# GWAS of retinal vessel tortuosity identifies 85 novel loci recovering variants associated with disease

Mattia Tomasoni[1,2], Ninon Mounier[2,3], Eleonora Porcu[2,4],

Tanguy Corre[1,2,3], Hana Abouzeid[5,6], Murielle Bochud[3], Sven Bergmann[1,2,7]

1 Dept. of Computational Biology, University of Lausanne, Lausanne, Switzerland
2 Swiss Institute of Bioinformatics, Lausanne, Switzerland
3 Center for Primary Care and Public Health (Unisanté), University of Lausanne, Lausanne, Switzerland
4 Center for Integrative Genomics, University of Lausanne, Lausanne, Switzerland
5 Division of Ophthalmology, Geneva University Hospitals, Switzerland
6 Clinical Eye Research Center Memorial Adolphe de Rothschild,Geneva, Switzerland
7 Dept. of Integrative Biomedical Sciences, University of Cape Town, Cape Town, South Africa

Corresponding authors:
mattia.tomasoni@unil.ch
sven.bergmann@unil.ch

## Abstract

Fundus pictures of the eye allow for non-invasive inspection of the microvasculature system of the retina which is informative on cardiovascular health. Automated image processing enables the extraction of morphometric properties of this system as quantitative features that can be used for modelling disease risks.

Here we report the results of the largest genome-wide association study (GWAS) of retinal vessel tortuosity conducted to date using data from the UK Biobank (*N*=63,899). We identified 87 loci associated with this trait (85 of which are novel). The heritability of the trait was $h^2$=0.23 (0.02). We carried out a replication study on a small independent population-based cohort, SKIPOGH (*N*=436). While the power of this study was too small to replicate individual hits, the effect size estimates correlated significantly between the two studies (Pearson correlation *r*=0.55, *p*=4.6E-6). Using LD score regression, we showed that the alleles associated with retinal vessel tortuosity point to a common genetic architecture of this trait with CVD and related traits.

Our results shed new light on the genetics of cardiovascular risk factors and disease.

# Table of contents

**Introduction** **2**

**Results** **2**

Automated processing of retina images defines retinal tortuosity phenotype 2

Retinal vessel tortuosity GWAS identifies 85 novel loci 2

Trend of effect sizes replicates in the SKIPOGH cohort 3

Tortuosity variants are associated with numerous diseases 3

Genetic signal is shared with hypertension and CVD 3

**Discussion** **4**

GWAS Analysis 4

**Materials and Methods** **5**

Definition of tortuosity 5

UK Biobank phenotypes 5

Data extraction 5

UK Biobank genotype data 6

The SKIPOGH study 6

Genome-wide association analysis 6

Heritability 7

Shared genetic architecture with disease 7

**Supplementary Material** **7**

**Figures** **8**

**Tables** **9**

**Author contributions** **11**

**Acknowledgements** **11**

**Funding** **11**

**References** **11**

## Introduction

The fundus of the eye is covered by blood vessels which are essential for bringing oxygen and nutrients to the various tissues of the retina. Fundus photography allows easy and non-invasive inspection of this retinal microvasculature, and it is well known that there are disease-related changes in the morphometric properties of the vessels. The fundus is of interest beyond the field of ophthalmology, since pathological changes in the retinal vessels often reflect those in microvasculature of other organs of major importance. Indeed, based on the homology between the microvasculature in the retina and that found in other organs, retinal analysis has the potential of becoming a powerful screening tool for diseases elsewhere in the body, notably the brain[1,2,3,4,5], kidney[6,7] and ear[8]. The retinal microvasculature can therefore provide signs of systemic disease, including increased risk of diabetes[9,10,11,12], obesity [13] and cardiovascular disease (CVD)[14,15,16,17,18], specifically stroke[16,19,19,20,21], coronary heart disease[22], coronary artery disease[23], hypertension[11,24,25,20,26,27,28,29,30,31,32,33,34], atherosclerosis[19,20,35] and myocardial infarction[36,37]. It is also informative of specific eye conditions such as Plus disease in the case of retinopathy of prematurity[38,39,38].

In recent years, measuring retinal features in large genotyped cohorts has paved the way for studying the genetic underpinning of these phenotypes and Genome Wide Association Studies (GWAS) have already identified a number of loci[40,41] associated with renal vessel size[42,4344], optic disc morphology[45,46] and vessel tortuosity[23].

In this study we carried out the largest GWAS on median retinal vessel tortuosity to date, confirming two known variants and discovering 85 new ones. Our discovery cohort was the UK Biobank, which provides a large collection of retinal images suitable for automatic analysis of morphometric properties of the vasculature of the human eye[47]. We used data from the much smaller, yet independent, population-based cohort SKIPOGH[48,49] for replication analysis. Many of the variants we identified with our GWAS have previously been associated with other traits, specifically CVD and some of its risk factors.

## Results

### Automated processing of retina images defines retinal tortuosity phenotype

We applied the ARIA[50] software for automated processing of 175,821 images from 63,899 individuals available in the UK Biobank. We modified ARIA to operate in batch mode, annotating the blood vessels in each image by extracting a list of points along the midline of each vessel. Using these data, we measured the tortuosity for each vessel (or annotated segment thereof) in terms of the so-called Distance Factor, i.e. the ratio between the path length along the vessel and the distance between its start and end point, as was first suggested in[51] used the median retinal vessel tortuosity over all annotated vessels (averaged over multiple images if available) as trait (see Methods for more details).

### Retinal vessel tortuosity GWAS identifies 85 novel loci

Applying linear regression of quantile-normalized median retinal vessel tortuosity on the genotypes of the matching subjects imputed to a panel of 15M genetic variants, we identified 6481 significantly associated SNPs (see Supplementary File 1). Applying LD pruning with a threshold of $R^2 < 0.01$ within a window of 500K bases to define independence, we obtained a list of 87 independent lead SNPs (the top 10 are listed in table 1, ordered by statistical significance, and a full list can be found in Supplementary File 2).

Two of the 6481 significant variants, namely rs1808382 and rs7991229, had previously been associated with retinal vessel tortuosity[23], while the remaining 85 SNPS represent novel loci associated with this trait (see figure 1 for a Manhattan plot of our genome-wide signals). According to the GWAS Catalogue, there is a third known locus for retinal vessel tortuosity, namely rs73157566. Yet, the association signal of this SNP was only marginally significant and it did not replicate in the replication cohort of the study which reported it[23]. We also did not replicate this finding (for details about these three variants and the respective genomic regions, see Supplementary Material on known associations with retinal vessel tortuosity).

We also performed LD score regression to estimate heritability, obtaining a $h^2 = 0.2293$ (SE = 0.0229). We did not observe any significant genomic inflation (slope = 1.0135, SE = 0.0103).

## Trend of effect sizes replicates in the SKIPOGH cohort

We attempted replication of our lead SNPs from the UKBB analysis in the SKIPOGH cohort[48,49] (436 individuals, multiple images per eye, for a total of 1,352 images). 60 out of the 87 lead SNPs were available for comparison. Given the limited sample size of the replication cohort, we lacked power to replicate the individual associations found in the discovery cohort, as none of them survived Bonferroni correction ($p = 0.05/60 = 8.3E-4$). Nevertheless, the effect size estimates using SKIPOGH data showed good concordance with those from the UKBB (see Supplementary File 3). First, 42 of 60 lead SNPs had the same sign of their effect size estimate in both studies (binomial test $p = 5.3E-4$). Second, we observed a Pearson correlation of $r = 0.55$ ($p = 4.6E-6$) across these estimates (see figure 2). Both results stay significant when removing outliers (see Supplementary Material, replication of effect sizes without outliers).

## Tortuosity variants are associated with numerous diseases

A shared genetic basis of retinal tortuosity and Coronary Artery Disease had already been noted for locus rs1808382 (mapped to the ACTN4/CAPN12 genes), underlining the usefulness of retinal vascular traits as biomarkers for cardiovascular diseases[23]. We replicated this finding and asked to what extent it also applies to the large panel of new variants we associated with retinal vessel tortuosity. Querying the GWAS Catalogue[52] for our hits revealed 9 loci linked to genes that had been reported as genome-wide significant in associations with other diseases including coronary heart disease, myocardial infarction, arterial hypertension, type 2 diabetes, chronic lymphocytic leukemia, Alzheimer's disease, diverticular disease, glaucoma and myopia (see table 2). Besides these 9 loci, we also uncovered 26 additional SNPs with pleiotropic effects on various diseases which could not be confidently mapped to a specific gene (see full list in Supplementary Material variants associated with disease outcome). We next expanded our query to include phenotypes known to confer a disease risk. We report a list of 12 loci linked to genes influencing both tortuosity and disease risk factors (see table 3). Furthermore, we uncovered another 9 SNPs showing similar pleiotropic properties, which could not be confidently mapped to a specific gene (see Supplementary Material).

SNP-level statistics were aggregated to obtain gene-wise association scores using the tool Pascal[53]. The results of our gene-wise association are summarized in figure 3: red squares mark disease genes, reported in table 2. Similarly, green squares indicate risk factor genes from table 3.

## Genetic signal is shared with hypertension and CVD

We extended our analysis of overlap with known genetic signal beyond variants with the same rsID, considering SNPs belonging to the same LD block ($R^2 > 0.8$). Figure 4 shows

how many of the variants identified by our retinal tortuosity GWAS had previously been reported as being associated with any of the large range of complex traits in the GWAS Catalogue[52]. A number of traits that stand out. Firstly, both diastolic (49) and systolic blood pressure (46) have been associated with many of our newly identified loci. Second, also pulse pressure and BMI share 19 associated loci. We note that both elevated blood pressure and BMI are well-known CVD risk factors, which we purposefully did not use as covariates in our GWAS, so as to be able to study the overlap in signal, even though correcting for these traits would have left the association signals largely unchanged (see Supplementary Material on Analysis of potential confounders). Furthermore, we observe a sizable number of tortuosity-associated variants overlapping with coronary artery disease variants, in line with what has recently been reported by a smaller scale GWAS on retinal vessel tortuosity[23]. For two additional phenotypes with a sizable overlap of trait associated variants, namely blood protein levels and bone mineral density, the relationship to tortuosity is less obvious, but might point to common pathogenic mechanisms. Similarly, for some of the other traits sharing several associated variants, notably, Type I and Type II diabetes, colorectal cancer, cholesterol, lung function, skin and eye pigmentation, and autoimmune diseases, there could be joint genetically modulated pathways, but some of these common associations may also just be spurious (see Supplementary File 4 for the full list of phenotypes and references to publications).

# Discussion

## GWAS Analysis

Adapting the retina image processing tool ARIA to run on a cluster facilitated the extraction of median retinal vessel tortuosity estimates for close 64 thousand subjects of the UK Biobank, enabling a GWAS for this trait with substantially increasing power compared to previous studies. This gain in power resulted in the identification of 85 novel loci and the replication of 2 out of 3 associations known from previous studies, providing a substantially improved picture of the genetic architecture of this trait.

We detected pleiotropic effects of 10 tortuosity variants associated with disease, specifically CVD-related diseases (Coronary Artery disease, Coronary Heart disease, Myocardial infarction, Hypertension), systemic diseases (diabetes, chronic lymphocytic leukemia, Alzheimer's disease) and ophthalmological conditions (myopia, glaucoma). Our results only link these diseases through common associated genetic variants to retinal vessel tortuosity. While one might speculate that this trait may reflect pathological developments of blood vessels that are causally upstream of some of these diseases, establishing such causal links will require more work, including the application of Mendelian randomization[54,55].

This study was subject to several limitations. First, our tortuosity measurements combine those of arteries and veins, while most ophthalmological studies distinguish between arterial and venular tortuosity. This compromise was made because we could not fully automate vessel classification. Indeed this is a difficult problem that still seems to require some expert input at least for some images or vessels, which would have prevented us from analysing such a large set of retinal images. Yet, the noise introduced by mixing arteries and veins, apparently was outweighed by the gain in sensitivity we achieved, as evidenced by the large number of associated loci. Subsequent studies using vessel annotations distinguishing their type may test whether these loci obtain different effect sizes for arterial and venular tortuosity. The second limitation of our study was that the software tool we used estimated tortuosity by the Distance Factor[51], a global measure which may not be ideal to capture vessel pathology (that may be better described by more local measures such as curvature[56], based on which, potentially more disease-relevant measures have been proposed[57,58]). We provide statistics about the distributions of vessel lengths in our dataset and repeated our GWAS using only a subset of relatively short vessel segments on which we measured Distance Factor tortuosity, but observed no dramatic change in the observed effect size

estimates of our top hits (see Supplementary Material on Tortuosity of short vessels). Also, we did not adjust for spherical equivalent refractive error, which might have confounded our measurements to some degree. The third limitation of our study was the small size of the replication cohort, which prevented us from replicating any individual hits. Nevertheless, the effects sizes in the replication study correlated strongly with those in the discovery cohort, providing independent evidence that they were not driven by any artifact specific to UKBB.

In conclusion, our highly powered GWAS on median retinal vessel tortuosity identified 85 novel loci in the maintenance of the microvasculature system, or failure thereof, as precursors or symptoms of complex diseases.

## Materials and Methods

### Definition of tortuosity

Our study assessed tortuosity using the measurements provided by the ARIA software[50]. We estimated tortuosity as the total vessel length divided by the euclidean distance between the vessel segment endpoints.

A number of measures have been designed to estimate vascular tortuosity. The measure adopted by the ARIA tool, in particular, is reported in a recent review as the AOC measure (Arc Over Chord ratio)[56]. In an earlier work on retinal vascular tortuosity, this measure was referred to as Tau 1[59] (an equivalent formulation in which a unit of one is subtracted). The measure was originally proposed (in the context of the femoral artery) by Smedby et al. as the Distance Factor[51]

### UK Biobank phenotypes

Our data was collected as part of the UK Biobank effort. The UK Biobank is a large-scale study that includes over half a million volunteers from the UK (502,505 participants, collection years 2006-2010). It includes a repeated assessment phase (20,000 participants, collection years 2012-2013). The age of participants ranges between 40 and 69 (median age 59), roughly balanced between sexes (229,122 males and 273,383 females).

175,821 fundus eye images (87,562 images of left eyes and 88,259 images of right eyes) were available at the time of required data extraction. We processed all images, including those from the reassessment time point. Other phenotypes were used to correct biases in the genetic associations (age, sex, PCs of genotypes) or to study correlation with disease and lifestyle (all other phenotypes). The following health statistics are of interest to interpret the medical implications of our analysis: 26,989 participants (5.3%) reported being diabetics, 16,787 (3.3%) reported being diagnosed with angina, 12,226 participants (2.4%) had a heart attack, 10,472 (2.0%) had deep-vein thrombosis (DVT, blood clot in leg), 8,216 (1.6%) were diagnosed with stroke, median DBP was 81 mm/Hg, median SBP was 136 mm/Hg, 227,360 (45.2%) received medication.

### Data extraction

The tortuosity phenotype measure was extracted using a modified version of the software ARIA by Peter Bunkehad[50]. We modified this tool to run in batch mode dumping vessel statistics to disk in the process, processing images without the need for human interaction. The ARIA parameters used for vessel extraction were the default applied by the software for tests to images from the database REVIEW[60].

We now describe the phenotype extraction quality control procedure. The ARIA tool was used to perform segmentation of blood vessels and measurement of both their tortuosity and diameters. The software is designed to perform vessel diameter measurements at regular

intervals along the centerline of each vessel, so that the number of measured diameters could be used as a proxy for the total length of the vascular system depicted in one image. Images for which less than 11 thousand equally spaced diameters could be measured were discarded. This threshold was (conservatively) set by visual inspection to discriminate lower quality images that were too dark, too light or out of focus. Roughly two out of three images passed this quality control, with a total of 120,363 images being sent forward in the pipeline.

Postprocessing of the data consisted in averaging the values derived from the left and right eye of each participant (for the resulting distribution, refer to Supplementary Material on Tortuosity estimates in UKBB)

The data extraction pipeline was written in python and bash and was run on a cluster using SLURM.

## UK Biobank genotype data

Around 488,000 participants were genotyped on Axiom arrays for a total of 805,426 markers. From this, about 96 million genotypes were imputed using a combined reference panel from the 1000 Genomes and UK10K projects[61]. The annotation used to report variant positions is the Genome Reference Consortium Human genome build 37 (known as GRChb37 or hg37). We subset the genotypes using the software BGENIX, shrinking the list of investigated variants to those that have been assigned an rsID (around 15 million SNPs). An additional Quality Control was performed via a postprocessing step on the GWAS output. We filtered out SNPs with MAF < 5E-4 (which given the sample size of 63,899 subjects, translates to having an average of over 30 individuals having at least one minor allele). We filtered out SNPs with imputation quality < 0.3, as used in Ref. [[62]].

## The SKIPOGH study

We performed replication of the GWAS results in the SKIPOGH cohort[48,49] (Swiss Kidney Project On Genes in Hypertension) which is a Swiss family-based population-based cohort that includes 1'042 participants (493 males and 549 females), aged between 18 and 96 years old, which have been extensively phenotyped at baseline and in a 3-year follow-up. Participants were recruited from three different locations in Switzerland, namely, Bern, Geneva and Lausanne. The genotyping was performed with the Illumina 2.5 omni chip, followed by an imputation based on HRC v1.1 panel using Minimach3. The annotation used to report variant positions is the Genome Reference Consortium Human genome build 37 (GRChb37).

## Genome-wide association analysis

The raw tortuosity measures extracted from the image data were transformed in order to correct for confounding effects that would bias the genetic association analysis (for an overview of the confounder analysis that we performed, please refer to Supplementary Material on Analysis of potential confounders). Only variables that showed a statistically significant correlation to tortuosity were corrected for. Specifically, we applied the linear model:

tortuosity ~ age + sex + genetic PCs.

A rank-based inverse normal transformation was applied to the residuals of this linear model and the GWAS was run on the output as a univariate linear regression without confounders (refer to Supplementary Material to inspect the resulting distributions).

The genetic association study was run using this software BGENIE[63]. The (unpruned) output consisted of 6481 significant SNPs (see Supplementary File 1).

The list of independent SNPs was calculated by performing LD pruning using the LDpair function of the R package LDlinkR[64], selecting as a reference panel "GBR" (Great Britain) from the 1k genome project. Two SNPs were considered as independent if they had LD $R^2 < 0.01$ or were more than 500K bases apart (see Supplementary File 2). This resulted in 87 independent lead SNPs. LD pruning was repeated with an alternative LD threshold of R2 < 0.1 (for comparison with other GWAS studies, where such value is used) resulting in a list of 124 significant SNPs (see Supplementary File 5).

The association analysis at replication stage for the SKIPOGH cohort was performed using the Emmax function of the Epacts software in order to account for family structure by using the kinship matrix in the model. Additionally, the recruitment center was included as covariable.

Summary plots were generated using the R packages qqman[65] and the GWASTools[66].

## Heritability

We carried out LD score regression using the software LDHub[67]. The portion of phenotypic variance cumulatively explained by the SNPs was $h^2$=0.2293 (0.0229). The measure of inflation was lambda_GC=1.1364; lambda GC measures the effect of confounding and polygenicity acting on the trait. The mean chi-square statistic mean_$X^2$=1.2941. The LD Score regression intercept was 1.0135 (0.0103); an intercept close 1 indicates little influence of confounders (mostly of population stratification). The ratio of the proportion of the inflation in the mean $X^2$ that is not due to polygenicity was 0.0459 (0.0352); a ratio close to 0 is desirable as it indicates low inflation from population stratification.

## Shared genetic architecture with disease

SNP variants overlap with disease phenotypes (same rsID) was analysed using the EBI's GWAS Catalogue[52]. We report independent SNPs in the tortuosity GWAS that are part of the GWAS Catalogue because associated to disease (or to a disease-related phenotype). This analysis was extended to genes and pathways using FUMA[68]. We list independent SNPs in the tortuosity GWAS who were in LD with SNPs that had already been reported in the GWAS Catalogue (see Supplementary File 4 and Supplementary File 6).

## Supplementary Material

Please refer to the Supplementary material file.

## Figures

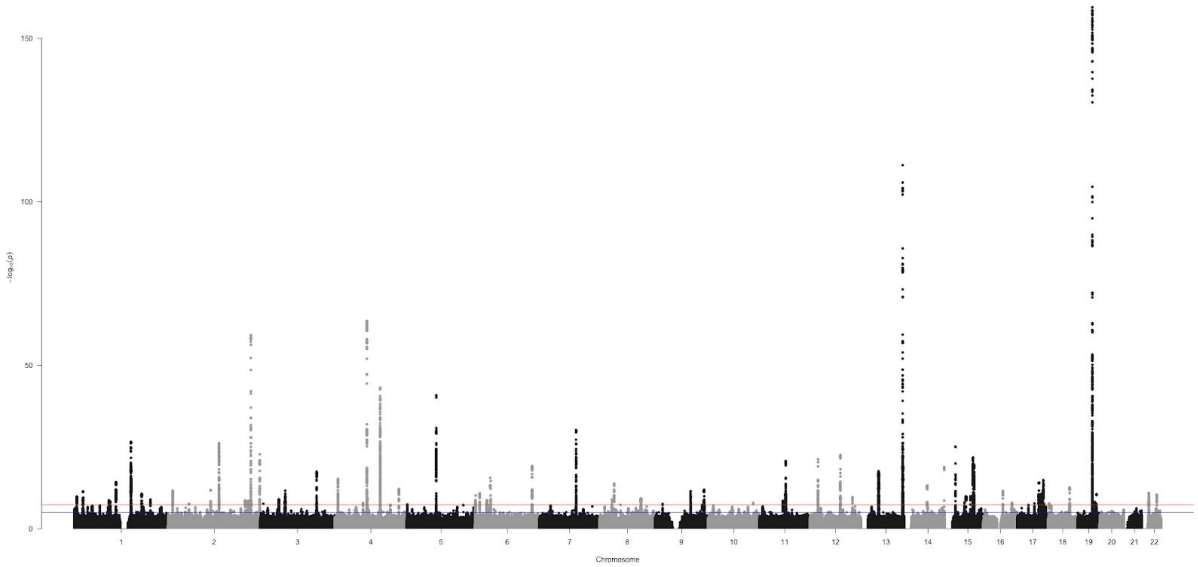


Figure 1: Manhattan plot of genome-wide association study for retinal vessel tortuosity corrected for phenotypic variables that showed a statistically significant association, i.e. age, sex, and a subset of principal components of genotypes (PCs: 1,2,5,6,7,8,16,17,18). Refer to Supplementary Material analysis for correlation with potential confounders. The red line is the genome-wide significance level (P = 5E-8). For a zoom of the genomic location, refer to Supplementary Material on known associations with retinal vessel tortuosity.

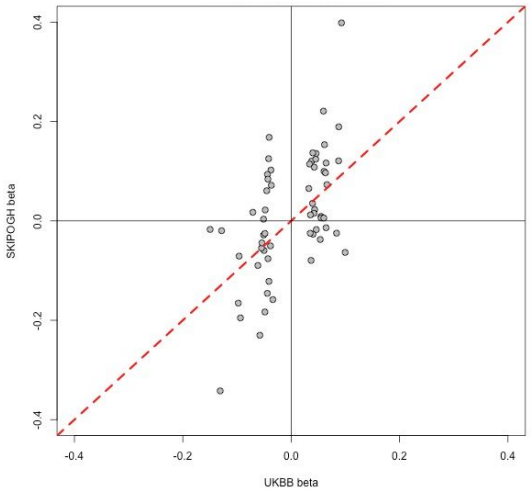


Figure 2: Statistically significant correlation between the measured effect sizes in the discovery cohort (UKBB, N=63,899) and replication cohort (SKIPOGH, N=436). We considered all lead (independent) SNPs in the UKBB. Of the 87, we could find 60 with matching rsIDs in SKIPOGH. The resulting correlation has Pearson r = 0.55 and P = 4.6E-6.

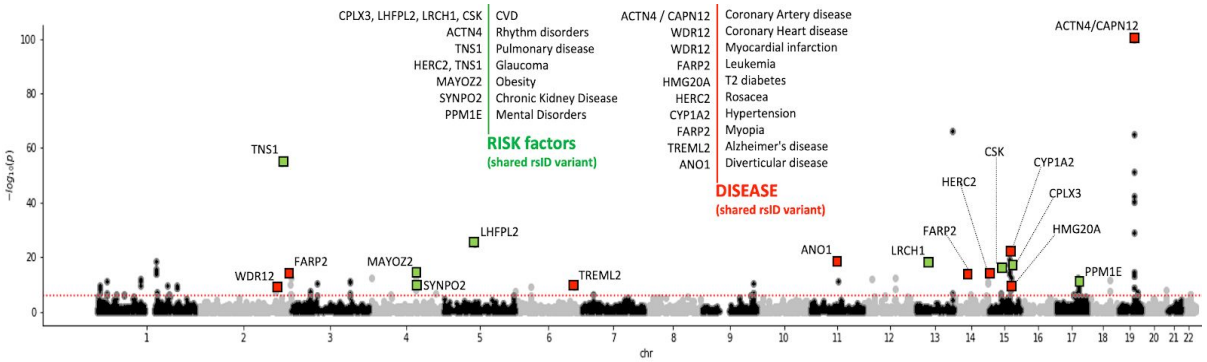


Figure 3: Gene-based Manhattan Plot. SNP-wise effect sizes were aggregated onto 21,722 protein coding genes. Genome-wide significance (dashed line in the plot) was defined at P = 0.05/21722 = 2.3E-6. Gene-based tests were computed by Pascal. Red squares indicate disease genes, to which we were able to map a SNP having pleiotropic effects on both retinal tortuosity and a disease (details are provided in table 2). Green squares indicate genes associated with disease risk (table 3).

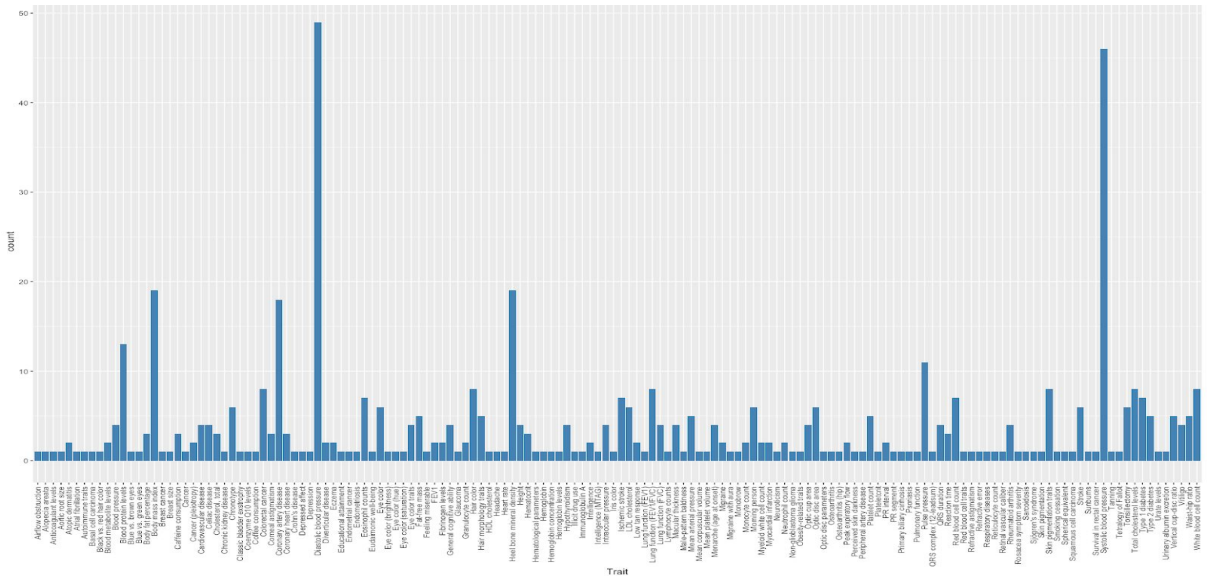


Figure 4: Number of variants shared with other traits reported in the GWAS Catalogue (SNPs in LD, R2 > 0.8). The largest peaks are blood pressure (SBP, DBP), followed by pulse pressure (i.e. SBP-DBP), BMI, coronary artery disease. Further traits with sizable number of shared associations are type I and type II diabetes, colorectal cancer, cholesterol, lung function, skin and eye pigmentation and autoimmune diseases.

## Tables

| CHR | SNP | BP | EA | NEA | FREQ | BETA | -log P |
|---|---|---|---|---|---|---|---|
| 19 | rs16972767 | 39153044 | G | A | 0.47369 | -0.14961 | 159.5 |
| 13 | rs9559797 | 111085411 | G | C | 0.58026 | -0.12837 | 111.2 |
| 4 | rs7661961 | 85029494 | A | G | 0.40333 | -0.096169 | 63.509 |
| 2 | rs2571461 | 218670945 | T | G | 0.60199 | -0.093652 | 59.24 |
| 4 | rs1026501 | 120013271 | G | C | 0.72423 | 0.087523 | 43.156 |
| 5 | rs784420 | 77987524 | A | G | 0.28180 | 0.083675 | 40.774 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 5 | rs341924 | 77944618 | G | T | 0.50123 | 0.065901 | 30.795 |
| 7 | rs13226484 | 96486765 | G | A | 0.56737 | 0.064379 | 30.21 |
| 1 | rs10788873 | 150250534 | T | A | 0.53124 | 0.060519 | 26.558 |
| 2 | rs1966628 | 134274625 | C | T | 0.39761 | -0.061464 | 26.128 |

Table 1: The 10 most significant retinal tortuosity lead SNPs, ordered by p-value. For full results, refer to list of 87 independent lead SNPs (Supplementary File 2). Abbreviations: CHR, Chromosome; BP, base pair position; EA, effect allele; NEA, non-effect allele; FREQ, allele frequency of effect allele; BETA, effect size estimate; Chromosomal positions are in GRCh37 coordinates.

| SHARED SNP | ref | DISEASE GWAS | GENE | GENE P-VALUE |
|---|---|---|---|---|
| rs6725887 (*) | [69] | Coronary heart disease | WDR12 | 5.62E-09 (1.03E-10) |
| rs6725887 (*) | [70] | Myocardial infarction | WDR12 | 5.62E-09 (1.03E-10) |
| rs936226 | [71] | Hypertension | CYP1A2 | 5.02E-20 (2.01E-13) |
| rs7119 | [72] | Type 2 diabetes | HMG20A | 1.21E-08 (3.85E-10) |
| rs757978 | [73] | Chronic lymphocytic leukemia | FARP2 | 6.24E-13 (6.22E-12) |
| rs9381040 | [74] | Alzheimer's disease | TREML2 | 8.33E-10 (3.33E-12) |
| rs875107 | [75] | Diverticular disease | ANO1 | 3.14E-19 (1.59E-23) |
| rs7588567 | [76] | Glaucoma | NCKAP5 | 0.033 (9.80E-18) |
| rs2753462 | [77] | Myopia | FARP2 | 6.24E-13 (6.22E-12) |

Table 2: List of variants identified by the UKBB tortuosity GWAS which were independently found to be associated with disease outcome in an independent study. We report only exact variants (same rsID in both tortuosity and disease GWAS). We report only variants which we could confidently map to a gene. Gene p-values were computed by using two tools, and both results are given for comparison: the first value was computer by Pascal [53] and the second (in parenthesis) by MAGMA [78]. Variants associated with more than one disease are marked with a star.

| SHARED SNP | ref | RISK FACTOR (diseases) | GENE | GENE P-VALUE |
|---|---|---|---|---|
| rs6495122 | [79] | Diastolic blood pressure (CVD) | CPLX3 | 6.88E-15 (4.14E-10) |
| rs1378942 (*) | [80] | Diastolic blood pressure (CVD) | CSK | 2.66E-15 (2.89E-20) |
| rs17355629 | [81] | Pulse pressure (CVD) | LRCH1 | 9.92E-19 (1.33E-19) |
| rs35252676 | [82] | Pulse pressure (CVD) | LHFPL2 | 1.44E-25 (2.03E-42) |
| rs9555695 | [83] | Waist-hip ratio (obesity) | COL4A2 | 4.42E-05 (1.80E-72) |
| rs7655064 | [83] | Waist-hip ratio (obesity) | MYOZ2 | 1.10E-12 (1.42E-12) |
| rs11083475 | [84] | Heart rate (rhythm disorders) | ACTN4 | 1.00E-100 (7.18E-63) |
| rs1378942 (*) | [85] | Mean arterial pressure (CVD) | CSK | 2.66E-15 (2.89E-20) |
| rs2571445 | [86] | Lung function (pulmonary disease) | TNS1 | 2.09E-55 (5.75E-60) |
| rs9303401 | [87] | Cognitive ability (mental disorders) | PPM1E | 1.16E-09 (2.72E-12) |
| rs17263971 | [88] | eGFR (Chronic Kidney Disease) | SYNPO2 | 3.19E-10 (5.85E-23) |
| rs3791979 | [89] | Intraocular pressure (open angle glaucoma) | TNS1 | 2.09E-55 (5.75E-60) |
| rs12913832 | [90] | Intraocular pressure (open angle glaucoma) | HERC2 | 3.00E-15 (1.82E-10) |

Table 3: List of variants identified by the UKBB tortuosity GWAS which were independently found to be associated with disease risk factors in an independent study. We report only

exact variants (same rsID in both tortuosity and disease GWAS). We report only variants which we could confidently map to a gene. Variants associated with more than one disease are marked with a star.

## Author contributions

MT, MuB and SB designed the study. MT performed tortuosity measurements of the raw image data from the UKBB and the SKIPOGH cohort and post-processed it. MT carried out the GWAS and subsequent bioinformatics analyses, with the guidance of SB, NM and EP. TC performed the replication analysis in SKIPOGH. HAZ provided medical and ophthalmological expertise. MiB helped in quality controlling the retina images. MT and SB wrote the manuscript.

## Acknowledgements

Thanks to Micha Hersch for inspiring this project to the UK Biobank team for their support and responsiveness, and to all UKBB participants for sharing their data.

## Funding

This work was supported by the Swiss National Science Foundation grant no. FN 310030_152724/1 to SB. The SKIPOGH study was supported by the Swiss National Science Foundation (#FN 33CM30-124087 to MuB).

Conflict of Interest statement: None declared.

## References

1. MacCormick, I. J. C., Czanner, G. & Faragher, B. Developing retinal biomarkers of neurological disease: an analytical perspective. *Biomark. Med.* **9**, 691–701 (2015).
2. Patton, N. *et al.* Retinal vascular image analysis as a potential screening tool for cerebrovascular disease: a rationale based on homology between cerebral and retinal microvasculatures. *J. Anat.* **206**, 319–348 (2005).
3. Liao, H., Zhu, Z. & Peng, Y. Potential Utility of Retinal Imaging for Alzheimer's Disease: A Review. *Front. Aging Neurosci.* **10**, 188 (2018).
4. Dumitrascu, O. M. & Qureshi, T. A. Retinal Vascular Imaging in Vascular Cognitive Impairment: Current and Future Perspectives. *J. Exp. Neurosci.* **12**, 1179069518801291 (2018).
5. Baker, M. L., Hand, P. J., Wang, J. J. & Wong, T. Y. Retinal signs and stroke: revisiting the link between the eye and brain. *Stroke* **39**, 1371–1379 (2008).
6. Sabanayagam, C. *et al.* A deep learning algorithm to detect chronic kidney disease from retinal photographs in community-based populations. *The Lancet Digital Health* **2**, e295–e302 (2020).
7. Park, H. C. *et al.* Diabetic retinopathy is a prognostic factor for progression of chronic kidney disease in the patients with type 2 diabetes mellitus. *PLoS One* **14**, e0220506 (2019).
8. Nash, B., Carlson, M. L. & Van Gompel, J. J. Microvascular decompression for tinnitus: systematic review. *J. Neurosurg.* **126**, 1148–1157 (2017).
9. Weiler, D. L., Engelke, C. B., Moore, A. L. O. & Harrison, W. W. Arteriole tortuosity associated with diabetic retinopathy and cholesterol. *Optom. Vis. Sci.* **92**, 384–391 (2015).
10. Gulshan, V. *et al.* Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA* **316**,

2402–2410 (2016).
11. Smith, W. *et al.* Retinal arteriolar narrowing is associated with 5-year incident severe hypertension: the Blue Mountains Eye Study. *Hypertension* **44**, 442–447 (2004).
12. Mookiah, M. R. K. *et al.* Application of different imaging modalities for diagnosis of Diabetic Macular Edema: A review. *Comput. Biol. Med.* **66**, 295–315 (2015).
13. Wang, J. J. *et al.* Retinal vessel diameters and obesity: a population-based study in older persons. *Obesity* **14**, 206–214 (2006).
14. Poplin, R. *et al.* Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nat Biomed Eng* **2**, 158–164 (2018).
15. Grunwald, J. E. *et al.* Retinopathy and the risk of cardiovascular disease in patients with chronic kidney disease (from the Chronic Renal Insufficiency Cohort study). *Am. J. Cardiol.* **116**, 1527–1533 (2015).
16. Ikram, M. K., Ong, Y. T., Cheung, C. Y. & Wong, T. Y. Retinal vascular caliber measurements: clinical significance, current knowledge and future perspectives. *Ophthalmologica* **229**, 125–136 (2013).
17. Flammer, J. *et al.* The eye and the heart. *Eur. Heart J.* **34**, 1270–1278 (2013).
18. Cuspidi, C., Sala, C. & Grassi, G. Updated classification of hypertensive retinopathy: which role for cardiovascular risk stratification? *Journal of hypertension* vol. 33 2204–2206 (2015).
19. Seidelmann, S. B. *et al.* Retinal Vessel Calibers in Predicting Long-Term Cardiovascular Outcomes: The Atherosclerosis Risk in Communities Study. *Circulation* **134**, 1328–1338 (2016).
20. Kawasaki, R. *et al.* Retinal vessel diameters and risk of hypertension: the Multiethnic Study of Atherosclerosis. *J. Hypertens.* **27**, 2386–2393 (2009).
21. Ikram, M. K. *et al.* Retinal vessel diameters and risk of stroke: the Rotterdam Study. *Neurology* **66**, 1339–1343 (2006).
22. Liew, G. *et al.* Fractal analysis of retinal microvasculature and coronary heart disease mortality. *Eur. Heart J.* **32**, 422–429 (2011).
23. Veluchamy, A. *et al.* Novel Genetic Locus Influencing Retinal Venular Tortuosity Is Also Associated With Risk of Coronary Artery Disease. *Arterioscler. Thromb. Vasc. Biol.* **39**, 2542–2552 (2019).
24. Wong, T. & Mitchell, P. The eye in hypertension. *Lancet* **369**, 425–435 (2007).
25. Cheung, C. Y.-L. *et al.* Retinal vascular tortuosity, blood pressure, and cardiovascular risk factors. *Ophthalmology* **118**, 812–818 (2011).
26. Grosso, A., Veglio, F., Porta, M., Grignolo, F. M. & Wong, T. Y. Hypertensive retinopathy revisited: some answers, more questions. *Br. J. Ophthalmol.* **89**, 1646–1654 (2005).
27. Wong, T. Y., Shankar, A., Klein, R., Klein, B. E. K. & Hubbard, L. D. Prospective cohort study of retinal vessel diameters and risk of hypertension. *BMJ* **329**, 79 (2004).
28. Wong, T. Y., Klein, R., Klein, B. E. K., Meuer, S. M. & Hubbard, L. D. Retinal vessel diameters and their associations with age and blood pressure. *Invest. Ophthalmol. Vis. Sci.* **44**, 4644–4650 (2003).
29. Dimmitt, S. B. *et al.* Usefulness of ophthalmoscopy in mild to moderate hypertension. *Lancet* **1**, 1103–1106 (1989).
30. Leung, H. *et al.* Impact of current and past blood pressure on retinal arteriolar diameter in an older population. *J. Hypertens.* **22**, 1543–1549 (2004).
31. Wang, J. J. *et al.* Hypertensive retinal vessel wall signs in a general older population: the Blue Mountains Eye Study. *Hypertension* **42**, 534–541 (2003).
32. Wong, T. Y. *et al.* Retinal arteriolar diameter and risk for hypertension. *Ann. Intern. Med.* **140**, 248–255 (2004).
33. Wong, T. Y. *et al.* Retinal microvascular abnormalities and blood pressure in older people: the Cardiovascular Health Study. *Br. J. Ophthalmol.* **86**, 1007–1013 (2002).
34. Ikram, M. K. *et al.* Retinal vessel diameters and risk of hypertension: the Rotterdam Study. *Hypertension* **47**, 189–194 (2006).

35. Sharrett, A. R. *et al.* Retinal arteriolar diameters and elevated blood pressure: the Atherosclerosis Risk in Communities Study. *Am. J. Epidemiol.* **150**, 263–270 (1999).
36. Woo, S. C. Y., Lip, G. Y. H. & Lip, P. L. Associations of retinal artery occlusion and retinal vein occlusion to mortality, stroke, and myocardial infarction: a systematic review. *Eye* **30**, 1031–1038 (2016).
37. Rim, T. H. *et al.* Retinal vein occlusion and the risk of acute myocardial infarction development: a 12-year nationwide cohort study. *Sci. Rep.* **6**, 22351 (2016).
38. Heneghan, C., Flynn, J., O'Keefe, M. & Cahill, M. Characterization of changes in blood vessel width and tortuosity in retinopathy of prematurity using image analysis. *Med. Image Anal.* **6**, 407–429 (2002).
39. Poletti, E., Grisan, E. & Ruggeri, A. Image-level tortuosity estimation in wide-field retinal images from infants with Retinopathy of Prematurity. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **2012**, 4958–4961 (2012).
40. Smith, W. *et al.* Retinal arteriolar narrowing is associated with 5-year incident severe hypertension: the Blue Mountains Eye Study. *Hypertension* **44**, 442–447 (2004).
41. Mookiah, M. R. K. *et al.* Application of different imaging modalities for diagnosis of Diabetic Macular Edema: A review. *Comput. Biol. Med.* **66**, 295–315 (2015).
42. Nannini, D. R. *et al.* A Genome-Wide Association Study of Vertical Cup-Disc Ratio in a Latino Population. *Invest. Ophthalmol. Vis. Sci.* **58**, 87–95 (2017).
43. Jensen, R. A. *et al.* Novel Genetic Loci Associated With Retinal Microvascular Diameter. *Circ. Cardiovasc. Genet.* **9**, 45–54 (2016).
44. Ikram, M. K. *et al.* Four novel Loci (19q13, 6q24, 12q24, and 5q14) influence the microcirculation in vivo. *PLoS Genet.* **6**, e1001184 (2010).
45. Springelkamp, H. *et al.* Meta-analysis of Genome-Wide Association Studies Identifies Novel Loci Associated With Optic Disc Morphology. *Genet. Epidemiol.* **39**, 207–216 (2015).
46. Han, X. *et al.* Genome-wide association analysis of 95 549 individuals identifies novel loci and genes influencing optic disc morphology. *Hum. Mol. Genet.* **28**, 3680–3690 (2019).
47. MacGillivray, T. J. *et al.* Suitability of UK Biobank Retinal Images for Automatic Analysis of Morphometric Properties of the Vasculature. *PLoS One* **10**, e0127914 (2015).
48. Pruijm, M. *et al.* Heritability, determinants and reference values of renal length: a family-based population study. *Eur. Radiol.* **23**, 2899–2905 (2013).
49. Ponte, B. *et al.* Reference values and factors associated with renal resistive index in a family-based population study. *Hypertension* **63**, 136–142 (2014).
50. Bankhead, P., Scholfield, C. N., McGeown, J. G. & Curtis, T. M. Fast retinal vessel detection and measurement using wavelets and edge location refinement. *PLoS One* **7**, e32435 (2012).
51. Smedby, O. *et al.* Two-dimensional tortuosity of the superficial femoral artery in early atherosclerosis. *J. Vasc. Res.* **30**, 181–191 (1993).
52. Buniello, A. *et al.* The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.* **47**, D1005–D1012 (2019).
53. Lamparter, D., Marbach, D., Rueedi, R., Kutalik, Z. & Bergmann, S. Fast and Rigorous Computation of Gene and Pathway Scores from SNP-Based Summary Statistics. *PLoS Comput. Biol.* **12**, e1004714 (2016).
54. Burgess, S., Small, D. S. & Thompson, S. G. A review of instrumental variable estimators for Mendelian randomization. *Stat. Methods Med. Res.* **26**, 2333–2355 (2017).
55. Smith, G. D. & Ebrahim, S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int. J. Epidemiol.* **32**, 1–22 (2003).
56. Abdalla, M., Hunter, A. & Al-Diri, B. Quantifying retinal blood vessels' tortuosity — Review. *2015 Science and Information Conference (SAI)* (2015)

doi:10.1109/sai.2015.7237216.
57. Trucco, E., Azegrouz, H. & Dhillon, B. Modeling the tortuosity of retinal vessels: does caliber play a role? *IEEE Trans. Biomed. Eng.* **57**, 2239–2247 (2010).
58. Lisowska, A., Annunziata, R., Loh, G. K., Karl, D. & Trucco, E. An experimental assessment of five indices of retinal vessel tortuosity with the RET-TORT public dataset. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **2014**, 5414–5417 (2014).
59. Hart, W. E., Goldbaum, M., Côté, B., Kube, P. & Nelson, M. R. Automated measurement of retinal vascular tortuosity. *Proc. AMIA Annu. Fall Symp.* 459–463 (1997).
60. Al-Diri, B. *et al.* REVIEW - a reference data set for retinal vessel profiles. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **2008**, 2262–2265 (2008).
61. Chou, W.-C. *et al.* A combined reference panel from the 1000 Genomes and UK10K projects improved rare variant imputation in European and Chinese samples. *Sci. Rep.* **6**, 39313 (2016).
62. Pistis, G. *et al.* Rare variant genotype imputation with thousands of study-specific whole-genome sequences: implications for cost-effective study designs. *Eur. J. Hum. Genet.* **23**, 975–983 (2015).
63. Bycroft, C. *et al.* Genome-wide genetic data on ~500,000 UK Biobank participants. *bioRxiv* 166298 (2017) doi:10.1101/166298.
64. Myers, T. A., Chanock, S. J. & Machiela, M. J. LDlinkR: An R Package for Rapidly Calculating Linkage Disequilibrium Statistics in Diverse Populations. *Frontiers in Genetics* vol. 11 (2020).
65. Turner, S. D. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *Bioinformatics* (2014).
66. Gogarten, S. M. *et al.* GWASTools: an R/Bioconductor package for quality control and analysis of genome-wide association studies. *Bioinformatics* **28**, 3329–3331 (2012).
67. Zheng, J. *et al.* LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics* **33**, 272–279 (2017).
68. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat. Commun.* **8**, 1826 (2017).
69. Mehta, N. N. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Circulation. Cardiovascular genetics* vol. 4 327–329 (2011).
70. Kathiresan, S. *et al.* Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. *Nat. Genet.* **41**, 334–341 (2009).
71. German, C. A., Sinsheimer, J. S., Klimentidis, Y. C., Zhou, H. & Zhou, J. J. Ordered multinomial regression for genetic association analysis of ordinal phenotypes at Biobank scale. *Genet. Epidemiol.* **44**, 248–260 (2020).
72. Sim, X. *et al.* Transferability of type 2 diabetes implicated loci in multi-ethnic cohorts from Southeast Asia. *PLoS Genet.* **7**, e1001363 (2011).
73. Slager, S. L. *et al.* Common variation at 6p21.31 (BAK1) influences the risk of chronic lymphocytic leukemia. *Blood* **120**, 843–846 (2012).
74. Lambert, J. C. *et al.* Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat. Genet.* **45**, 1452–1458 (2013).
75. Maguire, L. H. *et al.* Genome-wide association analyses identify 39 new susceptibility loci for diverticular disease. *Nat. Genet.* **50**, 1359–1365 (2018).
76. Osman, W., Low, S.-K., Takahashi, A., Kubo, M. & Nakamura, Y. A genome-wide association study in the Japanese population confirms 9p21 and 14q23 as susceptibility loci for primary open angle glaucoma. *Hum. Mol. Genet.* **21**, 2836–2842 (2012).
77. Tedja, M. S. *et al.* Genome-wide association meta-analysis highlights light-induced signaling as a driver for refractive error. *Nat. Genet.* **50**, 834–848 (2018).
78. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput. Biol.* **11**, e1004219 (2015).
79. Levy, D. *et al.* Genome-wide association study of blood pressure and hypertension. *Nat.*

*Genet.* **41**, 677–687 (2009).
80. Newton-Cheh, C. *et al.* Genome-wide association study identifies eight loci associated with blood pressure. *Nat. Genet.* **41**, 666–676 (2009).
81. Giri, A. *et al.* Trans-ethnic association study of blood pressure determinants in over 750,000 individuals. *Nat. Genet.* **51**, 51–62 (2019).
82. Warren, H. R. *et al.* Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk. *Nat. Genet.* **49**, 403–415 (2017).
83. Kichaev, G. *et al.* Leveraging Polygenic Functional Enrichment to Improve GWAS Power. *Am. J. Hum. Genet.* **104**, 65–75 (2019).
84. den Hoed, M. *et al.* Identification of heart rate-associated loci and their effects on cardiac conduction and rhythm disorders. *Nat. Genet.* **45**, 621–631 (2013).
85. Wain, L. V. *et al.* Genome-wide association study identifies six new loci influencing pulse pressure and mean arterial pressure. *Nat. Genet.* **43**, 1005–1011 (2011).
86. Wain, L. V. *et al.* Novel insights into the genetics of smoking behaviour, lung function, and chronic obstructive pulmonary disease (UK BiLEVE): a genetic association study in UK Biobank. *Lancet Respir Med* **3**, 769–781 (2015).
87. Seshadri, S. *et al.* Genetic correlates of brain aging on MRI and cognitive test measures: a genome-wide association and linkage analysis in the Framingham Study. *BMC Med. Genet.* **8 Suppl 1**, S15 (2007).
88. Stapleton, C. P. *et al.* The impact of donor and recipient common clinical and genetic variation on estimated glomerular filtration rate in a European renal transplant population. *Am. J. Transplant* **19**, 2262–2273 (2019).
89. Khawaja, A. P. *et al.* Genome-wide analyses identify 68 new loci associated with intraocular pressure and improve risk prediction for primary open-angle glaucoma. *Nature Genetics* vol. 50 778–782 (2018).
90. Craig, J. E. *et al.* Multitrait analysis of glaucoma identifies new risk loci and enables polygenic prediction of disease susceptibility and progression. *Nat. Genet.* **52**, 160–166 (2020).