Vascular endothelial growth factor (VEGF) is thought to have angiogenic and proatherosclerotic properties. Observationally VEGF is positively associated with cardiovascular disease (CVD), although the association may not be linear. However, it is unclear whether the association is due to confounding or reverse causation because VEGF may be a symptom of acute myocardial infarction. VEGF is pharmacologically modifiable, so assessment of its potential as a target of intervention for CVD prevention is important, and clarifying its etiological role may improve understanding of CVD. Trials have shown that statins lower VEGF, raising the possibility that one of the additional benefits of statins may be due to effects on VEGF. Little evidence from randomized controlled trials concerning the effects of VEGF on CVD in the general population is available. Previous randomized controlled trials among patients with ischemic disease did not provide strong evidence for clinical efficacy of VEGF.

Mendelian randomization studies use genetic predictors randomly allocated during conception, analogous to the randomization process in a randomized controlled trial. Genetic variants are unlikely to be affected by factors such as lifestyle or socioeconomic position, which commonly confound observational studies. Furthermore, genetic variants are unlikely to be affected by disease outcomes, and hence, Mendelian randomization is more resistant to reverse causation. As a result, this design may provide more credible evidence concerning the role of VEGF in CVD than observational studies. Mendelian randomization studies may also provide evidence more relevant to effects in the general population because randomized controlled trials are often

**Background**—Vascular endothelial growth factor (VEGF) has angiogenic and possibly proatherosclerotic properties. Observationally it is positively associated with cardiovascular disease, although these observations could be confounded or due to reverse causation. We assessed ischemic heart disease (IHD) risk by genetically predicted VEGF, ie, using Mendelian randomization.

**Methods and Results**—Single nucleotide polymorphisms (SNPs) predicting VEGF level, at genome-wide significance, were applied to the CARDioGRAMplusC4D 1000 Genomes-based genome-wide association study IHD case (n=60 801)-control (n=123 504) study. We obtained unconfounded estimates using instrumental variable analysis by combining the Wald estimates for each SNP using inverse variance weighting and Mendelian randomization–Egger regression. Based on 9 SNPs independently predicting VEGF (rs1740073 [C6orf223], rs2375981 [KCNV2], rs2639990 [ZADH2], rs4782371 [ZFP1M1], rs6921438 [LOC100132354], rs7043199 [VLDLR-AS1], rs10761741 [JMJD1C], rs6993770 [ZFP2M], and rs114694170 [MEF2C]), VEGF was unrelated to IHD (odds ratio 0.99, 95%CI 0.96-1.02) using inverse variance weighting. However, Mendelian randomization–Egger regression suggested an inverse relation of VEGF with IHD (odds ratio 0.95, 95%CI 0.91-0.99), although the association was not evident after excluding the lead SNP (rs6921438) or additionally excluding the pleiotropic SNP (rs6993770).

**Conclusions**—Our study does not provide strong evidence for a positive effect of VEGF on IHD but does not rule out the possibility that some specific types of VEGF, for which genetic predictors have not yet been identified, might play a role. (J Am Heart Assoc. 2017;6:e005619. DOI: 10.1161/JAHA.117.005619.)

**Key Words:** ischemic heart disease • Mendelian randomization • vascular endothelial growth factor
conducted in patient populations only. Meta-analysis of candidate gene studies examining the relation of VEGF-related genetic polymorphisms to coronary artery disease found no evidence supporting a causal effect of VEGF on ischemic heart disease (IHD), but these studies were of limited size (=630 participants in the analysis for IHD) and so were potentially underpowered as well as open to publication bias. To provide a more definitive answer about the role of VEGF in CVD, both as a target of intervention and from an etiological perspective, we conducted a Mendelian randomization study using the largest genome-wide association study (GWAS) to date of VEGF (n=16 112), applied to a large IHD case (n=60 801 for IHD)-control (n=123 504) study.

Methods

Study Design

This is a Mendelian randomization study, ie, an instrumental variable analysis with a genetic instrument, and has 3 key assumptions: (1) the genetic instrument predicts the exposure; (2) the genetic instrument is not associated with confounders of the exposure-outcome relation; and (3) the genetic instrument does not affect the outcome other than via its influence on the exposure of interest (exclusion restriction, ie, absence of horizontal pleiotropy). We attempted to address these assumptions first by choosing as genetic instruments for VEGF genetic variants that strongly predicted VEGF, second by using genetic instruments because genetic instruments are unlikely to be associated with confounders of the exposure-outcome relation, and third by checking whether the genetic variants predicting VEGF are known to have any horizontal pleiotropic effects on IHD, ie, effects on IHD via pathways that do not involve VEGF. In the context of Mendelian randomization, if an instrument’s pleiotropic effect is mediated via VEGF (ie, vertical pleiotropy), then it would not violate the instrumental variable assumption. Furthermore, there are also other assumptions such as negligible measurement error for the gene-exposure relation (no measurement error assumption), if the variance of the estimates does not take into account the variance of the genetic variants on exposure, which is likely to be satisfied in large samples.

Genetically Predicted VEGF

From a GWAS of VEGF, based on 16 112 adults of European ancestry with a mean age of 54.8 years using the 1000 Genomes reference data, single nucleotide polymorphisms (SNPs) that predicted VEGF (per log-transformed pg/mL) at genome-wide significance ($P<5 \times 10^{-8}$) were obtained. Correlations among these SNPs were evaluated from the $r^2$ for linkage disequilibrium obtained from SNP Annotation and Proxy Search (http://www.broad.mit.edu/mpg/snap/ldsearch.php) using 1000 Genomes (pilot 1, CEU) reference data. To rule out the possibility of violation of the exclusion restriction assumption by pleiotropic SNPs that affect IHD through exposures other than VEGF, we also cross-checked any other phenotypes these SNPs were associated with via a comprehensive genotype-to-phenotype cross-reference, Ensembl (Release 87) (http://www.ensembl.org/index.html). Ensembl gives any known traits associated with SNPs reported in another database such as the NHGRI-EBI GWAS catalog, which curates SNP-phenotype relations from published genome-wide association studies, usually using the conventional threshold for genome-wide significance of $5 \times 10^{-8}$. Here, traits known to cause IHD (eg, lipids, blood pressure) were considered as horizontal pleiotropic effects and violated the exclusion restriction assumption. However, we cannot rule out the possibility that such traits indicate vertical pleiotropy (ie, the genetic instrument has multiple traits that are all mediated via VEGF), which does not invalidate the exclusion-restriction assumption, or that the selected SNPs may have unmeasured/unexpected pleiotropic effects that can invalidate the inverse variance weighting analysis. We also conducted Mendelian randomization (MR)-Egger regression to reduce the risk of biases due to potentially inappropriate inclusion/exclusion of SNPs.

Genetic Predictors of Ischemic Heart Disease

Data on coronary artery disease and myocardial infarction have been contributed by CARDIoGRAMplusC4D investigators and have been downloaded from www.CARDIOGRAMPLUSC4D.ORG. CARDIoGRAMplusC4D 1000 Genomes-based GWAS is a meta-analysis of GWAS of IHD case-control studies of people of mainly European (77%), South Asian, and East Asian descent imputed using the 1000 Genomes phase 1 v3 training set with 38 million variants. The study interrogated 9.4 million
variants and included 60,801 IHD cases and 123,504 controls. Case status was defined by an inclusive coronary artery disease diagnosis such as myocardial infarction, acute coronary syndrome, chronic stable angina, or coronary stenosis >50%. These were ascertained in various ways, such as medical records, clinical diagnosis, procedures that indicate coronary artery disease, medications or symptoms that indicate angina, or self-reports as described elsewhere.

Statistical Analyses

Estimates of the effect of VEGF on IHD were obtained from separate sample instrumental variable analyses. We calculated SNP-specific Wald estimates and obtained the variance using the Feiler theorem. We used inverse variance weighting with fixed effects to combine the SNP-specific estimates for uncorrelated SNPs (ie, with $r^2<0.05$), which is a common approach for separate sample instrumental variable analyses using summary data from GWAS. From the analysis we reported the odds ratio per log-transformed increase in VEGF for IHD with a 95% CI. However, inverse variance weighting will give biased estimates if some of the instruments are invalid.

Sensitivity Analyses

MR-Egger Method

We conducted MR-Egger analysis, which will give an unbiased estimate even if all instruments are invalid (eg, presence of directional pleiotropy). However, MR-Egger only gives valid estimate if the Instrument Strength Independent of Direct Effect (InSIDE) assumption holds. Unfortunately, the assumption cannot be tested empirically because it would be violated if many genetic instruments influence the same unmeasured confounder of an exposure outcome relation. MR-Egger may also be susceptible to effect estimate dilution due to violation of a no-measurement-error assumption for instrument on exposure. To assess the degree of such dilution, we assessed the heterogeneity of the relation of genetic instruments on exposure ($I^2_{C_X}$) and adjusted the MR-Egger estimate using the simulation extrapolation (SIMEX) method if $I^2_{C_X}$ was less than 90%. Furthermore, a very low $I^2_{C_X}$ would suggest that MR-Egger may be less robust but nevertheless serve as an indicator for the validity of this test. From the MR-Egger regression, we also reported the intercept and the $P$ value, which indicate the presence of overall directional pleiotropy if $P$ value is $<0.05$.

Exclusion of Horizontal Pleiotropic SNPs/Lead SNP

We repeated the analyses, ie, inverse variance weighting and MR-Egger with the exclusion of SNPs exhibiting horizontal pleiotropy or the lead SNP, contributing more than 50% weight in the overall analysis, as additional sensitivity analyses to examine their impact on the overall estimate.

Power Calculation

We used the online calculator for power calculation of MR studies (http://cnsgenomics.com/shiny/mRnd/) to estimate power. In the original GWAS the variance explained by all 10 SNPs ranged from 19% to 52%. However, one of the SNPs (rs34528081) was excluded because it was not genotyped in CARDIoGRAMplusC4D 1000 Genomes and had no proxy. Assuming the remaining 9 SNPs explained only the lower bound of the total variance (ie, 19%), the sample size allows estimation of an effect size of odds ratio 0.97 for IHD per standard deviation of VEGF at 80% power with 5% significance.

All statistical analyses were conducted using R version 3.2.3 (R Foundation for Statistical Computing, Vienna, Austria) using the package “TwoSampleMR” from Github and scripts in the cited reference. The relevant scripts can be found in Data S1.

This analysis of publicly available data does not require ethical approval.

Results

Based on the most recent GWAS of VEGF, 10 SNPs reached genome-wide significance: rs1740073 (C6orf223), rs2375981 (KCNV2), rs2639990 (ZADH2), rs4782371 (ZFPM1), rs6921438 (LOC100132354), rs7043199 (VLDLR-AS1), rs10761741 (JMJD1C), rs6993770 (ZFPM2), rs114694170 (MEF2C), and rs34528081 (VEGFA), which explained up to 52% of the VEGF phenotypic variance. Among them, rs34528081 (VEGFA) was not genotyped in the CARDIoGRAMplusC4D 1000 Genomes–based GWAS, and no proxy for rs34528081 could be identified; hence, it could not be included in the analyses. The remaining 9 SNPs were largely uncorrelated, although rs1740073 was minimally correlated with rs6921438 ($r^2<0.001$) and rs2375981 with rs7043199 ($r^2=0.001$), so all 9 SNPs were used. The SNP rs6993770 (ZFPM2) was potentially pleiotropic because it is associated with platelets according to Ensembl ($P<4\times10^{-17}$) and so might affect IHD other than via VEGF. The lead SNP (rs6921438 in LOC100132354) contributed 82% of the weight in the inverse variance-weighting analysis. Information extracted concerning these SNPs is given in Table S1.

Table shows the MR estimates for VEGF on IHD using different methodological approaches. The inverse variance-weighted estimate showed no clear association of VEGF with IHD (odds ratio 0.99 per log-transformed pg/mL, 95% CI 0.96–1.02) using all 9 SNPs. Figure 1 shows that rs6921438
Vascular Endothelial Growth Factor and Heart

**Table.** Estimates of the Effect of Genetically Predicted VEGF (per Log-Transformed pg/mL)^10 on IHD\(^{11}\) Obtained From Mendelian Randomization Using Different Methodological Approaches and Exclusions for Pleiotropic SNPs

<table>
<thead>
<tr>
<th>Exclusion</th>
<th>Odds Ratio (95% CI)</th>
<th>Odds Ratio (95% CI)</th>
<th>Intercept (P Value)</th>
<th>(I^2_{GX})</th>
</tr>
</thead>
<tbody>
<tr>
<td>All 9 SNPs</td>
<td>0.99 (0.96 to 1.02)</td>
<td>0.95 (0.91 to 0.99)</td>
<td>0.016 (0.02)</td>
<td>99.8%</td>
</tr>
<tr>
<td>Excluding rs6993770 (ie, 8 SNPs)</td>
<td>0.98 (0.96 to 1.01)</td>
<td>0.95 (0.92 to 0.99)</td>
<td>0.012 (0.09)</td>
<td>99.8%</td>
</tr>
<tr>
<td>Excluding rs6921438 (ie, 8 SNPs)</td>
<td>1.08 (1.02 to 1.15)</td>
<td>1.06 (0.90 to 1.24)</td>
<td>0.003 (0.78)</td>
<td>97.1%</td>
</tr>
<tr>
<td>Excluding rs6993770 and rs6921438 (ie, 7 SNPs)</td>
<td>1.05 (0.98 to 1.13)</td>
<td>1.01 (0.85 to 1.20)</td>
<td>0.006 (0.63)</td>
<td>96.9%</td>
</tr>
</tbody>
</table>

\(I^2_{GX}\) indicates the percentage of the total variation that cannot be attributed to chance. 26

**Discussion**

This first MR study examining the relation of VEGF with IHD found little evidence for a causal role of VEGF in IHD, consistent with meta-analysis of candidate gene studies.\(^9\) Therefore, the positive relation of VEGF with CVD seen in observational studies is unlikely to be causal.\(^3\)

The function of VEGF is diverse, ranging from angiogenesis, vascular permeability, and tumorigenesis to possibly atherosclerosis,\(^6,20\) which in turn has led to the development of corresponding therapies for several diseases, including cancer, CVD, and the retinopathy of prematurity.\(^21\) Although VEGF-targeted therapies have had benefits in cancer and macular degeneration, albeit causing some side effects including hypertension and ocular inflammation,\(^20,22-25\) benefits are less apparent in the treatment of ischemia.\(^8\) The Framingham Heart Study suggested an inverted U-shaped relation of VEGF with CVD, which could reflect a complex action of VEGF on CVD or could be an artifact of confounding and reverse causation.\(^5\) Other smaller studies also showed conflicting results, where people with CVD did not always have higher VEGF.\(^26,27\) Given the largely null findings from our study, VEGF could possibly be a biomarker or symptom of CVD rather than a cause. Similarly, randomized controlled trials did not show that VEGF improved clinical outcomes, although the effects were studied primarily in ischemic patients.\(^6\) as it is typical of smaller clinical trials. Nevertheless, better understanding of VEGF, such as its role in angiogenesis, and how the drugs should be designed to bring about its intended effects, may help to explain the null results in previous trials.\(^28\)

Statins decrease cardiovascular risk more than would be expected from their effects on lipids, suggesting that statins may have multiple effects in addition to lowering low-density lipoprotein cholesterol.\(^29\) Although statins lower VEGF, our study suggests the additional benefits of statins in reducing CVD risk do not appear to be primarily due to statin influence on VEGF although this is speculative and could only be confirmed in other study designs such as a mechanistic randomized controlled trial.\(^5\) Previous studies primarily focused on the effects of VEGF-A although other classes of VEGF may have unknown effects relevant to CVDs.\(^1\) For example, VEGF-B may have cytoprotective properties, and animal studies have provided some evidence of clinical utility such as delayed dilated cardiomyopathy progression, whereas VEGF-C might be higher in those with ischemic cardiomyopathy.\(^1\) Further investigation of the cardiovascular effects of VEGF sub-types may provide additional insight into their etiologic role in CVDs with corresponding implications for drug development.\(^30\)

The most influential SNP in this MR study of VEGF on IHD is rs6921438 (LOC100132354), which is located downstream of VEGF and close to C6orf23 but encodes a currently uncharacterized protein.\(^31\) so it is difficult to know why it is relevant; rs6921438 might also be related to lipids, although the \(P\) values did not reach genome-wide significance (1.2 \times 10^{-7}) for high-density lipoprotein cholesterol and

\(LOC100132354\) was most influential in the analysis. However, the estimate was not substantially different when rs6921438 (LOC100132354) or rs6993770 (ZFPM2) or both SNPs were excluded. The MR-Egger estimate gave an inverse association of VEGF with IHD (odds ratio 0.95, 95%CI 0.91-0.99) based on 9 SNPs or with rs6993770 (ZFPM2) excluded, but the MR-Egger estimate was null after exclusion of rs6921438 (LOC100132354) or both SNPs. The MR-Egger intercept \(P\) value (\(P=0.02\) using all 9 SNPs) suggested directional pleiotropy, which is also reflected in the scatter-plot of the genetic association of outcome against genetic association of exposure for each genetic instrument included in this study (Figure 2). However, this was no longer evident after excluding rs6993770 (ZFPM2) (\(P=0.09\)), excluding rs6921438 (LOC100132354) (\(P=0.78\)), or excluding both rs6993770 (ZFPM2) and rs6921438 (LOC100132354) (\(P=0.63\)). \(I^2_{GX}\) was >90% in all analyses, suggesting that dilution of the MR-Egger estimate due to violation of the no-measurement-error assumption was limited.
1.5 × 10⁻⁴ for low-density lipoprotein cholesterol. Some of the gene regions including the other VEGF-related SNPs are associated with other phenotypes. IMID1C is related to liver function, platelet counts, sex-hormone-binding globulin, and androgen levels. ZFPM1 may be related to heart and coronary vessel development. MEFC2 is related to neurodevelopment. Because the SNPs used to predict VEGF lack definitive characterization of their full functional effects, we used MR-Egger regression and exclusion of potentially pleiotropic SNPs to reduce the likelihood of bias as sensitivity analyses. We found the inverse association of VEGF with IHD that was no longer evident using MR-Egger regression excluding rs6921438 (LOC100132354). The rs6921438 is very influential (Figure 1) and may have been driving any inverse association. Whether the potential inverse association of VEGF with IHD is meaningful awaits clarification of the functional role, or otherwise, of rs6921438 and the other SNPs predicting VEGF in IHD.

Although we used separate sample instrumental variable analysis with genetic instruments, which is less susceptible to residual confounding than observational studies, limitations exist. First, MR has stringent assumptions. We chose SNPs that strongly predicted VEGF in GWAS. Because genetic variants are randomly allocated during conception, the genetic variants are unlikely to be associated with potential confounders. We also used several MR techniques.
Vascular Endothelial Growth Factor and Heart  Au Yeung et al

including using MR-Egger regression, although we did not use a weighted median because 82% of the weight came from the lead SNP for VEGF (rs6921438 in LOC100132354) when a weighted median estimate gives consistent estimates only if more than 50% of the weight is from valid instruments.  

We also repeated the analyses excluding rs6921438 (LOC100132354) to check if the estimates were driven by this lead SNP. We searched comprehensively from genotyping to phenotype to identify potentially pleiotropic effects and excluded potentially horizontal pleiotropic SNPs to reduce the likelihood of biases due to violation of the instrumental variable analysis. However, we are limited by current knowledge and lack of access to summary statistics of all genome-wide association studies of cardiovascular risk factors to check for potential genetic associations, so we cannot exclude the possibility that our estimates are biased by currently unknown pleiotropic effects. Nevertheless, we conducted MR-Egger regression, which is more robust to the inclusion of invalid SNPs, and generally found no evidence for the relation of VEGF and IHD risk, with estimates close to null. Furthermore, we were unable to examine potential nonlinearity of VEGF on IHD because we only used summary statistics in this study, whereas the existing method for assessing nonlinearity in MR requires individual-level data. 

Nevertheless, this could be further explored in the UK Biobank once it accumulates enough IHD cases. Genomic control in the GWAS reduced the likelihood of confounding by population stratification. The SNPs included in this study were replicated, and we used estimates from the discovery and replication stages combined. Hence, the results should be less susceptible to the winner’s curse. However, a small proportion (5%) of the 184 305 participants in CARDIoGRAMplusC4D 1000 Genomes–based GWAS come from the VEGF GWAS (n=9548), and any resulting bias due to winner’s curse could underestimate the association between VEGF and IHD risk. Channeling bias is also unlikely because the genetic variants were randomly allocated at conception and hence should not be determined by other factors. Second, we assumed the genetic association with VEGF was present in CARDIoGRAMplusC4D 1000 Genomes–based GWAS, which is likely, as both studies mainly included adults of European descent. 

Third, we did not apply the VEGF SNPs to CARDIoGRAMplusC4D metaochip, as it only had 3 VEGF SNPs (rs6921438, rs6993770, and rs10761741). Based on these 3 VEGF-related SNPs, the association of VEGF with IHD was 1.00 (95%CI 0.96-1.05) using inverse variance weighting with fixed effects, and the MR-Egger regression estimate was 0.96 (95%CI 0.62-1.48) similar to the analyses using CARDIoGRAMplusC4D 1000 Genomes–based GWAS, although the MR-Egger regression estimates had a wider confidence interval. Nevertheless, differences in sample size and population characteristics between the 2 CARDIoGRAM GWAS might have contributed to any differences. Fourth, we were unable to include rs34528081 (VEGFA) in the analyses because it was not genotyped in the CARDIoGRAMplusC4D 1000 Genomes–based GWAS, and no proxy SNPs could be identified for it. VEGFA is the main genetic locus-determining serum VEGF, so we cannot rule out the possibility that this SNP could produce a different estimate for the effect of VEGF on IHD or that VEGF-A may have a different effect from other types of VEGF. Nevertheless, we have also included other SNPs, which may have better predictive power than VEGFA, based on the estimates for these SNPs on VEGF compared to the estimates for rs34528081 (VEGFA) on VEGF. Fifth, our study does not provide direct evidence on the role of VEGF in ischemia treatment, given that participants included in this study are not only patients suffering from ischemia. However, causal effects are usually consistent.

This MR study suggests that the observed positive association of VEGF with IHD is unlikely to be causal. Further MR studies using individual-level data may be useful to delineate any potential nonlinearity between VEGF and IHD risk and to identify the effects across classes of VEGF.

Acknowledgments

The authors would like to thank CARDIoGRAMplusC4D for making summary statistics publicly available.

Disclosures

None.

References


DOI: 10.1161/JAHA.117.005619
Vascular Endothelial Growth Factor and Heart


SUPPLEMENTAL MATERIAL
# R script for the paper (JAHA, 2017)
## Inverse variance weighting with fixed effect
rm(list = ls())
#install.packages("meta")
library(meta)

#Create dataframe for analysis (Post allele harmonization)
snp <- c("rs1740073", "rs2375981", "rs2639990", "rs4782371", "rs6921438", "rs6993770", "rs7043199", "rs10761741", "rs114694170")
bx <- c(0.09,0.21,0.11,0.07,0.64,0.16,0.1,0.08,0.15)
bxse <- c(0.01,0.01,0.018,0.011,0.008,0.01,0.013,0.009,0.023)
by <- c(0.012844,0.008435,0.013316,0.005059,-0.019603,0.027433,-0.005214,0.006922,0.001106)
byse <- c(0.0109920,0.0096472,0.0212769,0.0112289,0.0094876,0.0101545,0.0142791,0.0092442,0.0254852)
VEGF  <- data.frame(snp, bx, bxse, by, byse)

#All SNPs
# Calculating IV estimates manually using Feiller's theorem
VEGF$IV_estimate <- by/bx
VEGF$IV_SE <- sqrt(((by/bx)^2)*((byse^2/(by^2))+(bxse^2/(bx^2))))
results<-metagen(VEGF$IV_estimate, VEGF$IV_SE, sm="OR", VEGF$snp)
forest(results, leftlabs=c("SNP", "Logodds", "SE"), comb.random = FALSE, hetstat = FALSE)

#Excluding rs6993770
VEGF_1<-subset(VEGF, snp!="rs6993770")
results<-metagen(VEGF_1$IV_estimate, VEGF_1$IV_SE, sm="OR", studlab=VEGF_1$snp)

#Excluding rs6921438
VEGF_2<-subset(VEGF, snp!="rs6921438")
results<-metagen(VEGF_2$IV_estimate, VEGF_2$IV_SE, sm="OR", studlab=VEGF_2$snp)

#Excluding rs6993770 and rs6921438
VEGF_3<-subset(VEGF, snp!="rs6993770"&snp!="rs6921438")
results<-metagen(VEGF_3$IV_estimate, VEGF_3$IV_SE, sm="OR", studlab=VEGF_3$snp)

#Scatterplot
plot(VEGF$bx, VEGF$by, xlab="Genetic association with VEGF", ylab="Genetic association with IHD")
abline(lm(VEGF$by~VEGF$bx, weights = VEGF$byse^-2))

####MR-Egger
rm(list = ls())
#install.packages("devtools")
library(devtools)
#install_github("MRCIEU/MRInstruments")
#install_github("MRCIEU/TwoSampleMR")
library(TwoSampleMR)
library(MRInstruments)
#install.packages("plyr")
library(plyr)

#MR-Egger
# Create dataframe for analysis (Post harmonization)

**Exposure Dataframe**

```r
SNP <- c("rs1740073", "rs2375981", "rs2639990", "rs4782371", "rs6921438", "rs6993770", "rs7043199", "rs10761741", "rs114694170")
beta <- c(0.09, 0.21, 0.11, 0.07, 0.64, 0.16, 0.1, 0.08, 0.15)
se <- c(0.01, 0.01, 0.018, 0.011, 0.008, 0.01, 0.013, 0.009, 0.023)

VEGF <- data.frame(SNP, beta, se, effect_allele)
Exp_data <- format_data(VEGF, type="exposure")
```

**Outcome Dataframe**

```r
SNP <- c("rs1740073", "rs2375981", "rs2639990", "rs4782371", "rs6921438", "rs6993770", "rs7043199", "rs10761741", "rs114694170")
beta <- c(0.012844, 0.008435, 0.013316, 0.005059, -0.019603, 0.027433, -0.005214, 0.006922, 0.001106)
se <- c(0.0109920, 0.0096472, 0.0212769, 0.0112289, 0.0094876, 0.0101545, 0.0142791, 0.0092442, 0.0254852)

IHD <- data.frame(SNP, beta, se, effect_allele)
Out_data <- format_data(IHD, type="outcome")
```

```r
dat <- harmonise_data(
  exposure_dat = Exp_data,
  outcome_dat = Out_data,
  action=1
)
```

## All SNPs

```r
res <- mr(dat)
res$OR <- exp(res$b)
res$LCI <- exp(res$b - 1.96*res$se)
res$UCI <- exp(res$b + 1.96*res$se)
res
```

```r
egg.int <- mr_pleiotropy_test(dat) # Test for directional pleiotropy (MR-Egger)
```

```r
# I2
Isq = function(y, s)
  k = length(y)
  w = 1/s^2; sum.w = sum(w)
  mu.hat = sum(y*w)/sum.w
  Q = sum(w*(y-mu.hat)^2)
  Isq = (Q - (k-1))/Q
  Isq = max(0, Isq)
  return(Isq)
}
```

```r
Isq(dat$beta.exposure/dat$se.outcome, dat$se.exposure/dat$se.outcome)
```

## Excluding rs6993770

```r
dat_1 <- subset(dat, SNP!="rs6993770")
res <- mr(dat_1)
res$OR <- exp(res$b)
res$LCI <- exp(res$b - 1.96*res$se)
res$UCI <- exp(res$b+1.96*res$se)
res
gg.int<-mr_pleiotropy_test(dat_1) #Test for directional pleiotropy (MR-Egger)

#I2
Isq(dat_1$beta.exposure/dat_1$se.outcome,dat_1$se.exposure/dat_1$se.outcome)

#Excluding rs6921438
dat_2 <- subset(dat, SNP!="rs6921438")
res<-mr(dat_2)
res$OR <- exp(res$b)
res$LCL <- exp(res$b-1.96*res$se)
res$UCI <- exp(res$b+1.96*res$se)
res
gg.int<-mr_pleiotropy_test(dat_2) #Test for directional pleiotropy (MR-Egger)

#I2
Isq(dat_2$beta.exposure/dat_2$se.outcome,dat_2$se.exposure/dat_2$se.outcome)

#Excluding rs6993770 and rs6921438
dat_3 <- subset(dat, SNP!="rs6921438"&SNP!="rs6993770")
res<-mr(dat_3)
res$OR <- exp(res$b)
res$LCL <- exp(res$b-1.96*res$se)
res$UCI <- exp(res$b+1.96*res$se)
res
gg.int<-mr_pleiotropy_test(dat_3) #Test for directional pleiotropy (MR-Egger)

#I2
Isq(dat_3$beta.exposure/dat_3$se.outcome,dat_3$se.exposure/dat_3$se.outcome)
Table S1. Characteristics of single nucleotide polymorphisms (SNPs) used in the Mendelian Randomization analysis of the effect of vascular endothelial growth factor (VEGF) (per log transformed pg/ml)\(^1\) on ischemic heart disease (IHD)\(^2\)

<table>
<thead>
<tr>
<th>SNP</th>
<th>Effect allele/Non effect allele</th>
<th>Increase in exposure (log transformed pg/ml)</th>
<th>P value</th>
<th>Increase in log odds per effect allele (SE)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs1740073</td>
<td>T/C</td>
<td>0.09 (0.01)</td>
<td>4.4x10^{-17}</td>
<td>0.0128 (0.011)</td>
<td>0.24</td>
</tr>
<tr>
<td>rs2375981</td>
<td>C/G</td>
<td>0.21 (0.01)</td>
<td>9.49x10^{-99}</td>
<td>0.008 (0.0096)</td>
<td>0.38</td>
</tr>
<tr>
<td>rs2639990</td>
<td>T/C</td>
<td>0.11 (0.018)</td>
<td>5.85x10^{-10}</td>
<td>0.013 (0.021)</td>
<td>0.53</td>
</tr>
<tr>
<td>rs4782371</td>
<td>G/T</td>
<td>0.07 (0.011)</td>
<td>1.26x10^{-9}</td>
<td>0.005 (0.011)</td>
<td>0.65</td>
</tr>
<tr>
<td>rs6921438</td>
<td>G/A</td>
<td>0.64 (0.008)</td>
<td>1.66x10^{-1449}</td>
<td>-0.020 (0.009)</td>
<td>0.04</td>
</tr>
<tr>
<td>rs6993770</td>
<td>A/T</td>
<td>0.16 (0.01)</td>
<td>3.83x10^{-55}</td>
<td>0.027 (0.010)</td>
<td>0.007</td>
</tr>
<tr>
<td>rs7043199</td>
<td>T/A</td>
<td>0.10 (0.013)</td>
<td>4.16x10^{-14}</td>
<td>-0.005 (0.014)</td>
<td>0.72</td>
</tr>
<tr>
<td>rs10761741</td>
<td>T/G</td>
<td>0.08 (0.009)</td>
<td>2.99x10^{-19}</td>
<td>0.007 (0.009)</td>
<td>0.45</td>
</tr>
<tr>
<td>rs114694170</td>
<td>C/T</td>
<td>0.15 (0.023)</td>
<td>1.09x10^{-11}</td>
<td>0.001 (0.025)</td>
<td>0.97</td>
</tr>
</tbody>
</table>
Supplemental References:


Vascular Endothelial Growth Factor and Ischemic Heart Disease Risk: A Mendelian Randomization Study
Shiu Lun Au Yeung, Hugh Simon Hung San Lam and C. Mary Schooling

*J Am Heart Assoc.* 2017;6:e005619; originally published August 1, 2017;
doi: 10.1161/JAHA.117.005619

The *Journal of the American Heart Association* is published by the American Heart Association, 7272 Greenville Avenue,
Dallas, TX 75231
Online ISSN: 2047-9980

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://jaha.ahajournals.org/content/6/8/e005619

Subscriptions, Permissions, and Reprints: The *Journal of the American Heart Association* is an online only Open Access publication. Visit the Journal at [http://jaha.ahajournals.org](http://jaha.ahajournals.org) for more information.