



Choi, Y., Lee, S. J., Spiller, W., Jung, K. J., Lee, J-Y., Kimm, H., Back, J. H., Lee, S., & Jee, S. H. (2019). Causal associations between serum bilirubin levels and decreased stroke risk: A two-sample Mendelian randomization study. *Arteriosclerosis, Thrombosis, and Vascular Biology*, 40(2), 437-445.  
<https://doi.org/10.1161/ATVBAHA.119.313055>.

Publisher's PDF, also known as Version of record

License (if available):  
CC BY

Link to published version (if available):  
[10.1161/ATVBAHA.119.313055](https://doi.org/10.1161/ATVBAHA.119.313055).

[Link to publication record in Explore Bristol Research](#)  
PDF-document

This is the final published version of the article (version of record). It first appeared online via American Heart Association at <https://www.ahajournals.org/doi/10.1161/ATVBAHA.119.313055> . Please refer to any applicable terms of use of the publisher.

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:  
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>



# Causal Associations Between Serum Bilirubin Levels and Decreased Stroke Risk

## A Two-Sample Mendelian Randomization Study

Yoonjeong Choi,\* Sun Ju Lee,\* Wes Spiller, Keum Ji Jung, Ji-Young Lee, Heejin Kimm, Joung Hwan Back, Sunmi Lee, Sun Ha Jee

**OBJECTIVE:** A number of epidemiological studies have reported that decreased serum bilirubin, an endogenous antioxidant, is associated with cardiovascular disease. However, previous Mendelian randomization analyses conducted using a single sample have shown no evidence of association.

**APPROACH AND RESULTS:** A 2-sample summary Mendelian randomization study was performed by obtaining exposure and outcome data from separate nonoverlapping samples. We utilized data from the KoGES (Korean Genome and Epidemiology Study;  $n=25\,406$ ) and KCPS-II (Korean Cancer Prevention Study-II;  $n=14\,541$ ) biobank for serum bilirubin and stroke, respectively. Using KoGES, a total of 1784 single nucleotide polymorphisms associated with serum bilirubin levels were discovered using a genome-wide significance threshold ( $P < 5 \times 10^{-8}$ ), of which 10 single nucleotide polymorphisms were identified as independent ( $R^2 < 0.005$ ) and adopted as genetic instruments. From KCPS-II, total and ischemic stroke cases were identified ( $n=1\,489$  and  $n=686$ ), with 12366 acting as controls. Various 2-sample summary Mendelian randomization methods were employed, with Mendelian randomization estimates showing an inverse causal association between serum bilirubin levels and total stroke risk (odds ratio, 0.481 [95% CI, 0.234–0.988];  $P=0.046$ ). This association increased in magnitude when restricting the analysis to ischemic stroke cases (odds ratio, 0.302 [95% CI, 0.105–0.868];  $P=0.026$ ).

**CONCLUSIONS:** Our findings provide evidence of significant causal relationship between high levels of bilirubin and decreased stroke risk in Korean population in agreement with observational approaches. This highlights the potential for bilirubin to serve as a therapeutic target for oxidative stress-related diseases such as stroke and suggests that previous findings were not a consequence of unmeasured confounding.

**Key Words:** bilirubin ■ cardiovascular diseases ■ causality ■ polymorphism, single nucleotide ■ stroke

# FIRST PROOF ONLY

Serum circulating bilirubin is widely known as a secondary degradation product of heme, which is released from the breakdown of aging red blood cells.<sup>1</sup> Under physiological conditions, unconjugated plasma bilirubin concentrations in humans typically range from  $\approx 5$  to 17  $\mu\text{mol/L}$ .<sup>2</sup> Abnormally high concentrations of bilirubin are considered toxic and are linked to clinical symptoms of kernicterus<sup>3</sup> or neurological dysfunction due to the deposition of bilirubin in the brain.<sup>2</sup>

However, mild to moderately elevated serum bilirubin concentrations can serve as a potent endogenous antioxidant, through the process of bilirubin oxidizing itself to biliverdin.<sup>1,4</sup> This is supported by studies finding concentrations of total serum bilirubin  $< 7 \mu\text{mol/L}$  to be a risk factor for oxidative stress-related diseases in humans, such as cardiovascular disease, diabetes mellitus, metabolic syndrome, certain cancers, and autoimmune diseases.<sup>4</sup>

Correspondence to: Sun Ha Jee, PhD, MPH, Department of Epidemiology and Health Promotion, Graduate School of Public Health, Yonsei University, 50-1 Yonsei-ro, Seodaemun-gu, Seoul 03722, Republic of Korea. Email [jsunha@yuhs.ac](mailto:jsunha@yuhs.ac)

\*These authors contributed equally to this article.

The online-only Data Supplement is available with this article at <https://www.ahajournals.org/doi/suppl/10.1161/ATVBAHA.119.313055>.

For Sources of Funding and Disclosures, see page xxx.

© 2019 The Authors. *Arteriosclerosis, Thrombosis, and Vascular Biology* is published on behalf of the American Heart Association, Inc., by Wolters Kluwer Health, Inc. This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution, and reproduction in any medium, provided that the original work is properly cited.

*Arterioscler Thromb Vasc Biol* is available at [www.ahajournals.org/journal/atvb](http://www.ahajournals.org/journal/atvb)

## Nonstandard Abbreviations and Acronyms

<b>Chr</b>	chromosome
<b>EA</b>	effect allele
<b>GWAS</b>	genome-wide association study
<b>ICAM-1</b>	intercellular adhesion molecule 1
<b>IVW</b>	inverse-variance weighted
<b>KCPS-II</b>	Korean Cancer Prevention Study-II
<b>KoGES</b>	Korean Genome and Epidemiology Study
<b>Ldlr<sup>-/-</sup></b>	low-density lipoprotein receptor-deficient mice
<b>MR</b>	Mendelian randomization
<b>N SNPs</b>	the numbers of SNPs
<b>OR</b>	odds ratio
<b>SNP</b>	single nucleotide polymorphism
<b>VCAM-1</b>	vascular cell adhesion molecule 1
<b>VSMC</b>	vascular smooth muscle cell

Stroke is one such form of oxidative stress-related disease<sup>5-7</sup> and has been widely reported to be inversely associated with serum bilirubin levels in both Western<sup>8,9</sup> and Asian populations.<sup>10,11</sup> However, such findings could result from unmeasured confounding factors, and as a consequence may not represent a true causal relationship. One strategy to correct for bias due to unmeasured confounding is Mendelian randomization (MR), which uses genetic variants associated with an exposure of interest to estimate the effect of genetically predicted serum bilirubin on stroke risk. This has the potential to provide unbiased causal association estimates, provided genetic variants are associated with the exposure of interest (IV1), are independent of confounders of the exposure and outcome (IV2), and are independent of the outcome after conditioning on the exposure (IV3). Associations violating assumptions IV2 and IV3 are generally referred to as horizontal pleiotropic associations.

In previous work, we performed a single sample MR analysis within a Korean population (n=5599), with estimates suggesting a possible inverse association between serum bilirubin and total stroke risk (hazard ratio, 0.63 [95% CI, 0.30–1.33];  $P=0.229$ ).<sup>12</sup> However, the sample size proved insufficient to provide sufficient statistical power to identify a causal relationship between serum bilirubin and total stroke risk at conventional significance thresholds.

Our previous study provided motivation to conduct a 2-sample summary MR study, using data from the KoGES (Korean Genome and Epidemiology Study; n=25 406) and KCPS-II (Korean Cancer Prevention Study-II; n=13 855). By using 2 nonoverlapping samples, it has been possible to increase the sample size by a factor of 7 in comparison to our previous study. Further, adopting the 2-sample MR framework allows for several sensitivity

## Highlights

- This Mendelian Randomization study revealed that serum bilirubin level is causally associated with decreased stroke risk.
- A magnitude of association became stronger when the outcome is limited to ischemic stroke.
- Bilirubin may be suggested as a potential therapeutic target to the oxidative stress-related diseases, such as stroke.

analyses to be performed, critically examining the underlying assumptions of the MR approach (IV1-IV3). In this present study, we re-examine whether elevated serum bilirubin levels are causally associated with decreased stroke risk using 2-sample summary MR, interpreting the findings within the context of previous observational and MR studies.



## MATERIALS AND METHODS

The data that support the findings of this study are available from the corresponding author on reasonable request.

### Selection of the Genetic Instrumental Variables

Candidate genetic instruments for total serum bilirubin levels were selected from KoGES. KoGES is a community-based cohort study, collecting data on ≈210 000 individuals from 2005 to 2014. From the KoGES cohort, 33 598 participants were initially selected with both genetic and phenotypic data, comprising 13 108 men and 20 490 women aged from 40 to 77. Among the 33 598 participants with genetic data, 25 406 participants were finally selected who had measures of serum bilirubin levels. Serum total bilirubin levels (both direct and indirect bilirubin) were measured through automated biochemical profiling, and the unit of bilirubin concentration was milligram per deciliter. General characteristics of KoGES population are shown in Table I in the [online-only Data Supplement](#) and further details concerning the KoGES can be obtained from the previously published cohort profile.<sup>13</sup>

Genetic data were obtained using the Korea Biobank Array (referred to as Korean Chip) available through the K-CHIP consortium. Korean Chip, containing about 830 000 single nucleotide polymorphisms (SNPs) specific to the Korean population, was designed by the Center for Genome Science, Korea National Institute of Health, Korea (4845-301, 3000–3031).<sup>14</sup> Standard quality control procedures ( $P$  value for Hardy-Weinberg equilibrium  $\geq 1.0 \times 10^{-6}$ , call rate  $\geq 95\%$ , and INFO  $\geq 0.8$ ) were used. Genetic data were imputed using SHAPEIT v.2<sup>15</sup> and Minimac v.3,<sup>16</sup> utilizing the 1000 Genome reference panel phase 3, as well as 397 samples of Korean reference genome provided by the Center for Genome Science of Korea National Institute of Health.

We performed a genome-wide association study (GWAS) to identify SNPs associated with serum total bilirubin levels, adjusting for age and sex (minor allele frequency  $\geq 0.01$ ,  $P$  value for Hardy-Weinberg equilibrium test  $\geq 0.0001$ ). The GWAS of

bilirubin was conducted without any transformation. The results are presented in Figure 1 in the [online-only Data Supplement](#). Applying a Bonferroni multiple testing correction, a total of 1784 SNPs were identified to be associated with serum total bilirubin levels at a genome-wide significance level ( $P < 5 \times 10^{-8}$ ). As 2-sample MR analyses generally require SNPs to be independent, we applied a linkage disequilibrium clumping algorithm to identify independent SNPs using an  $R^2$  threshold  $< 0.005$ . After excluding correlated SNPs by clumping algorithm, a total of 10 SNPs were selected as genetic instruments for genetically determined serum bilirubin levels.

A replication of the bilirubin GWAS was also performed using the KCPS-II biobank ( $n=13855$ ), and the Human Gene Ontology names of the 10 SNPs were explored from NCBI or PhenoScanner database.<sup>17</sup> Cross-referencing identified SNPs using online data resources such as PhenoScanner is valuable in evaluating agreement with previous GWAS analyses, as well as potentially highlighting associations which may violate the underlying MR assumptions. However, it should be acknowledged that as such resources are largely representative of Western populations, the extent to which they are applicable to non-Western populations must be taken into consideration.

Potential horizontal pleiotropic associations, violating assumptions IV2 and IV3 of the MR approach, were explored from the PhenoScanner database.<sup>17</sup> We defined a SNP below GWAS significance ( $P < 5 \times 10^{-8}$ , proxies for East Asian:  $R^2 > 0.8$ ) as being potentially pleiotropic, if the related phenotype had previously been reported as a cardiovascular risk factor.

## Data Sources for the Study Outcome

Total and ischemic stroke were selected as outcome variables, using data from the KCPS-II biobank. KCPS-II is comprised of 156701 individuals who visited 18 health examination centers in the Seoul and Gyeonggi districts of Korea for a health check-up from 2004 to 2013. The participants were aged from 20 to 84, including 94840 men and 61861 women. Further details about KCPS-II biobank are presented in a previously reported cohort profile,<sup>18</sup> and general characteristics of KCPS-II population are displayed in Table 1 in the [online-only Data Supplement](#). As with the KoGES sample, genetic data were produced using Korean Chip, using identical quality control and imputation methods.

Total stroke was defined as all forms of stroke including ischemic, hemorrhagic, and all other stroke subtypes, including both prevalent and incident cases of stroke. *International Classification of Diseases-Tenth Revision* codes were used to define stroke and stroke subtype cases, specifically: ischemic stroke, I63–I639; hemorrhagic stroke, I60–I629; and all stroke types, I60–I699.

Stroke cases were determined using hospital admission discharge records from 2005 to 2015 (median follow-up duration, 8.0 years), available from health insurance claims provided by the National Health Insurance Service.<sup>10</sup> A total of 1489 total stroke cases were identified (654 ischemic stroke, 340 hemorrhagic stroke, 32 ischemic-hemorrhagic combined stroke, and 463 all other stroke subtypes), and 686 ischemic stroke (654 ischemic stroke and 32 ischemic-hemorrhagic combined stroke). To conduct a GWAS of stroke risk, 12366 controls were used, fitting a logistic regression model adjusting for age and sex (minor allele frequency  $\geq 0.01$ ,  $P$  value for

Hardy-Weinberg equilibrium test  $\geq 0.0001$ ). Log odds for stroke risk were then used in subsequent MR analyses.

Signed written consent was obtained from all participants, and The Institutional Review Board of Yonsei University approved this study protocol (IRB approval number 4-2014-1008).

## Statistical Analysis

For each individual genetic instrument, an MR estimate can be obtained by dividing the instrument-outcome association (KCPS-II) by the instrument-exposure association (KoGES). This is referred to as a Wald ratio estimate, using association estimates obtained from GWASs. In this study, we specifically consider inverse variance weighted (IVW), MR-Egger, and weighted median regression approaches, which utilize the set of Wald ratio estimates provide more accurate estimates of causal effect, as well as test and correct for bias due to horizontal pleiotropic pathways.

The IVW estimate is obtained by regressing the set of instrument-outcome associations upon the instrument-exposure associations omitting an intercept and weighting each estimate by the inverse of the variance of the instrument-outcome association. This represents a weighted average of Wald ratio estimates. In the case of MR-Egger regression, an intercept is included, interpreting the intercept as an average horizontal pleiotropic effect across the set of genetic instruments and the slope of the regression as the corrected causal effect.<sup>19</sup>

It should be noted that as an additional parameter is estimated, MR-Egger is statistically less powerful than IVW. MR-Egger regression is also reliant on the magnitude of pleiotropic effects being independent of the strength of instrument-exposure association. Finally, weighted median regression essentially calculates a weighted median using the Wald ratio estimates, with the weighting corresponding to the inverse variance weights selected. This has the advantage of being robust to bias due to horizontal pleiotropic bias, provided the proportion of nonpleiotropic SNPs exceeds 50% with respect to their corresponding weighting.

Genetic instruments are required to be strongly associated with the exposure of interest (IV1), with the strength of association quantified using the  $F$ -statistic for the genetic instrument in a regression of the exposure on the instrument. We calculate  $F$ -statistics adopting the method put forward by Au Yeung et al.<sup>20</sup> The mean  $F$ -statistics for the 10 genetic instruments was 168, satisfying the threshold of  $F > 10$ , typically recommended for MR analyses.<sup>21</sup> For reference, an  $F$ -statistic of 10 represents a relative bias of 10% toward the null in a 2-sample MR setting.

Initially, Wald ratio estimates using each of the 10 genetic instruments were constructed by dividing the log odds of stroke risk by the  $\beta$  coefficient of bilirubin levels, obtained from GWASs of KCPS-II and KoGES. With this complete, IVW, MR-Egger, and weighted median regression models were applied, interpreting the estimated association as the effect of a genetically determined 1 mg/dL increase in serum bilirubin levels on the stroke risk.<sup>22</sup> Additionally, we performed leave-one-out analyses by systematically removing each SNP in turn and repeating the IVW analysis to assess if a single SNP is driving the association.<sup>23</sup> For the MR-Egger regression, a significant nonzero intercept is considered evidence of horizontal pleiotropic bias ( $P$  value for MR-Egger intercept  $< 0.05$ ), and where this is not



the case, the IVW estimate is preferable owing to the statistical power limitations of MR-Egger regression.

To improve the visualization of the IVW and MR-Egger estimates, we performed radial variants of the IVW and MR-Egger models. These models are similar to the conventional IVW and MR-Egger regression models, but regress the product of the Wald ratio estimate and the square root of the weighting for each genetic variant upon the square root of the genetic variants weighting. For radial IVW, an intercept is omitted in contrast to radial MR-Egger, as in the conventional setting. This has the advantage that genetic instruments which are subject to pleiotropic bias can be identified as outliers, based on their distance from the regression line. The distance of each point to the regression line is proportional to the individual contribution to global heterogeneity in effect estimates for the given genetic instrument.<sup>24</sup>

All MR analyses were calculated using the TwoSampleMR<sup>25</sup> and RadialMR<sup>24</sup> R packages, using R, version 3.5.1 (R Development Core Team, Vienna, Austria).

## RESULTS

### Genetic Instrumental Variables for Bilirubin

The associations for each genetic instrument with serum bilirubin levels and stroke risk are presented in Table 1, and all genetic instruments were replicated within KCPS-II biobank participants (Table II in the [online-only Data Supplement](#)). Of the 10 genetic instruments, only rs2119503 had a readily identifiable pleiotropic association, having previously been reported to be associated with Crohn disease in European populations ( $P=1.94\times 10^{-12}$ ). A proxy variant for rs2119503, rs6758317,<sup>26</sup> was also associated with Crohn disease in East Asian populations ( $P=4.50\times 10^{-21}$ ).<sup>27</sup> Although all instrumental variables were located on the chromosome 2 or 12, there was no overlap in terms of overarching gene. The associations of stroke risk for each genetic variant are shown in Table 1 on the log odds scale.

### Mendelian Randomization Analyses for Stroke Risk

IVW, MR-Egger, and weighted median regression were used to estimate causal associations between genetically predicted serum bilirubin levels and stroke risk (Table 2; Figure 1). Using IVW, we found evidence of causal relationships between bilirubin levels and stroke risk. Serum bilirubin levels seem to be negatively associated with both total and ischemic stroke risk (total stroke: IVW odds ratio [OR] per 1 mg/dL increase in serum bilirubin=0.481 [95% CI, 0.234–0.988];  $P=0.046$ ; Ischemic stroke: IVW OR per 1 mg/dL increase in serum bilirubin=0.302 [95% CI, 0.105–0.868];  $P=0.026$ ). MR-Egger and weighted median regression also showed directionally similar estimates, though they seem insufficiently powered to exceed conventional significance thresholds (total stroke: MR-Egger OR per 1 mg/dL increase in serum bilirubin, 0.534 [95% CI, 0.125–2.286];  $P=0.422$ ; weighted median OR per 1 mg/dL increase in serum bilirubin, 0.469 [95% CI, 0.174–1.262];  $P=0.134$ ; ischemic stroke: MR-Egger OR per 1 mg/dL increase in serum bilirubin, 0.320 [95% CI, 0.037–2.781];  $P=0.332$ ; weighted median OR per 1 mg/dL increase in serum bilirubin, 0.430 [95% CI, 0.102–1.821];  $P=0.252$ ).

Funnel plots display the individual Wald ratios for each SNP plotted against their precision, where asymmetry is indicative of directional horizontal pleiotropy. It should be noted, however, that assessing funnel plots with respect to symmetry is difficult when using a small number of genetic instruments (Figure 2). The MR-Egger intercepts show no evidence for significant directional pleiotropy, for both total and ischemic stroke ( $P=0.877$  and  $P=0.954$ , respectively; Table 2). These results suggest that directional pleiotropic effects are not present between bilirubin levels and either total or ischemic stroke risk.

**Table 1. List of Genetic Instruments for Serum Bilirubin Levels and Log Odds Ratios of Stroke Risk by Each Instrumental SNPs (GWAS Significance with  $P<5\times 10^{-8}$  and linkage disequilibrium threshold with  $R^2<0.005$ )**

No.	SNP	Gene	Chr.	EA	OA	EAF	Bilirubin $\beta$ (SE)	Total stroke log <sub>e</sub> OR (SE)	Ischemic stroke log <sub>e</sub> OR (SE)
1	rs12993249	<i>USP40</i>	2	G	A	0.378	0.060 (0.003)	0.016 (0.041)	0.007 (0.060)
2	rs12996139	<i>MSL3P1</i>	2	A	C	0.104	0.033 (0.004)	-0.002 (0.065)	-0.022 (0.094)
3	rs187332311	<i>UGT1A12P</i>	2	A	G	0.024	0.058 (0.009)	-0.033 (0.131)	-0.175 (0.201)
4	rs2119503*	<i>ATG16L1</i>	2	A	G	0.117	0.053 (0.004)	0.026 (0.062)	0.023 (0.088)
5	rs4149014	<i>SLCO1B1</i>	12	G	T	0.228	-0.024 (0.003)	0.040 (0.047)	0.036 (0.067)
6	rs4663344	<i>TRPM8</i>	2	T	G	0.380	-0.015 (0.003)	0.043 (0.041)	0.079 (0.059)
7	rs55686299	<i>MROH2A</i>	2	G	T	0.065	0.048 (0.005)	0.010 (0.081)	-0.016 (0.118)
8	rs55986512	-	12	A	G	0.089	0.062 (0.005)	-0.171 (0.075)	-0.233 (0.111)
9	rs6723506	<i>UGT1A1</i>	2	G	A	0.027	0.169 (0.008)	-0.226 (0.137)	-0.445 (0.214)
10	rs73233620	<i>SLCO1B3</i>	12	G	T	0.199	0.057 (0.003)	-0.072 (0.051)	-0.065 (0.074)

Chr indicates chromosome; EA, effect allele; OA, other allele; EAF, effect allele frequency; and OR, odds ratio.

\*rs2119503 previously had known potentially pleiotropic effect related to Crohn disease.

**Table 2. Mendelian Randomization for Serum Bilirubin on Stroke Risk**

	Total Stroke			Ischemic Stroke		
	OR Estimates* (95% CI)	P-Value	N SNPs	OR Estimates† (95% CI)	P-Value	N SNPs
IVW	0.481 (0.234–0.988)	0.046	10	0.302 (0.105–0.868)	0.026	10
MR-Egger regression	0.534 (0.125–2.286)	0.422	10	0.320 (0.037–2.781)	0.332	10
Intercept†	−0.006 (0.037)	0.877		−0.003 (0.055)	0.954	
Weighted Median	0.469 (0.174–1.262)	0.134	10	0.430 (0.102–1.821)	0.252	10

IVW indicates inverse-variance weighted; N SNPs, the numbers of SNPs; and OR, odds ratio.

\*Odds ratio per 1 mg/dL increase in serum total bilirubin.

†Intercept is presented as  $\beta$  coefficients with SEs.

## Effects of Individual Genetic Instruments in Relation to Stroke Risk

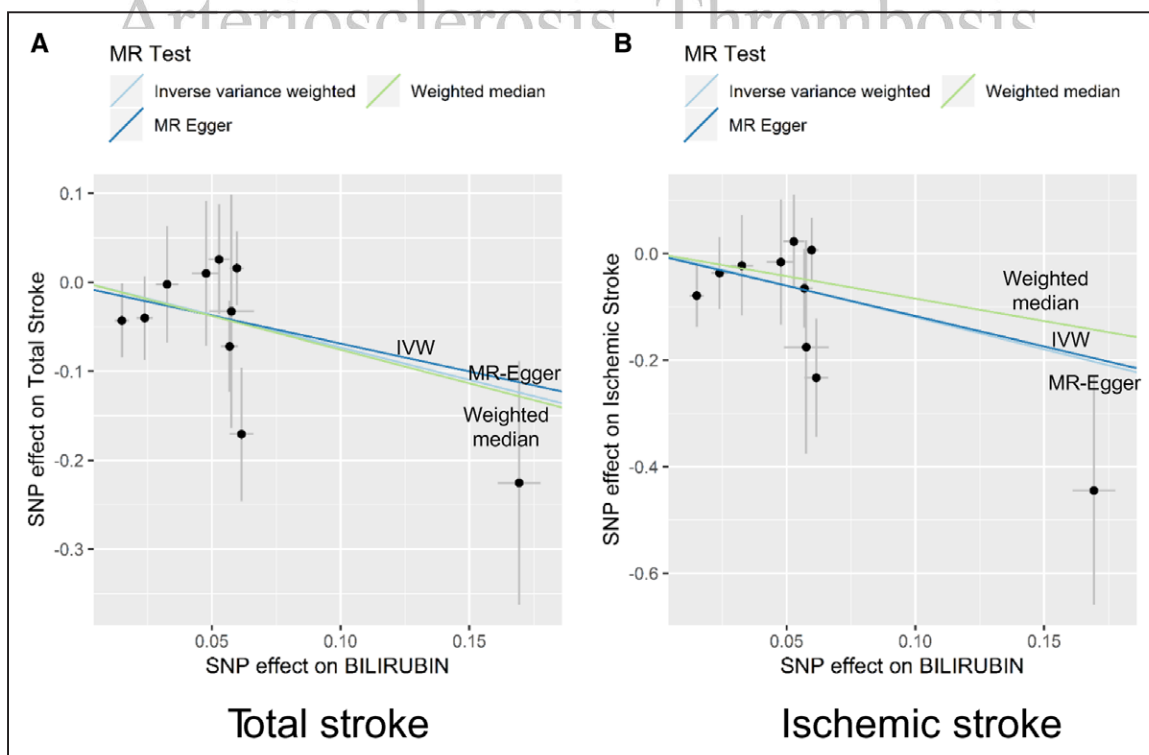
To compare the influence of individual SNPs on the overall causal estimate, we performed leave-one-out analyses. When we systematically removed each SNP and repeated the MR analyses, there did not appear to be a substantial difference in estimated causal effect. It is therefore likely that estimated effects are not the result of a single genetic instrument (Figure II in the [online-only Data Supplement](#)).

When using radial IVW and MR-Egger approaches, there did not appear to be evidence of outlying genetic variants. This result is in agreement with the leave-one-out and MR-Egger analyses, which provide no indication of pleiotropic effects which could bias the overall analyses (Figure 3).

## DISCUSSION

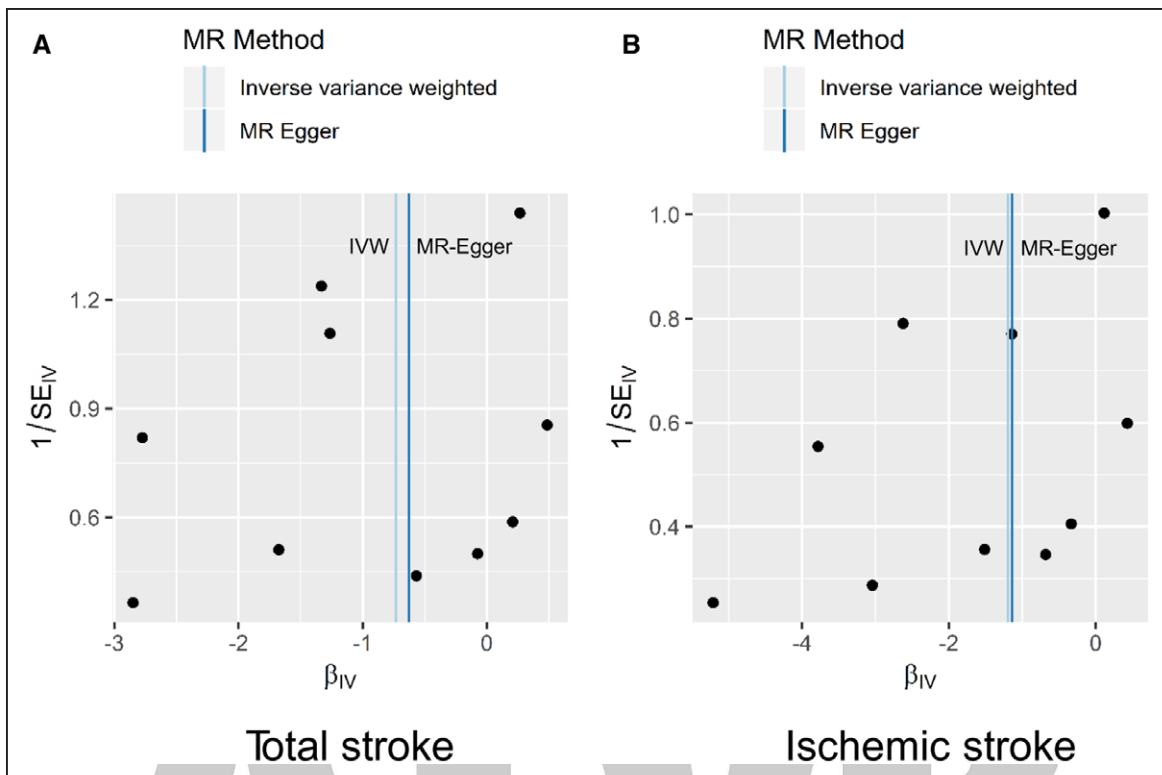
In this present study, we revealed that genetically determined serum total bilirubin levels are causally associated with decreased total stroke risk in a Korean population. This association increased in magnitude when restricting the analysis to ischemic stroke cases.

A number of observational studies in humans have shown evidence of strong associations between serum bilirubin levels and cardiovascular diseases. For example, serum total bilirubin concentrations were negatively correlated to arterial stiffness in Chinese men.<sup>28</sup> Likewise, serum bilirubin levels were observed to be inversely associated with coronary artery calcification and cardiovascular events in a German population.<sup>29</sup> In Chinese populations, serum total bilirubin levels were negatively associated with silent cerebral infarction, which increases



**Figure 1. Scatter plot to visualize causal effect of serum bilirubin on total stroke risk (A) and ischemic stroke risk (B).**

The slope of the straight line indicates the magnitude of the causal association. IVW indicates inverse-variance weighted; and MR, Mendelian randomization.



**Figure 2. Funnel plots to visualize overall heterogeneity of Mendelian randomization (MR) estimates for the effect of bilirubin on the total stroke (A) or ischemic stroke risk (B).** IVW indicates inverse-variance weighted.

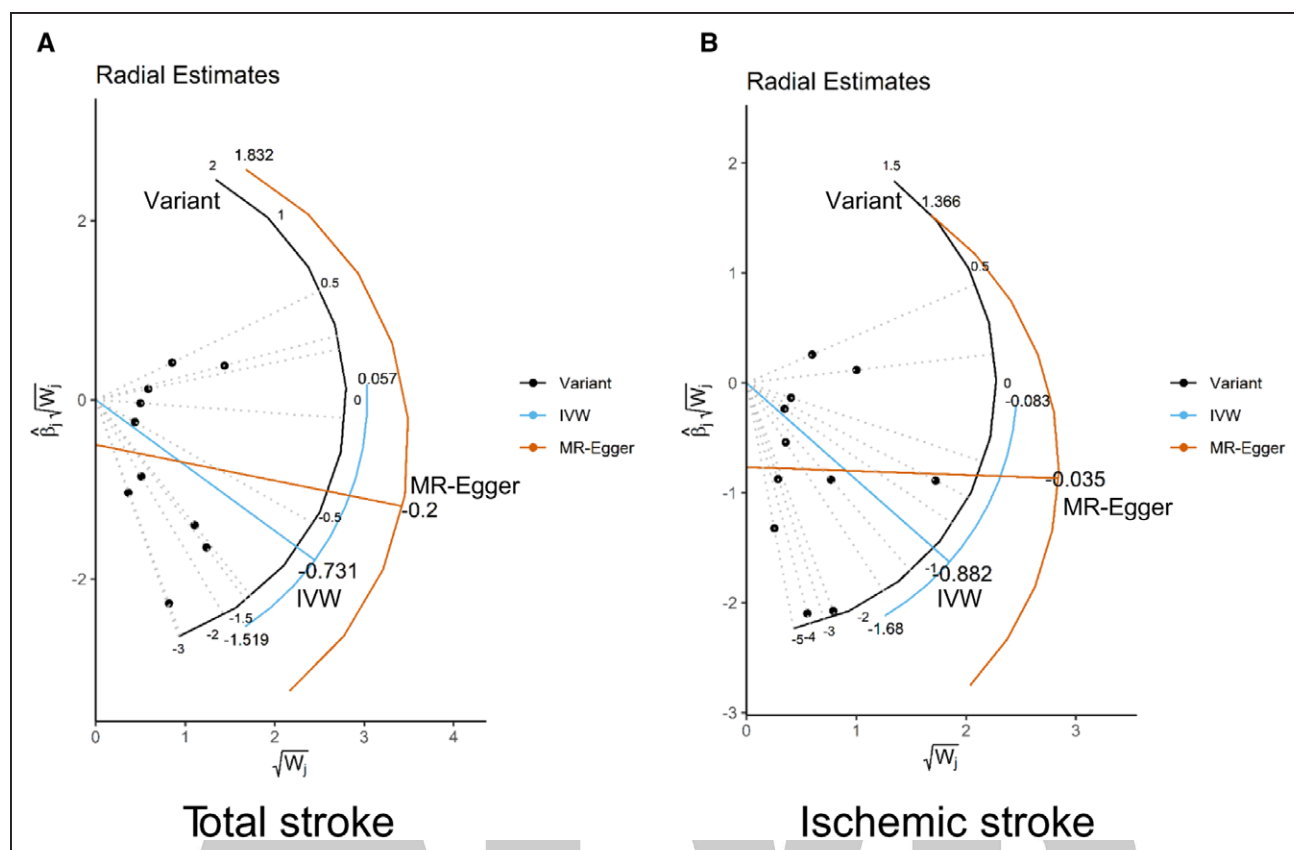
the risk of transient ischemia attack, symptomatic stroke, and cardiovascular diseases.<sup>30</sup> Regarding stroke risk, one National Health and Nutrition Examination Survey 1999 to 2004 study in the United States reported that serum total bilirubin levels were inversely associated with prevalent stroke and adverse stroke outcomes.<sup>9</sup> In the Korean population specifically, serum bilirubin concentrations were estimated to be negatively associated with ischemic stroke in men.<sup>10</sup>

Using 2-sample summary MR, we investigated the causal relationship between circulating serum bilirubin levels and stroke risk. Interestingly, we found that genetically increased bilirubin levels are causally associated with decreased total stroke risk. When we limited the outcome to ischemic stroke, the magnitude of relevance became stronger. These results are consistent with previous observational studies and support the argument that previous MR studies lacked statistical power to detect such associations. As previous MR studies showed evidence of the inverse association, albeit with *P* values exceeding conventional thresholds, we strongly support the argument put forward by Davey Smith and Sterne,<sup>31</sup> that findings should be interpreted with respect to the context of the study, and not be reduced to arbitrary thresholds.

Several experimental studies also support our findings. Vogel et al<sup>32</sup> have revealed that bilirubin prevents atherosclerotic lesion formation in low-density

lipoprotein receptor-deficient mice (*Ldlr*<sup>-/-</sup>) by inhibiting endothelial VCAM-1 (vascular cell adhesion molecule 1) and ICAM-1 (intercellular adhesion molecule 1)-mediated leukocyte migration. During atherosclerotic pathophysiology, binding of leukocyte integrins to endothelial VCAM-1 and ICAM-1 generates cellular reactive oxygen species production. These reactive oxygen species lead to alterations in the endothelial junctional structure that facilitate the transmigration of leukocytes. But bilirubin, a potent antioxidant, can remove the reactive oxygen species induced by the VCAM-1 and ICAM-1 activation that finally inhibit a migration of leukocyte and atherosclerotic lesion formation. This study group also has reported that bilirubin suppresses the transendothelial migration of murine lymphocytes in vitro<sup>33</sup> and also attenuates tissue injury in mouse models of VCAM-1-dependent inflammation.<sup>34</sup> Another study group has also suggested that bilirubin has antiproliferative effects on vascular smooth muscle cells in vitro through p53-dependent cell cycle arrest by hypophosphorylation of the retinoblastoma tumor suppressor protein in growth factor-stimulated vascular smooth muscle cells.<sup>35</sup> Collectively, the putative biological mechanisms linking serum bilirubin to stroke risk are bilirubin-mediated inhibition of lipid oxidation, immune cell migration, and vascular cell proliferation.<sup>36</sup>

The present 2-sample summary MR has several strengths compared with one-sample MR. First, in 2-sample MR design, exposure and outcome variables



**Figure 3. Radial plots to visualize individual outlier single nucleotide polymorphisms (SNPs) in the Mendelian randomization (MR) estimates for total stroke (A) or ischemic stroke risk (B).**

Black dots show valid SNPs and green dots display invalid outlier SNPs. There is no significant outlier SNP in these present plots. IVW indicates inverse-variance weighted.

are extracted from 2 independent populations, respectively. This is advantageous in cases where exposure and outcome data are not available in the same set of individuals.<sup>37</sup> In this study, serum bilirubin levels were measured in KoGES, but stroke cases were not followed in this population. However, using stroke cases in the KCPS-II biobank, a homogenous population in terms of ethnicity and demographic profile, MR analyses could be performed. Further, utilizing multiple samples improves the overall sample size and subsequent precision of causal effect estimates. As a consequence, we were able to identify a causal association between bilirubin levels and decreased stroke risk where our previous one-sample MR study was unable.<sup>12</sup> The 2-sample summary MR approach is also useful in not restricting researchers to using individual level data. In this study, summary estimates were extracted from available individual level data; however, the methods can be applied in principle using publicly available GWAS data, as demonstrated by the MR-Base online platform.<sup>38</sup>

Two sample summary MR also has several characteristics which improve upon analyses using a single sample. In cases where the instrument is not strong, violating the first MR assumption (IV1), effect estimates will be attenuated toward the null. This provides a more

conservative estimate, as opposed to the one-sample setting in which weak instruments result in bias toward the observational estimate. A related advantage with respect to the underlying MR assumptions is that the 2-sample study design is ideal for sensitivity analyses such as MR-Egger regression. As such, in performing a 2-sample summary MR, it is possible to perform a greater range of sensitivity analyses and arrive at more robust estimates of causal effect.

There are several limitations to this study. First, due to using summary data, any potential nonlinear relationships or stratification effects cannot be explored. Several previous epidemiological studies have argued that the association between bilirubin and cardiovascular risk resembles a U-shape<sup>39,40</sup> and differs by sex.<sup>10,28,40</sup> However, our previous work using quintile grouping by bilirubin level showed inversely linear associations between bilirubin and ischemic stroke risk in men, whereas all stroke cases showed no evidence of association. This result suggests that the observed U-shape association could potentially be driven by epidemiological studies failing to account for subtypes of CVD.<sup>10</sup> Second, the absence of horizontal pleiotropic pathways is difficult to assess and if violated would result in biased estimates using IVW regression. We used MR-Egger and weighted



median approaches to test for pleiotropic effects; however, each approach relies on a set of assumptions which are, at present untestable. Pleiotropic could be present where a majority of genetic instruments exhibit similar bias or where the instrument strength is associated with the magnitude of pleiotropic effect. The absence of outliers supports these assumptions holding, though the possibility cannot be discounted.

As a third limitation, the clumping algorithm used to identify correlated genetic variants relied primarily on European samples from the 1000 genomes project. This is owing to the lack of corresponding data for East Asian populations. However, as none of the genetic instruments shared a common gene, it seems probable that the conventional clumping algorithm is sufficient in this case. This is further supported by the inclusion of SNPs located on the *UGT1A1* and *SLCO1B1* genes, which encode essential proteins for the bilirubin metabolism. This shows that the genetic variants selected remain biologically plausible. Finally, it is important to emphasize that as this study focuses solely upon effects within a Korean population, the results cannot be immediately generalized to other ethnicities without further justification.

## CONCLUSIONS

In summary, we have shown evidence of causal relationships between circulating serum bilirubin levels and decreased stroke risk in Korean population. These results support the conclusions of previous observational studies between bilirubin and oxidative stress-related diseases such as stroke and atherosclerosis. This study therefore contributes to a growing body of evidence suggesting bilirubin could play a role as a therapeutic target to cardiovascular diseases.

## ARTICLE INFORMATION

Received June 10, 2019; accepted November 4, 2019.

### Affiliations

From the Department of Public Health, Graduate School, Yonsei University, Seoul, Korea (Y.C., S.H.J.); Department of Epidemiology, Institute for Health Promotion, Graduate School of Public Health, Yonsei University, Seoul, Korea (Y.C., S.J.L., K.J.J., J.-Y.L., H.K., S.H.J.); Population Health Science Institute, University of Bristol, Barley House, Oakfield Grove, United Kingdom (W.S.); and Health Insurance Policy Research Institute, National Health Insurance Service, Wonju, Korea (S.J.L., J.H.B., S.L.).

### Acknowledgments

We thank the staff and participants from the Health Insurance Policy Research Institute of National Health Insurance Service for their valuable contributions. This study was conducted with biobiosources from National Biobank of Korea, the Center for Disease Control and Prevention, Republic of Korea (KBN-2016-013).

### Sources of Funding

This study was funded by a grant from the Korean Health Technology R&D Project, Ministry of Health & Welfare, Republic of Korea (H114C2686).

### Disclosures

None.

## REFERENCES

- Sedlak TW, Snyder SH. Bilirubin benefits: cellular protection by a biliverdin reductase antioxidant cycle. *Pediatrics*. 2004;113:1776–1782. doi: 10.1542/peds.113.6.1776
- Stocker R. Antioxidant activities of bile pigments. *Antioxid Redox Signal*. 2004;6:841–849. doi: 10.1089/ars.2004.6.841
- Stocker R, Yamamoto Y, McDonagh AF, Glazer AN, Ames BN. Bilirubin is an antioxidant of possible physiological importance. *Science*. 1987;235:1043–1046. doi: 10.1126/science.3029864
- Gazzin S, Vitek L, Watchko J, Shapiro SM, Tiribelli C. A novel perspective on the biology of bilirubin in health and disease. *Trends Mol Med*. 2016;22:758–768. doi: 10.1016/j.molmed.2016.07.004
- Sierra C, Coca A, Schiffrin EL. Vascular mechanisms in the pathogenesis of stroke. *Curr Hypertens Rep*. 2011;13:200–207. doi: 10.1007/s11906-011-0195-x
- Allen CL, Bayraktutan U. Oxidative stress and its role in the pathogenesis of ischaemic stroke. *Int J Stroke*. 2009;4:461–470. doi: 10.1111/j.1747-4949.2009.00387.x
- Khoshnam SE, Winlow W, Farzaneh M, Farbood Y, Moghaddam HF. Pathogenic mechanisms following ischemic stroke. *Neurol Sci*. 2017;38:1167–1186. doi: 10.1007/s10072-017-2938-1
- Kunutsor SK, Bakker SJ, Gansevoort RT, Chowdhury R, Dullaart RP. Circulating total bilirubin and risk of incident cardiovascular disease in the general population. *Arterioscler Thromb Vasc Biol*. 2015;35:716–724. doi: 10.1161/ATVBAHA.114.304929
- Perlstein TS, Pande RL, Creager MA, Weuve J, Beckman JA. Serum total bilirubin level, prevalent stroke, and stroke outcomes: nhanes 1999–2004. *Am J Med*. 2008;121:781–788 e781. doi: 10.1016/j.amjmed.2008.03.045
- Kimm H, Yun JE, Jo J, Jee SH. Low serum bilirubin level as an independent predictor of stroke incidence: a prospective study in Korean men and women. *Stroke*. 2009;40:3422–3427. doi: 10.1161/STROKEAHA.109.560649
- Oda E, Kawai R. A possible cross-sectional association of serum total bilirubin with coronary heart disease and stroke in a Japanese health screening population. *Heart Vessels*. 2012;27:29–36. doi: 10.1007/s00380-011-0123-7
- Lee SJ, Jee YH, Jung KJ, Hong S, Shin ES, Jee SH. Bilirubin and stroke risk using a mendelian randomization design. *Stroke*. 2017;48:1154–1160. doi: 10.1161/STROKEAHA.116.015083
- Kim Y, Han BG; KoGES group. Cohort profile: the Korean Genome and Epidemiology Study (KoGES) consortium. *Int J Epidemiol*. 2017;46:1350. doi: 10.1093/ije/dyx105
- Moon S, Kim YJ, Han S, Hwang MY, Shin DM, Park MY, Lu Y, Yoon K, Jang HM, Kim YK, et al. The Korea biobank array: design and identification of coding variants associated with blood biochemical traits. *Sci Rep*. 2019;9:1382. doi: 10.1038/s41598-018-37832-9
- Delaneau O, Marchini J, Zagury JF. A linear complexity phasing method for thousands of genomes. *Nat Methods*. 2011;9:179–181. doi: 10.1038/nmeth.1785
- Fuchsberger C, Abecasis GR, Hinds DA. minimac2: faster genotype imputation. *Bioinformatics*. 2015;31:782–784. doi: 10.1093/bioinformatics/btu704
- Staley JR, Blackshaw J, Kamat MA, Ellis S, Surendran P, Sun BB, Paul DS, Freitag D, Burgess S, Danesh J, et al. PhenoScanner: a database of human genotype-phenotype associations. *Bioinformatics*. 2016;32:3207–3209. doi: 10.1093/bioinformatics/btw373
- Jee YH, Emberson J, Jung KJ, Lee SJ, Lee S, Back JH, Hong S, Kimm H, Sherliker P, Jee SH, et al. Cohort profile: The Korean Cancer Prevention Study-II (KCPS-II) biobank. *Int J Epidemiol*. 2018;47:385–386f. doi: 10.1093/ije/dyx226
- Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol*. 2015;44:512–525. doi: 10.1093/ije/dyv080
- Au Yeung SL, Borges MC, Lawlor DA. Association of genetic instrumental variables for lung function on coronary artery disease risk: a 2-Sample Mendelian Randomization Study. *Circ Genom Precis Med*. 2018;11:e001952. doi: 10.1161/CIRCGEN.117.001952
- Palmer TM, Lawlor DA, Harbord RM, Sheehan NA, Tobias JH, Timpson NJ, Davey Smith G, Sterne JA. Using multiple genetic variants as instrumental variables for modifiable risk factors. *Stat Methods Med Res*. 2012;21:223–242. doi: 10.1177/0962280210394459
- van Vliet NA, Noordam R, van Klippen JB, Westendorp RG, Bassett JD, Williams GR, van Heemst D. Thyroid stimulating hormone and bone mineral density: evidence from a two-sample mendelian randomization study and a Candidate Gene Association Study. *J Bone Miner Res*. 2018;33:1318–1325. doi: 10.1002/jbmr.3426

23. Gage SH, Jones HJ, Burgess S, Bowden J, Davey Smith G, Zammit S, Munafò MR. Assessing causality in associations between cannabis use and schizophrenia risk: a two-sample Mendelian randomization study. *Psychol Med*. 2017;47:971–980. doi: 10.1017/S0033291716003172
24. Sanderson E, Davey Smith G, Windmeijer F, Bowden J. An examination of multivariable mendelian randomization in the single-sample and two-sample summary data settings. *Int J Epidemiol*. 2018;48:713–727. doi: 10.1093/ije/dyy262
25. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, Laurin C, Burgess S, Bowden J, Langdon R, et al. The mr-base platform supports systematic causal inference across the human genome. *Elife*. 2018;7:e34408. doi: 10.7554/eLife.34408
26. Liu JZ, van Sommeren S, Huang H, Ng SC, Alberts R, Takahashi A, Ripke S, Lee JC, Jostins L, Shah T, et al; International Multiple Sclerosis Genetics Consortium; International IBD Genetics Consortium. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat Genet*. 2015;47:979–986. doi: 10.1038/ng.3359
27. Jostins L, Ripke S, Weersma RK, Duerr RH, McGovern DP, Hui KY, Lee JC, Schumm LP, Sharma Y, Anderson CA, et al; International IBD Genetics Consortium (IIBDGC). Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature*. 2012;491:119–124. doi: 10.1038/nature11582
28. Li Y, Meng SY, Meng CC, Yu WG, Wang RT. Decreased serum bilirubin is associated with arterial stiffness in men. *Nutr Metab Cardiovasc Dis*. 2013;23:375–381. doi: 10.1016/j.numecd.2011.09.004
29. Mahabadi AA, Lehmann N, Möhlenkamp S, Kälsch H, Bauer M, Schulz R, Moebus S, Jöckel KH, Erbel R, Heusch G. Association of bilirubin with coronary artery calcification and cardiovascular events in the general population without known liver disease: the Heinz Nixdorf Recall study. *Clin Res Cardiol*. 2014;103:647–653. doi: 10.1007/s00392-014-0697-z
30. Li RY, Cao ZG, Zhang JR, Li Y, Wang RT. Decreased serum bilirubin is associated with silent cerebral infarction. *Arterioscler Thromb Vasc Biol*. 2014;34:946–951. doi: 10.1161/ATVBAHA.113.303003
31. Sterne JA, Davey Smith G. Sifting the evidence—what's wrong with significance tests? *BMJ*. 2001;322:226–231. doi: 10.1136/bmj.322.7280.226
32. Vogel ME, Idelman G, Konanah ES, Zucker SD. Bilirubin prevents atherosclerotic lesion formation in low-density lipoprotein receptor-deficient mice by inhibiting endothelial vcam-1 and icam-1 signaling. *J Am Heart Assoc*. 2017;6
33. Keshavan P, Deem TL, Schwemberger SJ, Babcock GF, Cook-Mills JM, Zucker SD. Unconjugated bilirubin inhibits VCAM-1-mediated transendothelial leukocyte migration. *J Immunol*. 2005;174:3709–3718. doi: 10.4049/jimmunol.174.6.3709
34. Zucker SD, Vogel ME, Kindel TL, Smith DL, Idelman G, Avissar U, Kakarlapudi G, Masnovi ME. Bilirubin prevents acute DSS-induced colitis by inhibiting leukocyte infiltration and suppressing upregulation of inducible nitric oxide synthase. *Am J Physiol Gastrointest Liver Physiol*. 2015;309:G841–G854. doi: 10.1152/ajpgi.00149.2014
35. Ollinger R, Yamashita K, Bilban M, Erat A, Kogler P, Thomas M, Cszimadia E, Ushveva A, Margreiter R, Bach FH. Bilirubin and biliverdin treatment of atherosclerotic diseases. *Cell Cycle*. 2007;6:39–43. doi: 10.4161/cc.6.1.3700
36. Targher G. Risk of ischemic stroke and decreased serum bilirubin levels: is there a causal link? *Arterioscler Thromb Vasc Biol*. 2014;34:702–704. doi: 10.1161/ATVBAHA.114.303279
37. Zheng J, Baird D, Borges MC, Bowden J, Hemani G, Haycock P, Evans DM, Smith GD. Recent developments in mendelian randomization studies. *Curr Epidemiol Rep*. 2017;4:330–345. doi: 10.1007/s40471-017-0128-6
38. Lawlor DA. Commentary: two-sample Mendelian randomization: opportunities and challenges. *Int J Epidemiol*. 2016;45:908–915. doi: 10.1093/ije/dyw127
39. Troughton JA, Woodside JV, Young IS, Arveiler D, Amouyel P, Ferrières J, Ducimetière P, Patterson CC, Kee F, Yarnell JW, et al; PRIME Study Group. Bilirubin and coronary heart disease risk in the Prospective Epidemiological Study of Myocardial Infarction (PRIME). *Eur J Cardiovasc Prev Rehabil*. 2007;14:79–84. doi: 10.1097/01.hjr.0000230097.81202.9f
40. Breimer LH, Wannamethee G, Ebrahim S, Shaper AG. Serum bilirubin and risk of ischemic heart disease in middle-aged British men. *Clin Chem*. 1995;41:1504–1508.

# ATVB

## Arteriosclerosis, Thrombosis, and Vascular Biology

### FIRST PROOF ONLY