2 for these CVDs ranged from
21 0.015 to 0.088 and P < 0.05/22 for 21 CVDs (Figure 4B, Supplementary Table 10).

22

1 From 62 ETCs that were found to be significantly correlated with A10 and metformin use in
2 phenotypic analysis, we found 37 had significant heritability (Bonferroni corrected P-adjust <
3 0.05, $h^2 > 0$). We firstly used the cross-trait LDSC analysis to estimate the genetic
4 correlations between these 37 ETCs and the use of diabetes medications (Figure 5A). As
5 shown in Figure 5A, the use of A10 was found to be significantly correlated (Bonferroni
6 corrected P-adjust < 0.05) with 8 ETCs. Among these ETCs, five (II_T, aVF_T, V4_T,
7 aVL_R-area, and aVR_T-area) were also genetically correlated with metformin use. Detailed
8 information is given in Supplementary Table 11.

9

10 Next, the genetic correlation between the ETCs and cardiovascular diseases (CVDs) was
11 examined using 21 GWAS summary statistics of CVDs from FinnGen (Supplementary Table
12 1). The genetic correlations between the 37 ETCs and 21 CVDs (Bonferroni corrected
13 $P_{heritability} < 0.05/22$) were analyzed by cross-trait LDSC, indicating that 6 ETC-CVD pairs
14 had significant genetic correlations (absolute r_g ranging from 0.37 to 0.44, FDR < 0.05)
15 (Figure 5B, Supplementary Table 12). Among them, heart failure (HF) was significantly
16 correlated with increased III_S whereas ischemic heart disease (IHD) and coronary
17 atherosclerosis (CA) were found to be significantly correlated with decreased aVR_T-area.
18 The phenotypic correlation analysis also indicated that HF is significantly (Bonferroni
19 adjusted P-value = 0.009) correlated with increased III_S and IHD. Further, CA was found to
20 be significantly (P-adjust = 2.34×10^{-12} and P-adjust = 6.74×10^{-12}) correlated with
21 decreased aVR_T-area by phenotypic correlation analysis (Figure 3C).

22

1 Finally, the genetic correlations between CVDs and the use of diabetes medication were
2 estimated by cross-trait LDSC (Figure 5C), indicating 17 CVDs have significant genetic
3 correlation with the use of both A10 and metformin (Bonferroni adjusted $P < 0.05$,
4 Supplementary Table 13). Among them, 8 CVDs were found to be significantly (P -adjust $<$
5 0.05) correlated with both A10 and metformin use by phenotypic analysis (Figure 3D).

6

7 **Mendelian Randomization analysis to identify putative causal relationships between**
8 **ETCs, use of diabetes medications and CVDs.**

9 The putative causal relationship between the 14 ETCs and the use of diabetes medications
10 that have shown significant phenotypic (Figure 3A-B) and genetic correlations (Figure 5A)
11 was analyzed by means of seven different MR approaches (see Methods). Six out of seven
12 MR models provided evidence supporting use of A10 has a putative casual effect on aVR_T-
13 area with estimated liability-scale odds ratios (ORs) ranging from 0.74-0.92 (Figure 6)
14 (weighted median: $FDR = 0.018$, IVW: $FDR = 5.92 \times 10^{-4}$, weighted mode: $FDR = 0.018$,
15 GSMR: $FDR = 2.50 \times 10^{-5}$, MRlap: $FDR = 5.92 \times 10^{-4}$, and CAUSE: $FDR = 5.74 \times 10^{-3}$).
16 Furthermore, the use of metformin also had a putative causal effect ($FDR < 0.05$) on aVR_T-
17 area as indicated by six out of seven MR methods (the exception being MR-Egger), with
18 estimated liability-scale ORs ranging from 0.68-0.93 (Figure 6). These results suggested that
19 those samples associated with the use of A10 or metformin were prone to have an
20 approximately 12% smaller aVR_T-area as compared to individuals taking neither A10 or
21 metformin.

22

1 When the MR methods were employed to identify putative causal relationships between the
2 ETCs and CVDs, a causal effect (OR:1.02~1.03) of III_S on HF was identified by five MR
3 methods (Figure 7, Supplementary Table 14), suggesting that an increase of 1 μ V in III_S led
4 to an ~2% increase in the risk of HF. No putative causal effect was observed in the reverse
5 direction by MR analysis.

6

7 The causal relationship between CVDs and A10/metformin use was further explored.
8 Specifically, the role of genetics to ascertain the impact of A10 on IHD risk was supported by
9 six out of the seven MR methods (OR: 1.14-2.30; FDR = 3.60×10^{-8} - 2.54×10^{-2} for IHD1;
10 OR: 1.13-2.14; FDR = 4.80×10^{-7} - 4.52×10^{-2} for IHD2) (Figure 8, Supplementary Figure
11 6). Further, all seven MR methods provided consistent evidence in support of a causal effect
12 of metformin use on IHD (OR: 1.13-2.14 for IHD1; OR: 1.11-2.66 for IHD2) (Figure 8,
13 Supplementary Figure 6). In addition, the use of genetics to ascertain the impact of A10 or
14 metformin use also indicated a putative causal effect on CA (Figure 8, OR: 1.11-2.66 for A10;
15 OR: 1.13-3.85 for metformin). More results about the genetics of A10/metformin use and
16 CVDs are summarized in Supplementary Table 15.

17

18 Taken together, the use of metformin or A10 has a putative causal effect on aVR_T-area, as
19 well as on the risk of CA and IHD, whilst aVR_T-area exhibits a genetic correlation with IHD
20 and CA (only r_g was significant). In addition, the III_S has a putative causal effect on HF.
21 These relationships provided genetic evidence in support of the results of the phenotypic
22 analysis in Figure 3.

1

2 **Discussion**

3 In this study, GWAS were performed for 12-lead ECG that was characterized by 168 ETCs,
4 including not only the classical ETCs but also wave area-related features. The GWAS
5 identified 124 independent SNPs associated with at least one ETC. Half of these have not
6 been reported as being associated with CVDs or classical ETCs (Supplementary Material,
7 Supplementary Table 5). Many variants near genes with calcium or potassium channel
8 functions were identified as being associated with R-area (Supplementary Table 4) whereas
9 variants near genes known to be related with autism spectrum disorder or schizophrenia were
10 associated with S-area (Supplementary Table 4). Moreover, variants identified as being
11 associated with T-area tended to be located near sodium channel genes (Supplementary Table
12 4). For example, *HAND1* and *SAP30L* were found to be associated with many S-wave related
13 traits. *HAND1* is already known to be linked to developmental heart chamber defects(35) and
14 SNPs near to *HAND1* have been reported to be associated with the QR-interval(36). Another
15 gene, *KCND3*, was identified as being significantly associated with 29 ETCs. Although
16 previous studies have demonstrated that *KCND3* is associated with early repolarization on
17 electrocardiograms(37, 38), we provide further evidence to support a genetic association
18 between *KCND3* with various ETCs, suggesting an important role for *KCND3* in cardiac
19 function.

20

21 The heritability estimation of ETCs indicated that the highest heritability 0.192 (SE=0.027
22 and SE = 0.029) for the difference between R start and S end in voltage in aVR lead

1 (aVR_J_up) and the difference between R start and S end in voltage in I lead (I_J_up)
2 (Supplementary Table 7). The 10 highest heritability estimates of ETCs were all related to the
3 ST segment (Supplementary Table 7). Interestingly, we found that the 10 ETCs with the
4 lowest heritability included nine T-wave related features. Thus, T-wave related features may
5 be characterized by lower heritability than features associated with other ECG waves.

6
7 A recent clinical study investigated the effects of long-term metformin use on heart disease
8 outcome by following 3,234 participants for 21 years, and demonstrated that metformin had
9 no impact on cardiovascular survival(39). Another randomized clinical trial in 368 samples
10 investigating the effects of metformin on coronary heart disease indicated that taking
11 metformin had no effect on cardiovascular disease(40). This study has provided the first
12 genetic evidence that patients taking metformin ($r_g = -0.34$ and P-value = 3.00×10^{-4}) or
13 A10 ($r_g = -0.33$ and P-value = 8.82×10^{-6}) tend to display a decreased aVR_T-area. Further
14 MR analysis indicated that the genetic factors associated with the use of metformin or A10
15 have a putative causal effect in terms of an 11% decrease in T-wave area compared to
16 individuals not taking these medications.

17
18 This study had several limitations. Firstly, GWAS of ECG was based on the earliest ECG of
19 the samples in UKB. Further studies are required to leverage the longitudinal ECG data of
20 samples in order to consider ETC changes during aging. Secondly, the limited number of
21 samples for which ECG information was available reduced the power of our GWAS analysis
22 to detect SNPs associated with ETCs and in the MR analysis. Thirdly, all the analyses in this

1 study were based on data from Europeans because ECG and genetic data from other
2 ethnicities are quite limited.

3

4 **Conclusions**

5 GWAS was performed by extracting comprehensive traditional and non-traditional ETCs and
6 identifying 62 novel genetic loci as being significantly associated with them. The genetic
7 relationship between ETCs, the use of diabetes medication and CVDs was investigated, and
8 supported the putative causal effects of the use of diabetes medication on the aVR_T-area and
9 the risk of IHD and CA. These findings suggest new avenues for the use of ECG in the
10 monitoring or early diagnosis of IHD or CA risk in patients taking diabetes medications.

11

12 **Competing interests**

13 The authors declare that they have no competing interests.

14

15 **Funding**

16 The work was funded by the National Key Research and Development Program of China
17 (2020YFB0204803), the Natural Science Foundation of China (81801132, 81971190, and
18 61772566), Guangdong Key Field Research and Development Plan (2019B020228001 and
19 2021A1515010256) and Guangzhou Science and Technology Research Plan (202007030010).

20

21 **Acknowledgments**

22 Data for this study were obtained from the UK-Biobank (application number #65805).

1

References

- 1 1. Kirkman MS, Mahmud H, Korytkowski MT. Intensive Blood Glucose Control and Vascular Outcomes in
2 Patients with Type 2 Diabetes Mellitus. *Endocrinol Metab Clin North Am*. 2018;47(1):81-96.
- 3 2. Group AC, Patel A, MacMahon S, Chalmers J, Neal B, Billot L, et al. Intensive blood glucose control and
4 vascular outcomes in patients with type 2 diabetes. *N Engl J Med*. 2008;358(24):2560-72.
- 5 3. Duckworth W, Abraira C, Moritz T, Reda D, Emanuele N, Reaven PD, et al. Glucose control and vascular
6 complications in veterans with type 2 diabetes. *N Engl J Med*. 2009;360(2):129-39.
- 7 4. Riddle MC. Effects of intensive glucose lowering in the management of patients with type 2 diabetes
8 mellitus in the Action to Control Cardiovascular Risk in Diabetes (ACCORD) trial. *Circulation*. 2010;122(8):844-6.
- 9 5. Holman RR, Sourij H, Califf RM. Cardiovascular outcome trials of glucose-lowering drugs or strategies
10 in type 2 diabetes. *Lancet*. 2014;383(9933):2008-17.
- 11 6. Dormandy JA, Charbonnel B, Eckland DJ, Erdmann E, Massi-Benedetti M, Moules IK, et al. Secondary
12 prevention of macrovascular events in patients with type 2 diabetes in the PROactive Study (PROspective
13 pioglitAzone Clinical Trial In macroVascular Events): a randomised controlled trial. *Lancet*.
14 2005;366(9493):1279-89.
- 15 7. Lim LL, Chow E, Chan JCN. Cardiorenal diseases in type 2 diabetes mellitus: clinical trials and real-
16 world practice. *Nat Rev Endocrinol*. 2023;19(3):151-63.
- 17 8. AlGhatrif M, Lindsay J. A brief review: history to understand fundamentals of electrocardiography.
18 *Journal of community hospital internal medicine perspectives*. 2012;2(1).
- 19 9. Costa ECdS, Gonçalves AA, Areas MA, Morgabel RGB. Effects of metformin on QT and QTc interval
20 dispersion of diabetic rats. *Arquivos Brasileiros de Cardiologia*. 2008;90:254-60.
- 21 10. Emeka PM, Al-Ahmed A. Effect of metformin on ECG, HR and BP of rats administered with cardiotoxic
22 agent doxorubicin. *International Journal of Basic & Clinical Pharmacology*. 2017;6(5):1054-9.
- 23 11. Wu Y, Byrne EM, Zheng Z, Kemper KE, Yengo L, Mallett AJ, et al. Genome-wide association study of
24 medication-use and associated disease in the UK Biobank. *Nature communications*. 2019;10(1):1891.
- 25 12. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric
26 Genomics C, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association
27 studies. *Nature genetics*. 2015;47(3):291-5.
- 28 13. Hemani G, Bowden J, Davey Smith G. Evaluating the potential role of pleiotropy in Mendelian
29 randomization studies. *Hum Mol Genet*. 2018;27(R2):R195-R208.
- 30 14. Zhu Z, Zheng Z, Zhang F, Wu Y, Trzaskowski M, Maier R, et al. Causal associations between risk factors
31 and common diseases inferred from GWAS summary data. *Nature communications*. 2018;9(1):224.
- 32 15. Burgess S, Dudbridge F, Thompson SG. Combining information on multiple instrumental variables in
33 Mendelian randomization: comparison of allele score and summarized data methods. *Statistics in Medicine*.
34 2015;35(11):1880-906.
- 35 16. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization
36 with Some Invalid Instruments Using a Weighted Median Estimator. *Genet Epidemiol*. 2016;40(4):304-14.
- 37 17. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic
38 variants using summarized data. *Genet Epidemiol*. 2013;37(7):658-65.
- 39 18. Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via
40 the zero modal pleiotropy assumption. *Int J Epidemiol*. 2017;46(6):1985-98.
- 41 19. Morrison J, Knoblauch N, Marcus J, Stephens M, He X. Mendelian randomization accounting for
42 correlated and uncorrelated pleiotropic effects using genome-wide summary statistics. *bioRxiv*. 2020:682237.

-
- 1 20. Loh PR, Tucker G, Bulik-Sullivan BK, Vilhjálmsson BJ, Finucane HK, Salem RM, et al. Efficient Bayesian
2 mixed-model analysis increases association power in large cohorts. *Nature genetics*. 2015;47(3):284-90.
 - 3 21. Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, Gibbs RA, et al. A map of human genome
4 variation from population-scale sequencing. *Nature*. 2010;467(7319):1061-73.
 - 5 22. Loh PR, Danecek P, Palamara PF, Fuchsberger C, Y AR, H KF, et al. Reference-based phasing using the
6 Haplotype Reference Consortium panel. *Nature genetics*. 2016;48(11):1443-8.
 - 7 23. Mounier N, Kutalik Z. Bias correction for inverse variance weighting Mendelian randomization. 2021.
 - 8 24. Prince C, Sharp GC, Howe LD, Fraser A, Richmond RC. The relationships between women's
9 reproductive factors: a Mendelian randomisation analysis. *BMC medicine*. 2022;20(1):1-16.
 - 10 25. Schoeler T, Baldwin J, Allegrini A, Barkhuizen W, McQuillin A, Pirastu N, et al. Novel biological insights
11 into the common heritable liability to substance involvement: a multivariate genome-wide association study.
12 *Biological Psychiatry*. 2022.
 - 13 26. Tylee DS, Lee YK, Wendt FR, Pathak GA, Levey DF, De Angelis F, et al. An Atlas of Genetic Correlations
14 and Genetically Informed Associations Linking Psychiatric and Immune-Related Phenotypes. *JAMA psychiatry*.
15 2022.
 - 16 27. Jiang T, Gill D, Butterworth AS, Burgess S. An empirical investigation into the impact of winner's curse
17 on estimates from Mendelian randomization. *medRxiv*. 2022.
 - 18 28. Howe L, Rasheed H, Jones PR, Boomsma DI, Evans DM, Giannelis A, et al. Educational attainment,
19 health outcomes and mortality: a within-sibship Mendelian randomization study. *medRxiv*. 2022.
 - 20 29. Byrne EM, Zhu Z, Qi T, Skene NG, Bryois J, Pardinas AF, et al. Conditional GWAS analysis to identify
21 disorder-specific SNPs for psychiatric disorders. *Molecular psychiatry*. 2021;26(6):2070-81.
 - 22 30. Zhang H, Xiu X, Yang Y, Yang Y, Zhao H. Identification of Putative Causal Relationships Between Type 2
23 Diabetes and Blood-Based Biomarkers in East Asians by Mendelian Randomization. *Am J Epidemiol*. 2022.
 - 24 31. Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, et al. The NHGRI-EBI GWAS
25 Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic
26 Acids Res*. 2019;47(D1):D1005-d12.
 - 27 32. UKBiobank TOPMed-imputed PheWeb [Available from: <https://pheweb.org/UKB-TOPMed/>].
 - 28 33. Vösa U, Claringbould A, Westra H-J, Bonder MJ, Deelen P, Zeng B, et al. Unraveling the polygenic
29 architecture of complex traits using blood eQTL metaanalysis. *bioRxiv*. 2018:447367.
 - 30 34. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among
31 gene clusters. *Omics : a journal of integrative biology*. 2012;16(5):284-7.
 - 32 35. Hofbauer P, Jahnel SM, Papai N, Giesshammer M, Deyett A, Schmidt C, et al. Cardioids reveal self-
33 organizing principles of human cardiogenesis. *Cell*. 2021;184(12):3299-317 e22.
 - 34 36. Evans DS, Avery CL, Nalls MA, Li G, Barnard J, Smith EN, et al. Fine-mapping, novel loci identification,
35 and SNP association transferability in a genome-wide association study of QRS duration in African Americans.
36 *Hum Mol Genet*. 2016;25(19):4350-68.
 - 37 37. Teumer A, Trenkwalder T, Kessler T, Jamshidi Y, van den Berg ME, Kaess B, et al. KCND3 potassium
38 channel gene variant confers susceptibility to electrocardiographic early repolarization pattern. *JCI Insight*.
39 2019;4(23).
 - 40 38. Takayama K, Ohno S, Ding WG, Ashihara T, Fukumoto D, Wada Y, et al. A de novo gain-of-function
41 KCND3 mutation in early repolarization syndrome. *Heart Rhythm*. 2019;16(11):1698-706.
 - 42 39. Goldberg RB, Orchard TJ, Crandall JP, Boyko EJ, Budoff M, Dabelea D, et al. Effects of long-term
43 metformin and lifestyle interventions on cardiovascular events in the diabetes prevention program and its
44 outcome study. *Circulation*. 2022;145(22):1632-41.

1 40. Preiss D, Lloyd SM, Ford I, McMurray JJ, Holman RR, Welsh P, et al. Metformin for non-diabetic
2 patients with coronary heart disease (the CAMERA study): a randomised controlled trial. *Lancet Diabetes*
3 *Endocrinol.* 2014;2(2):116-24.

4

1 **Legends**

2

3 **Graphical abstract.** Firstly, genome-wide association studies (GWAS) were performed for 168 ECG trace
4 characteristics (ETCs) extracted from the 12-lead ECGs of 42,340 Europeans in the UK Biobank which
5 identified 124 independent single nucleotide polymorphisms (SNPs) significantly associated with at least
6 one ETC (62 were novel). Secondly, the phenotypic relationships were analyzed by regression model and
7 found significant phenotypic association between 62 ETCs, 8 CVDs and A10/metformin use, and between
8 125 ETCs and 10 CVDs. Thirdly, pairwise genetic correlations of diabetes medications use (ATC code:
9 A10 drugs, and metformin), T-wave area in lead aVR (aVR_T-area), and the risks of ischemic heart disease
10 (IHD) and coronary atherosclerosis (CA) were identified by linkage disequilibrium score regression
11 (LDSC). Then, Mendelian randomization (MR) support a putative causal effect of diabetes medications use
12 on decreasing aVR_T-area, and on increasing risk of IHD and CA.

13

14 **Figure 1.** Schematics for genetic and phenotypic analysis of relationships between ETCs, medication use,
15 and CVDs. Step 1: Genome-wide association study for 168 ETCs by BLOT-LMM; Step 2: Analyzing
16 phenotypic relationships among ETCs, taking A10/metformin and cardiovascular diseases, by employing
17 medical records from the UK-Biobank; Step 3: Analyzing genetic relationships between the ETCs,
18 medication usage and CVDs through LDSC and seven MR methods.

19

20 **Figure 2.** Genome-wide association studies for the ETCs. A. GWAS identified 124 independent SNPs as
21 being significantly associated with at least one ETC ($P\text{-value} < \frac{5 \times 10^{-8}}{168}$), half of which have been reported to
22 be associated with ECG or CVDs (red points), and 62 SNPs that were novel findings from this study (blue
23 points). The blue text on the Y axis refers to the ETCs of R/S/T wave amplitudes in each lead; the green

1 text on the *Y* axis refers to the ETCs of combinations of amplitudes; the red text on the *Y* axis denotes the
2 ETCs of wave areas for R/S/T in each lead; the brown text on the *Y* axis refers to the ETCs of time interval
3 in each lead. The background color represents different chromosomes. The chromosome index is shown on
4 the *X* axis. B. Manhattan plot of the ETCs and the smallest P-value of each SNP across all the traits is
5 shown. The red dashed line shows the threshold P-value = $5 \times 10^{-8} / 168$. Orange points represent SNPs with
6 P-value $< 5 \times 10^{-8} / 168$, whereas black points represent SNPs with LD $r^2 < 0.05$ with at least one SNP within
7 1000 bp.

8
9 **Figure 3.** Phenotypic relationship among ETCs, A10/metformin use, and CVDs. A. The coefficients of the
10 regression analysis between ETCs and A10 use. The markers on the horizontal axis denote the names of
11 leads, whilst the labels on the vertical axis denote the ETCs. The darker colors represent higher values of
12 coefficients (*beta*). B. Coefficients of the regression analysis between ETCs and metformin use. The
13 markers on the horizontal axis are the lead name, the labels on the vertical axis are the ETCs. The darker
14 colors represent higher values of coefficients (*beta*). C. Relationships between ETCs and CVDs. The
15 markers on the horizontal axis are the abbreviations of CVDs, whilst the labels on the vertical axis are the
16 ETCs. The darker colors represent lower P-values. D. Relationships between medicine use and CVDs. The
17 markers on the horizontal axis denote medicine use, the labels on the vertical axis are the CVDs. The
18 darker colors represent lower P-values. “*” denotes Bonferroni adjusted P-value < 0.05 .

19
20 **Figure 4.** The heritability (h^2) of ETCs and CVDs. A. The heritability of each ETC from different leads,
21 estimated by LDSC with unconstrained intercept. The markers on the horizontal axis are the lead name, the
22 labels on the vertical axis are the ETCs. The darker colors represent higher values of h^2 . “*” denotes a P-

1 value < 0.05/168. B. Heritability of 22 CVDs was calculated by single-trait LDSC. Red bars denote P-
2 values for heritability estimation lower than 0.05/22. The labels on the vertical axis are the Abbreviations
3 which can be checked in Supplementary Table 1.

4

5 **Figure 5.** Genetic correlations among ETCs, medication use and CVDs. A. Genetic correlation (r_g)
6 between ETCs and medication use. Blue and red bars represent metformin use and A10 use respectively. “*”
7 denotes a Bonferroni corrected P-value < 0.05. B. LDSC analysis was used to calculate genetic correlation
8 (r_g) between ETCs and CVDs. The larger and darker dots represent higher r_g . Blue and red represent
9 negative and positive correlations respectively. “*” denotes a FDR < 0.05. C. Genetic correlation (r_g)
10 between medication use and CVDs by LDSC analysis. The larger and darker dots represent higher absolute
11 r_g . Blue and red colors represent negative and positive correlations respectively. “*” means Bonferroni
12 corrected P-value < 0.05.

13

14 **Figure 6.** The causal relationship (in log-scale odds ratio [OR]) between A10 or metformin use and
15 aVR_T-area was estimated by seven MR methods. The value of ORs and 95% confidence intervals (CIs)
16 indicates the change of aVR_T-area for samples using A10 or metformin compared to the samples
17 associated with never having taken these two medicines. CAUSE: causal analysis using summary effect
18 estimates; GSMR: generalized summary-data-based Mendelian randomization; IVW: inverse variance
19 weighted. Red and blue colors denote different directions of MR analysis.

20

21 **Figure 7.** The causal relationship (in log-scale odds ratio [OR]) of III_S on HF was estimated by seven
22 MR methods in two directions. Odds ratios (ORs) and 95% confidence intervals (CIs) express the risk of

1 HF for the change of III_S compared to the samples without HF.

2

3 **Figure 8.** The causal relationship (in log-scale odds ratio [OR]) of A10 or metformin use on IHD or CA

4 was estimated by seven MR methods. Odds ratios (ORs) and 95% confidence intervals (CIs) express the

5 risk of IHD for samples associated with A10 or metformin use compared to the samples associated with

6 never take these two medicines. IHD1: I9_ISCHHEART, ischemic heart disease; CA: I9_CORATHER,

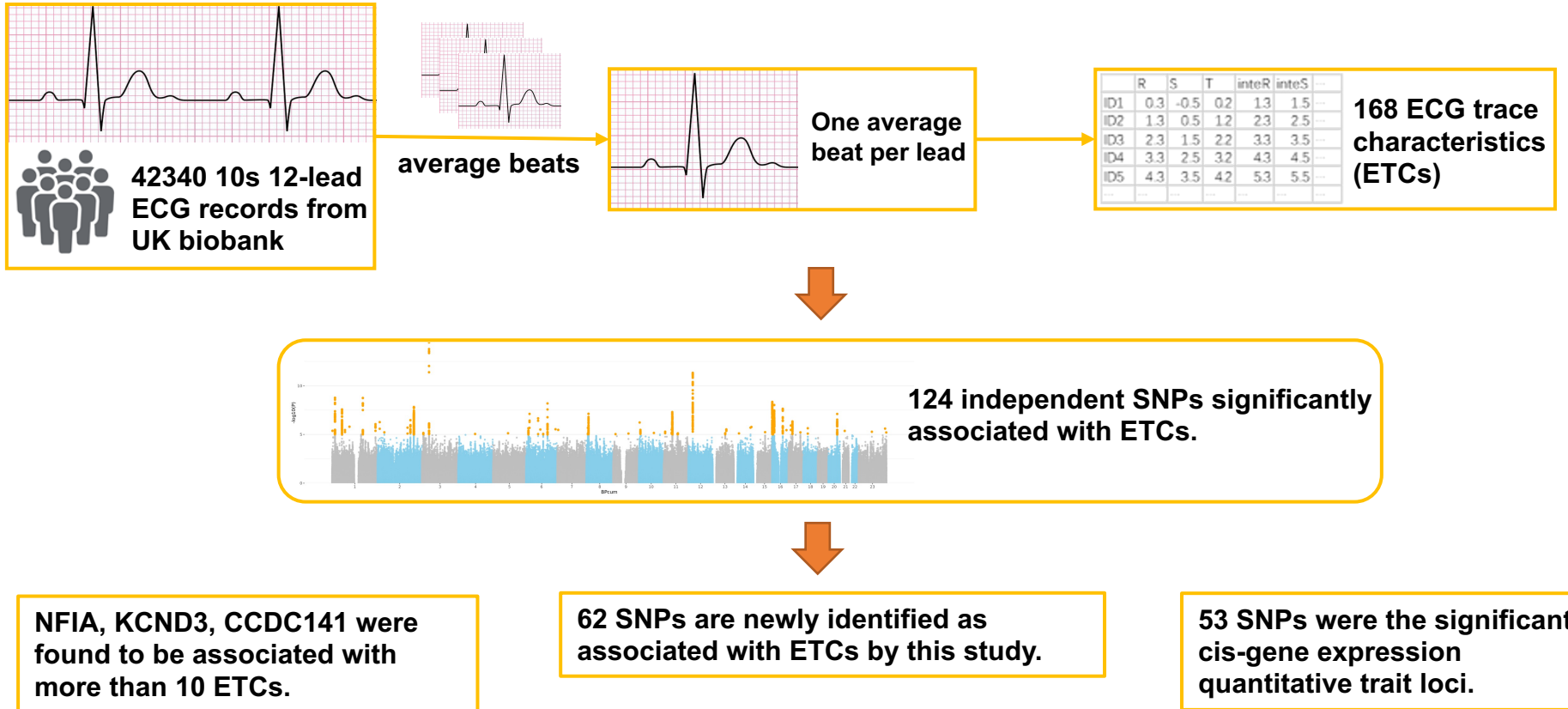
7 coronary atherosclerosis.

8

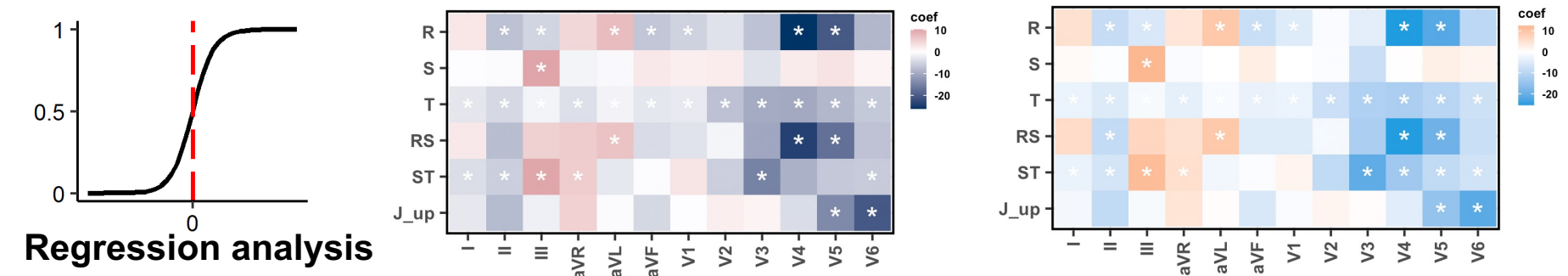
9

10

① Genome wide association study for ECG trace characteristics



② Phenotypic correlations among ETCs, medications use and cardiovascular diseases (CVDs)



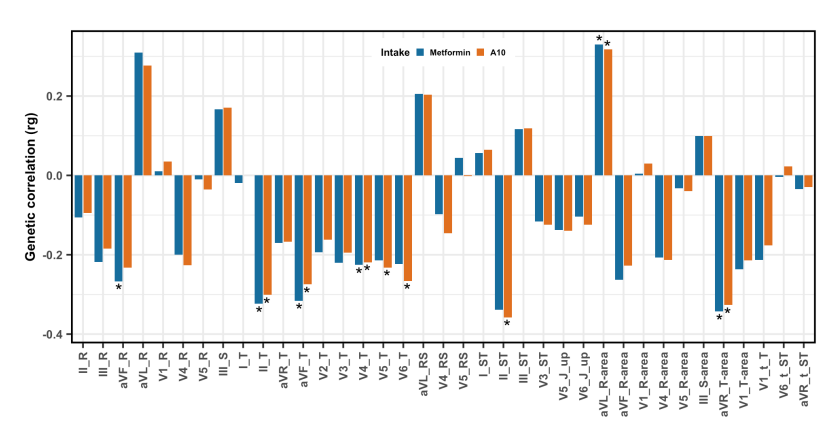
③ Genetic relationship among ETCs, medications use, and CVDs

37 ECG trace characteristics

2 diabetes medications use

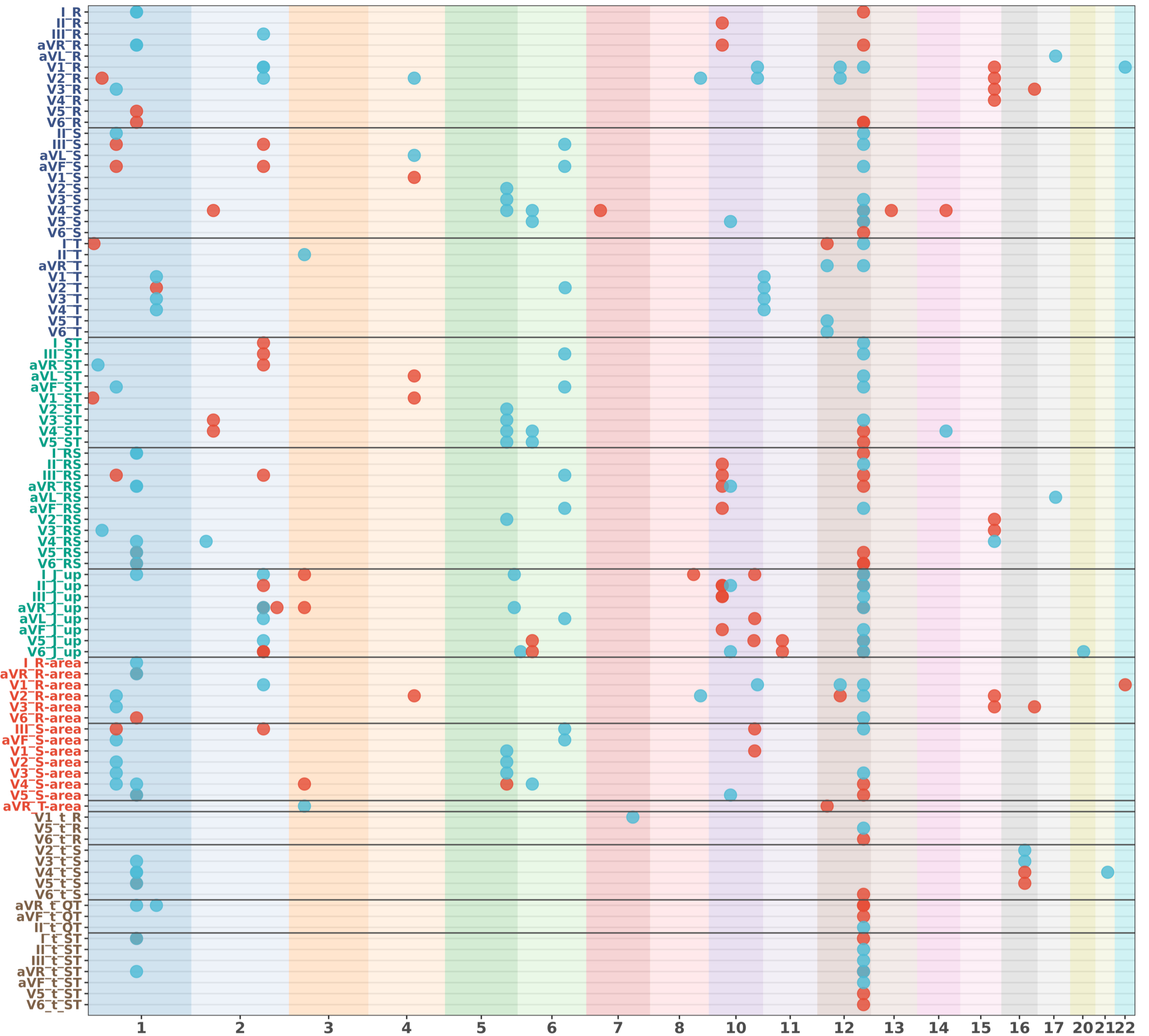
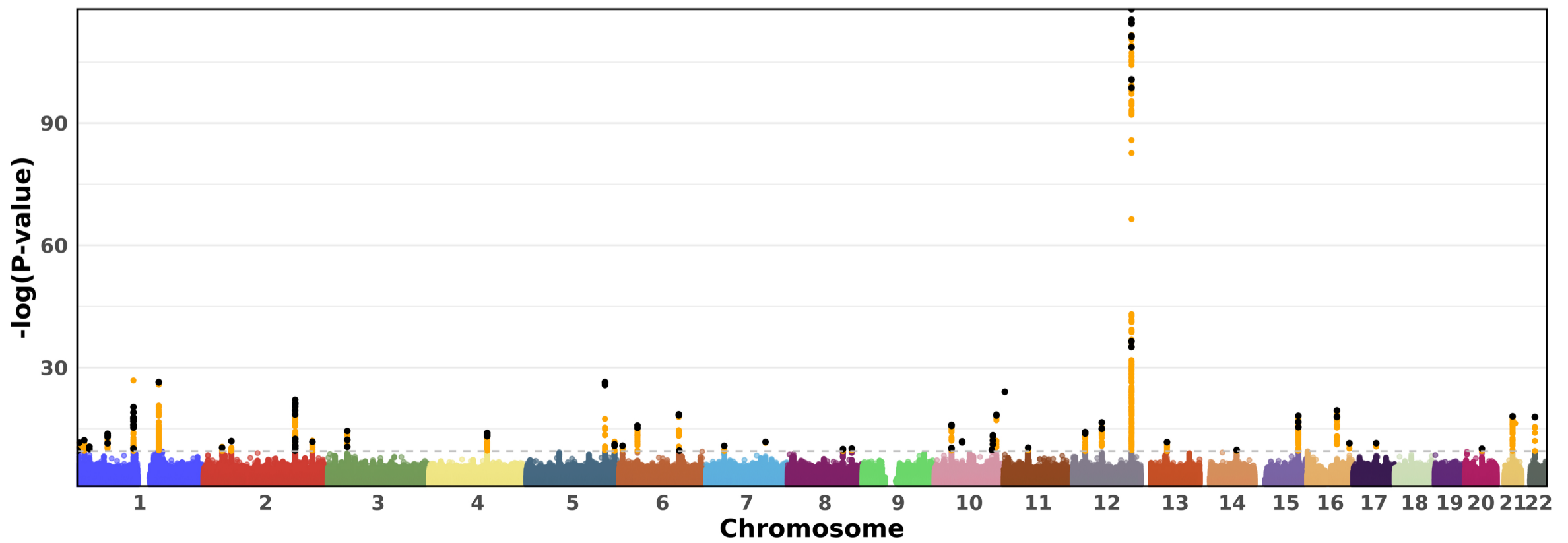
21 cardiovascular diseases

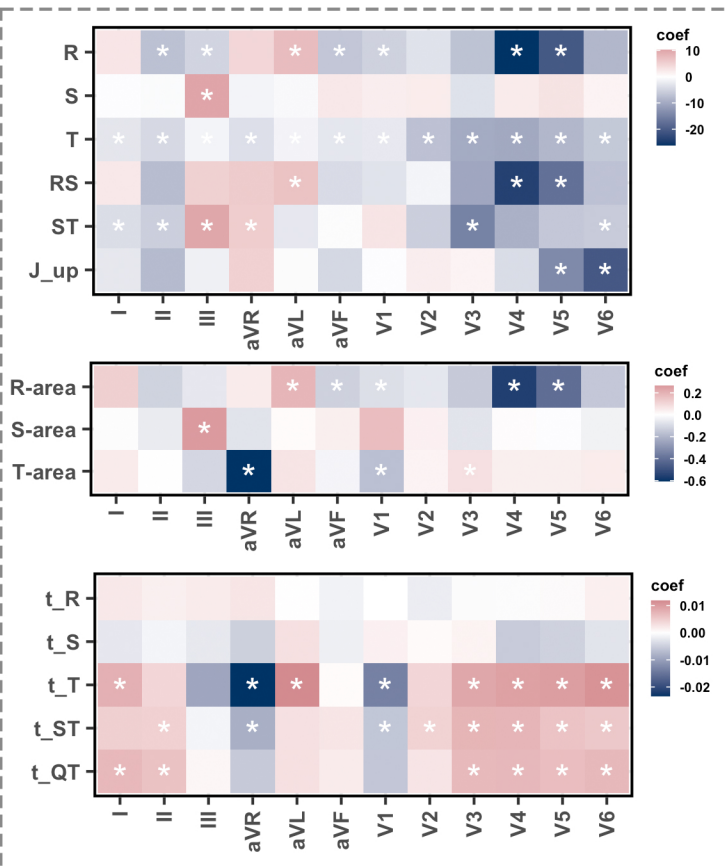
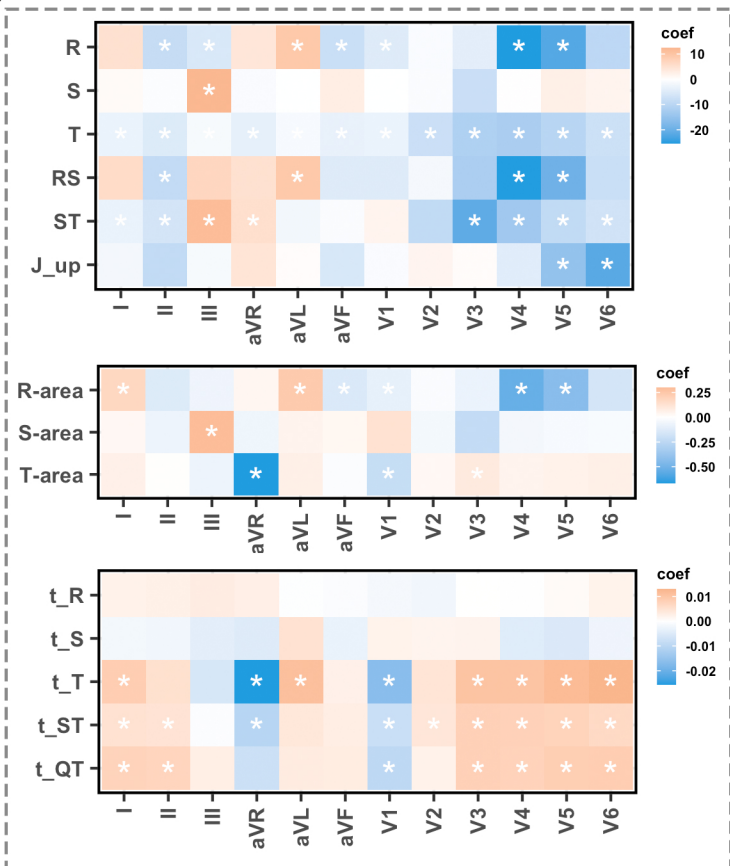
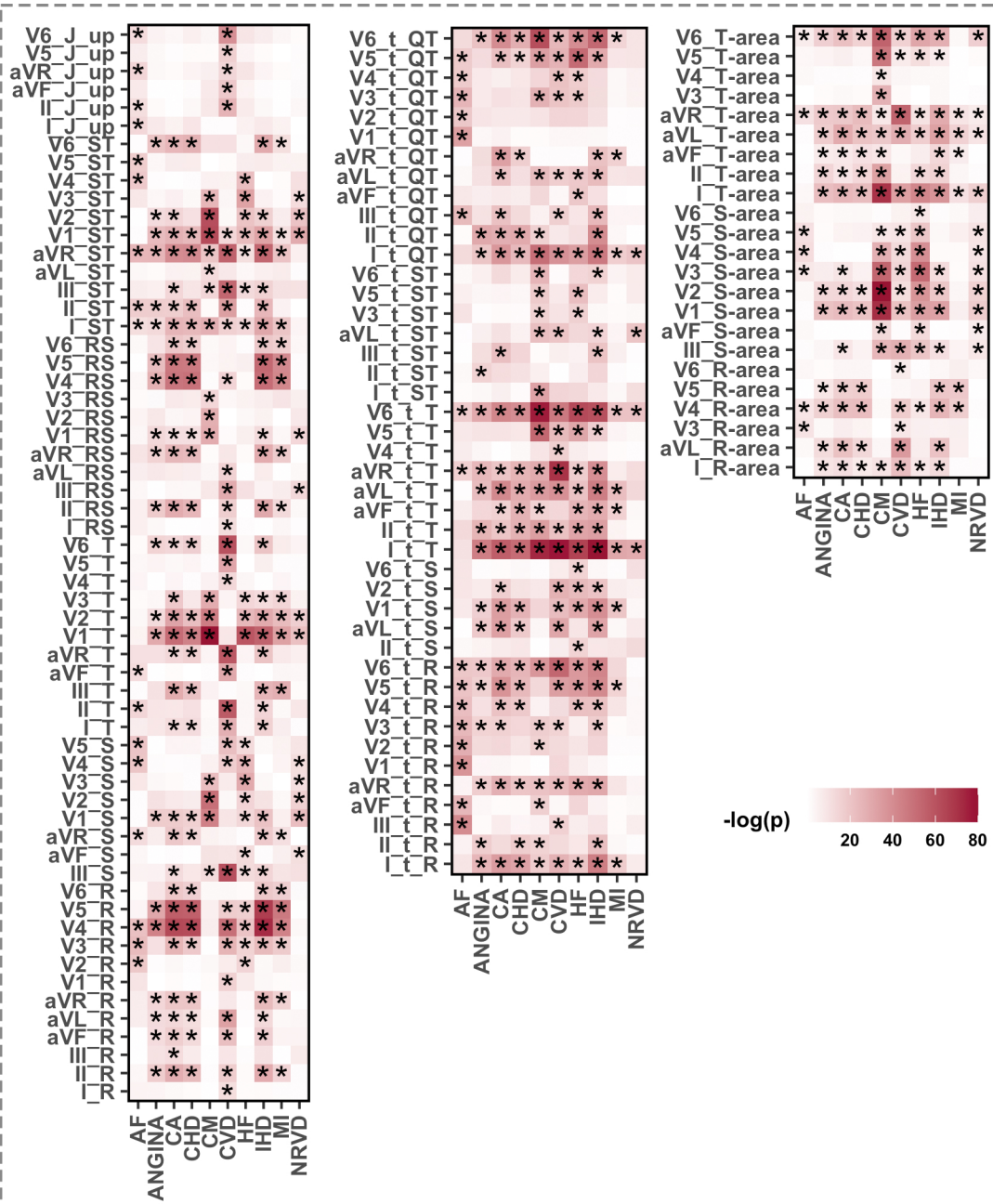
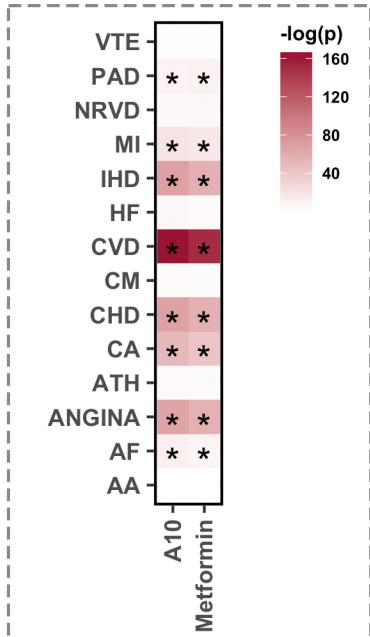
Genetic correlation estimated by Cross-trait LDSC

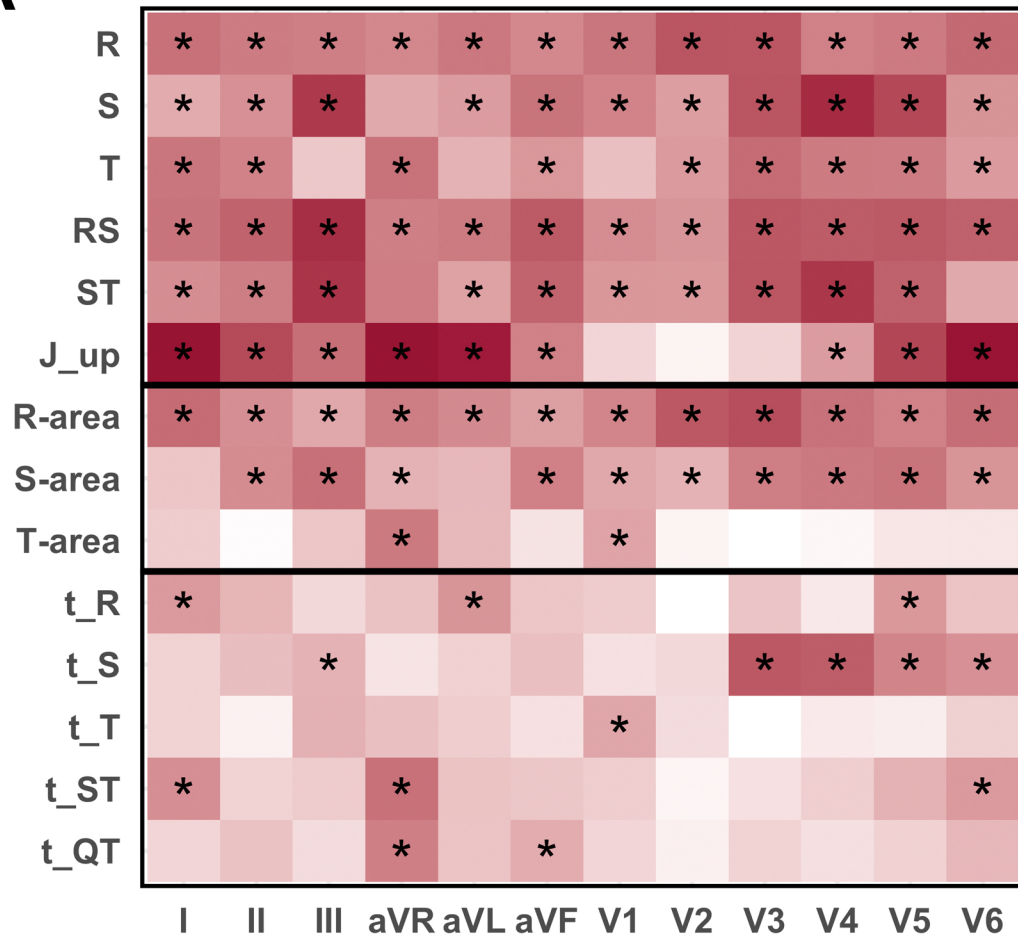
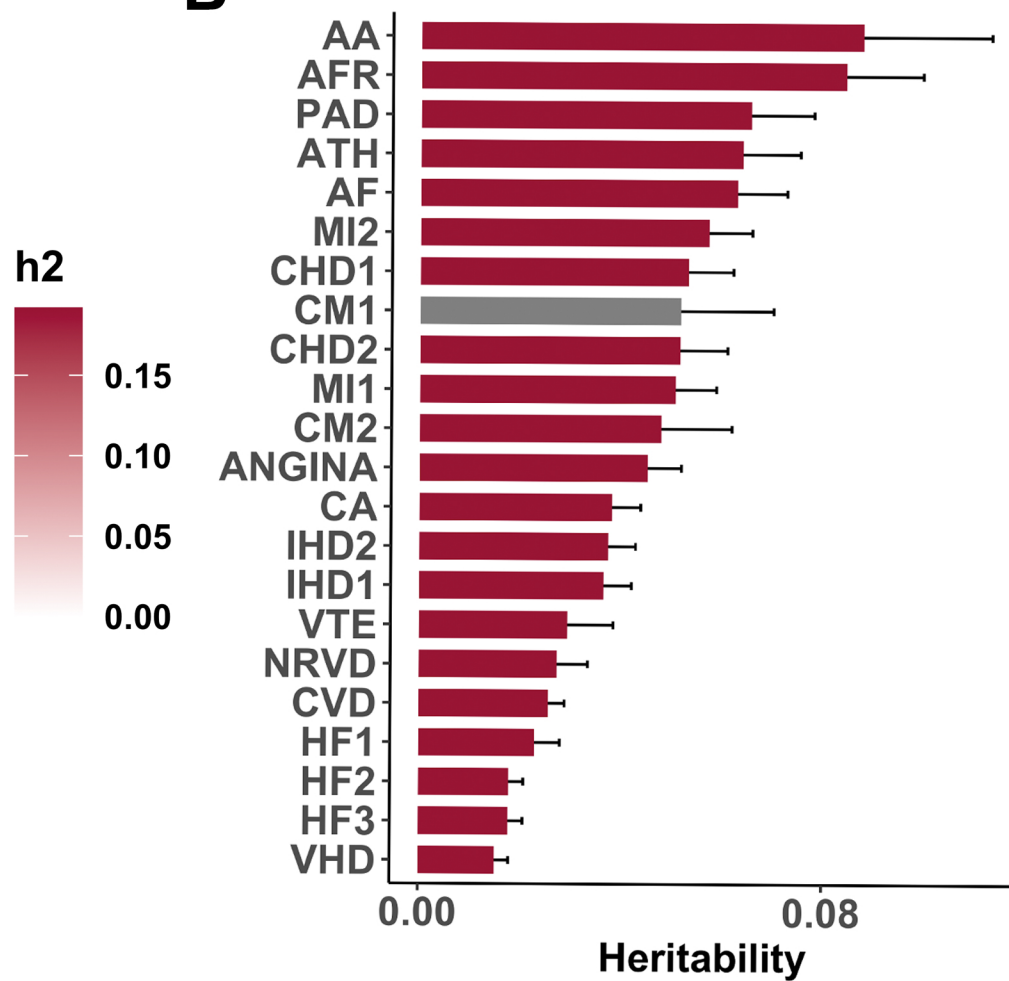


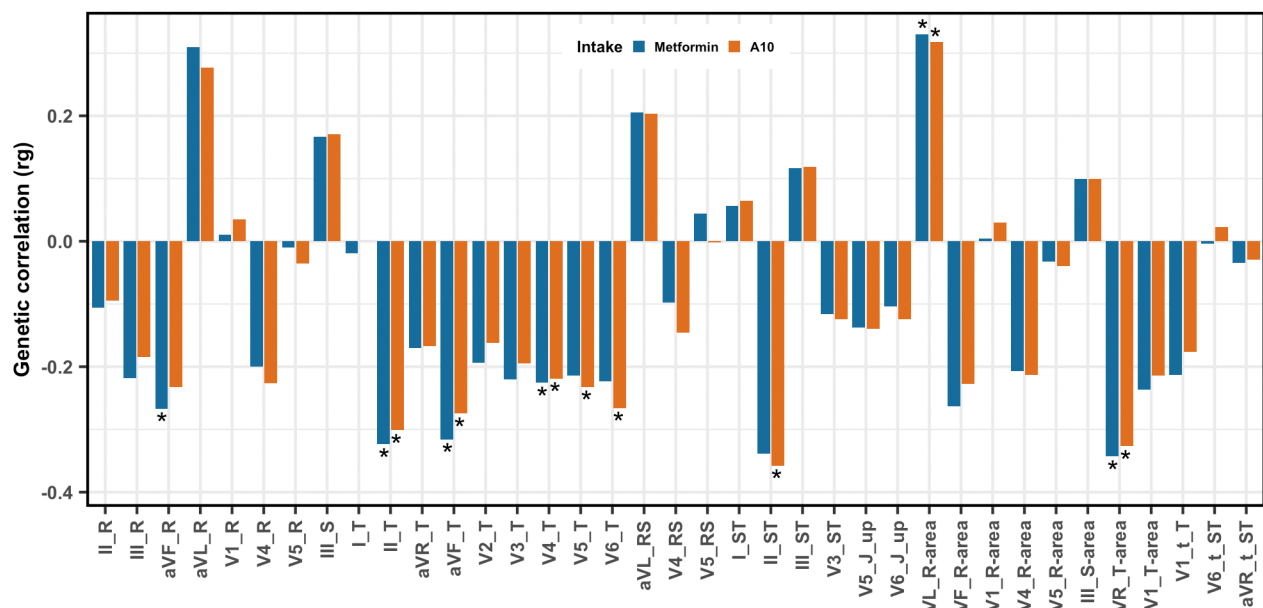
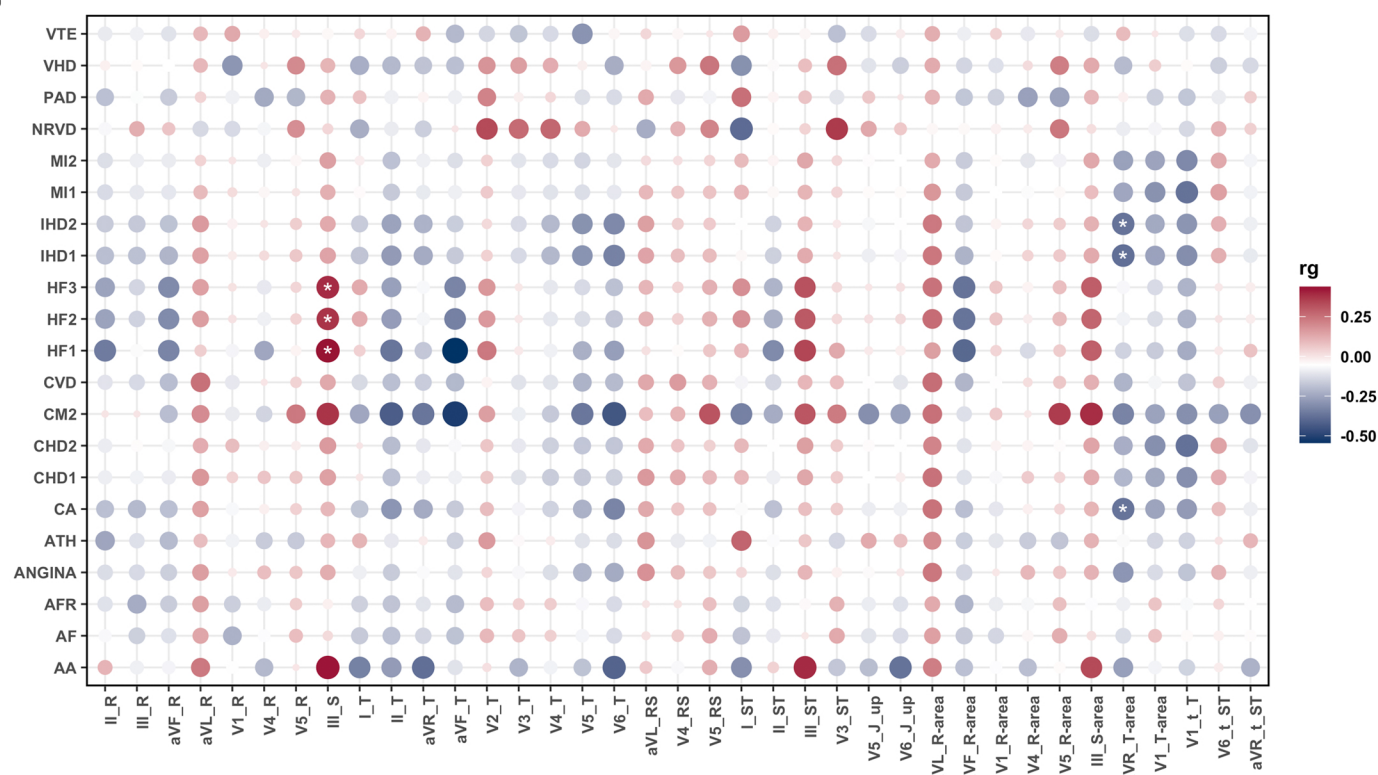
Putative causal relationship estimated by seven MR methods

Method	Exposure	Outcome	nSNP	P	Method	Exposure	Outcome	nSNP	P
MR Egger	Metformin	aVR_T-area	49	9.83e-02	MR Egger	A10	aVR_T-area	65	1.67e-01
Weighted median	Metformin	aVR_T-area	49	1.35e-02	Weighted median	A10	aVR_T-area	65	1.23e-02
IVW	Metformin	aVR_T-area	49	1.76e-03	IVW	A10	aVR_T-area	65	2.10e-04
Weighted mode	Metformin	aVR_T-area	49	2.73e-02	Weighted mode	A10	aVR_T-area	65	1.72e-02
GSMR	Metformin	aVR_T-area	48	1.06e-05	GSMR	A10	aVR_T-area	64	3.57e-06
MRlap	Metformin	aVR_T-area	49	2.14e-03	MRlap	A10	aVR_T-area	65	2.70e-04
CAUSE	Metformin	aVR_T-area	54	2.22e-02	CAUSE	A10	aVR_T-area	61	2.26e-01

A**B**

A**B****C****D**

A**B**

A**B****C**