This is the published version of a paper published in Journal of the American College of Cardiology.

Citation for the original published paper (version of record):

van der Laan, S., Fall, T., Soumaré, A., Teumer, A., Sedaghat, S. et al. (2016)
Cystatin C and cardiovascular disease: A mendelian randomization study.
Journal of the American College of Cardiology, 68(9): 934-945
https://doi.org/10.1016/j.jacc.2016.05.092

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:
http://urn.kb.se/resolve?urn=urn:nbn:se:du-23629
Cystatin C and Cardiovascular Disease

A Mendelian Randomization Study

Sander W. van der Laan, MSC, a,* Tove Fall, PhD, b,* Aicha Soumaré, PhD, c Alexander Teumer, PhD, d,e Sanaz Sedaghat, MSC, f Jens Baumert, PhD, g Delilah Zabaneh, PhD, h,i Jessica van Setten, PhD, a Ivana Isgum, PhD, j Tessel E. Galesloot, PhD, k Marcus Dör, MD, l Patrik K. Magnusson, PhD, l,1 Vilmantas Gedraitis, PhD, l,1 Anders Larsson, PhD, l,1 Andrew P. Morris, PhD, l,1,x Anja C. Morrison, PhD, y Nora Franceschini, MD, MPH, z Joshua C. Bis, PhD, z,x Maryam Kavousi, MD, PhD, z,1

ABSTRACT

BACKGROUND Epidemiological studies show that high circulating cystatin C is associated with risk of cardiovascular disease (CVD), independent of creatinine-based renal function measurements. It is unclear whether this relationship is causal, arises from residual confounding, and/or is a consequence of reverse causation.

OBJECTIVES The aim of this study was to use Mendelian randomization to investigate whether cystatin C is causally related to CVD in the general population.

METHODS We incorporated participant data from 16 prospective cohorts (n = 76,481) with 37,126 measures of cystatin C and added genetic data from 43 studies (n = 252,216) with 63,292 CVD events. We used the common variant rs911119 in CST3 as an instrumental variable to investigate the causal role of cystatin C in CVD, including coronary heart disease, ischemic stroke, and heart failure.

RESULTS Cystatin C concentrations were associated with CVD risk after adjusting for age, sex, and traditional risk factors (relative risk: 1.82 per doubling of cystatin C; 95% confidence interval [CI]: 1.56 to 2.13; p = 2.12 × 10⁻¹⁴). The minor allele of rs911119 was associated with decreased serum cystatin C (6.13% per allele; 95% CI: 5.75 to 6.50; p = 5.95 × 10⁻²¹), explaining 2.8% of the observed variation in cystatin C. Mendelian randomization analysis did not provide evidence for a causal role of cystatin C, with a causal relative risk for CVD of 1.00 per doubling cystatin C (95% CI: 0.82 to 1.22; p = 0.994), which was statistically different from the observational estimate (p = 1.6 × 10⁻³). A causal effect of cystatin C was not detected for any individual component of CVD.

CONCLUSIONS Mendelian randomization analyses did not support a causal role of cystatin C in the etiology of CVD. As such, therapeutics targeted at lowering circulating cystatin C are unlikely to be effective in preventing CVD. (J Am Coll Cardiol 2016;68:934–45) © 2016 The Authors. Published by Elsevier Inc. on behalf of the American College of Cardiology Foundation. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
Cystatin C (encoded by CST3 on 20p11.21) is a potent cysteine protease inhibitor that plays pleiotropic roles in human vascular pathophysiology, in particular regulating cathepsins S and K (1–3), and serves as a marker of renal function (4). Cathepsins are overexpressed in human atherosclerotic and aneurysmal lesions, giving rise to rupture-prone plaques by degrading the extracellular matrix (Figure 1) (1). Prospective epidemiological studies show a strong association between circulating cystatin C and risk of future coronary heart disease (CHD), ischemic stroke (IS), and heart failure (HF) (5,6). This association is also present in patients with subclinical atherosclerosis (7) or those at high risk of cardiovascular disease (CVD) (8–10), and is independent of renal function determined by formulae on the basis of creatinine measurements or other cardiovascular risk factors (5,11–14). Moreover, heritability analyses indicate that CVD and cystatin C concentrations have shared polygenic backgrounds (15).

The accumulating experimental and epidemiological evidence supports the hypothesis that cystatin C could play a causal role in CVD etiology independent of renal function.

From the *Laboratory of Experimental Cardiology, Division of Heart and Lungs, University Medical Center Utrecht, Utrecht, the Netherlands; **Department of Medical Sciences, Cardiovascular Epidemiology, Uppsala University, Uppsala, Sweden; ***INSERM U219 Team Vintage, University of Bordeaux, Bordeaux, France; ****Department SHIP-KEF, Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany; *****Deutsches Zentrum für Herz- und Kreislauforschung (DZHK), German Centre for Cardiovascular Research) partner site, Greifswald, Germany; ****Department of Epidemiology, Erasmus University Medical Center, Rotterdam, the Netherlands; ***Institute of Epidemiology II, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany; ***Department of Genetics, Environment and Evolution, University College London, London, United Kingdom; 1Department of Medical Sciences, Cardiovascular Epidemiology, Uppsala University, Uppsala, Sweden; 1Department of Medical Sciences, Uppsala University, Uppsala, Sweden; 2Department of Medical Sciences, Uppsala University, Uppsala, Sweden; 3Department of Biostatistics, University of Liverpool, Liverpool, United Kingdom; 4Welcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom; 5Department of Epidemiology, Human Genetics, and Environmental Sciences, University of Texas Health Science Center, Houston, Texas; 6Department of Epidemiology, University of North Carolina, Chapel Hill, North Carolina; 7Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, Washington; 8Department of Cardiology, Boston Veterans Administration Healthcare, West Roxbury, Massachusetts; 9National Heart, Lung, and Blood Institute Framingham Heart Study, Framingham, Massachusetts; 10Centre for Cardiovascular Genetics, Institute of Cardiovascular Sciences; 11University College London, London, United Kingdom; 12MRC Integrative Epidemiology Unit, School of Social and Community Medicine, University of Bristol, Bristol, United Kingdom; 13Department of Cardiology, Division Heart and Lungs, University Medical Center Utrecht, Utrecht, the Netherlands; 14National Institute for Health Research Cardiovascular Biomedical Research Unit, William Harvey Research Institute, Queen Mary University of London, London, United Kingdom; 15Institute of Stroke and Dementia Research, Klinikum der Universität München, Ludwig-Maximilians-University Munich, Munich, Germany; 16Munich Cluster for Systems Neurology (SyNergy), Munich, Germany; 17Departments of Neurology and Health Evaluation Sciences, University of Virginia, Charlottesville, Virginia; 18Institute for Integrative and Experimental Genomics, University of Lübeck, Lübeck, Germany; 19Department of Cardiovascular Sciences, University of Leicester, British Heart Foundation Cardiovascular Research Centre, Glenfield Hospital, Leicester, United Kingdom; 20National Institute for Health Research Leicester Cardiovascular Biomedical Research Unit, Glenfield Hospital, Leicester, United Kingdom; 21Deutsches Herzzentrum München, Technische Universität München, Munich, Germany; 22DZHK, German Centre for Cardiovascular Research, partner site Munich Heart Alliance, Munich, Germany; 23Department of Statistics, University of Oxford, Oxford, United Kingdom; 24The Genetic Epidemiology Research Group, Institute of Cardiovascular Science, University College London, London, United Kingdom; 25Bar’s Heart Centre, London, United Kingdom; 26Fann Institute of Health Informatics, University College London, London, United Kingdom; 27Department of Internal Medicine, Radboud University Medical Center, Nijmegen, the Netherlands; 28Institute for Epidemiology and Preventive Medicine, University of Regensburg, Regensburg, Germany; 29Department of Internal Medicine, Erasmus University Medical Center, Rotterdam, the Netherlands; 30Department of Internal Medicine II-Cardiology, University of Ulm Medical Center, Ulm, Germany; 31Durrer Center for Cardiogenetic Research, ICIN-Netherlands Heart Institute, Utrecht, the Netherlands; 32Metabolic and Renal Research Group, UIT The Arctic University of Norway, Tromse, Norway; 33Section of Nephrology, University Hospital of North Norway, Tromse, Norway; 34Department of Community Medicine, UIT The Arctic University of Norway, Tromse, Norway; 35Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, the Netherlands; 36Biological and Social Epidemiology, Institute for Social and Economic Research, University of Essex, Essex, United Kingdom; 37Department of Cardiology, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands; 38Department of Genetics, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands; 39Department of Epidemiology and Public Health, University College London, London, United Kingdom;
Mendelian randomization harnesses the properties of the genome to enable causal inference of a biomarker (16). Specifically, the invariant nature of the genome and the random distribution of alleles from parents to offspring at conception mean that genetic information is not influenced by disease status (reverse causality) and should be free from confounding by traditional risk factors. Thus, genetic variation that modulates serum concentrations of cystatin C could serve as an instrumental variable to assess the effect of lifelong elevated concentrations of cystatin C on disease risk, independent of potential confounders (16).

To this end, we established the Cystatin C Mendelian Randomization Consortium to investigate the causal relevance of serum cystatin C to CVD risk. From the published genome-wide association studies (GWAS), we identified common single nucleotide polymorphisms (SNPs) in the CST3 locus associated with circulating concentrations of cystatin C (17-20) and selected rs911119 as showing the strongest association, independent from other variants (18).

References:

1. van der Laan et al., JACC Vol. 68, No. 9, 2016.

Manuscript received December 18, 2015; revised manuscript received May 12, 2016, accepted May 18, 2016.
We robustly associated rs911119 with circulating cystatin C in 9 cohorts (8 of which have not participated in prior GWAS). Next, we evaluated the association of serum cystatin C with CVD in observational analyses of prospective cohorts. Finally, we used rs911119 as an instrument variable to test the causal effect of circulating cystatin C on CVD through Mendelian randomization.

**METHODS**

We included data from 15 general population–based cohorts and 1 randomized clinical trial (Table 1, Online Tables 1 and 2) (detailed study descriptions in Online Appendix). All participants provided informed consent, and the local ethics committees approved these studies.

**CONSORTIA DATA.** We included individual study summary statistics from the discovery stages of CARDIoGRAM (Coronary Artery Disease Genome-wide Replication and Meta-analysis), including 17 studies, 20,251 CHD cases, and 60,183 control subjects (21) and the METASTROKE collaboration (the first large meta-analysis of stroke GWAS data), consisting of 15 studies, 12,389 all-cause IS cases, and 62,004 control subjects (22). We also included the summary statistics from the C4D (Coronary Artery Disease Genetic Consortium) on CHD (23) (including 4 studies comprising 15,388 cases and 15,040 control subjects) and CHARGE-HF (Cohorts for Heart and Aging Research in Genomic Epidemiology–Heart Failure), the CHARGE GWAS on incident HF, which included 4 studies, 2,526 cases, and 18,400 control subjects from European descent (24). Additionally, we included consortia data on a number of cardiovascular traits (Online Table 3). For the primary outcome (CVD), we meta-analyzed genetic association results from the 16 individual cohorts, CARDIoGRAM, C4D, METAStROKE, and...
Cystatin C and Cardiovascular Disease

AUGUST 30, 2016

van der Laan et al.

TABLE 1 Characteristics of Prospective Cohorts

<table>
<thead>
<tr>
<th>Study</th>
<th>Total SNP</th>
<th>Cystatin C</th>
<th>CHD$</th>
<th>IS$</th>
<th>HF$</th>
<th>MI$</th>
<th>Male</th>
<th>Age (yrs)</th>
<th>Cystatin C (mg/dl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>JC</td>
<td>6,440</td>
<td>6,435</td>
<td>1,244</td>
<td>1,717</td>
<td>1,235</td>
<td>459</td>
<td>439</td>
<td>486</td>
<td>39.19 ± 5.52</td>
</tr>
<tr>
<td>EPIC-NL</td>
<td>6,265</td>
<td>5,192</td>
<td></td>
<td>1,967</td>
<td>1,430</td>
<td>537</td>
<td>1,430</td>
<td>22.39</td>
<td>53.80 ± 10.23</td>
</tr>
<tr>
<td>GOSH</td>
<td>1,478</td>
<td>1,479</td>
<td></td>
<td>493</td>
<td>111</td>
<td>235</td>
<td>233</td>
<td>42.08</td>
<td>51.08 ± 11.86</td>
</tr>
<tr>
<td>HRS</td>
<td>7,844</td>
<td>5,585</td>
<td>5,777</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.92 ± 0.24</td>
</tr>
<tr>
<td>KORA</td>
<td>4,856</td>
<td>1,867</td>
<td>4,676</td>
<td>540</td>
<td>341</td>
<td>255</td>
<td>341</td>
<td>49.53</td>
<td>49.75 ± 14.11</td>
</tr>
<tr>
<td>NBS</td>
<td>1,819</td>
<td>1,297</td>
<td></td>
<td>66</td>
<td>66</td>
<td>170</td>
<td>170</td>
<td>49.48</td>
<td>61.05 ± 10.26</td>
</tr>
<tr>
<td>PIVUS</td>
<td>1,016</td>
<td>949</td>
<td>1,004</td>
<td>255</td>
<td>175</td>
<td>71</td>
<td>75</td>
<td>49.90</td>
<td>70.20 ± 0.17</td>
</tr>
<tr>
<td>PREVEND</td>
<td>3,245</td>
<td>3,245</td>
<td>3,245</td>
<td>236</td>
<td>190</td>
<td>58</td>
<td>50.26</td>
<td>49.42</td>
<td>54.46 ± 12.25</td>
</tr>
<tr>
<td>PROSPER</td>
<td>5,244</td>
<td>5,150</td>
<td>2,561</td>
<td>2,034</td>
<td>779</td>
<td>211</td>
<td>762</td>
<td>48.13</td>
<td>75.34 ± 3.35</td>
</tr>
<tr>
<td>Rotterdam</td>
<td>7,983</td>
<td>5,974</td>
<td>3,906</td>
<td>3,579</td>
<td>1,934</td>
<td>1,328</td>
<td>1,625</td>
<td>38.90</td>
<td>73.06 ± 7.49</td>
</tr>
<tr>
<td>SHIP</td>
<td>3,224</td>
<td>3,224</td>
<td>3,212</td>
<td>114</td>
<td>19</td>
<td>87</td>
<td>134</td>
<td>48.08</td>
<td>54.46 ± 15.26</td>
</tr>
<tr>
<td>Tromsø</td>
<td>6,129</td>
<td>6,129</td>
<td>1,251</td>
<td>1,251</td>
<td>494</td>
<td>494</td>
<td>881</td>
<td>47.59</td>
<td>60.59 ± 10.25</td>
</tr>
<tr>
<td>TWINGENE</td>
<td>6,902</td>
<td>6,902</td>
<td>6,740</td>
<td>932</td>
<td>610</td>
<td>287</td>
<td>206</td>
<td>47.23</td>
<td>64.83 ± 8.26</td>
</tr>
<tr>
<td>ULSAM</td>
<td>1,221</td>
<td>1,107</td>
<td>1,193</td>
<td>503</td>
<td>285</td>
<td>175</td>
<td>220</td>
<td>47.59</td>
<td>60.59 ± 10.25</td>
</tr>
<tr>
<td>WHI</td>
<td>7,854</td>
<td>7,844</td>
<td>6,920</td>
<td>932</td>
<td>610</td>
<td>287</td>
<td>206</td>
<td>47.59</td>
<td>60.59 ± 10.25</td>
</tr>
<tr>
<td>Whitehall II</td>
<td>4,961</td>
<td>5,011</td>
<td>1,011</td>
<td>349</td>
<td>254</td>
<td>111</td>
<td>254</td>
<td>74.58</td>
<td>49.19 ± 5.99</td>
</tr>
<tr>
<td>Overall</td>
<td>76,481</td>
<td>61,261</td>
<td>37,126</td>
<td>19,394</td>
<td>11,552</td>
<td>7,057</td>
<td>3,009</td>
<td>8,673</td>
<td></td>
</tr>
</tbody>
</table>

Values are n, %, or mean ± SD. *Total number of individuals with genotype data. Genotypic data were available in 29,805 of the 37,126 individuals that had values for cystatin C, which we used to associate rs911119 with circulating cystatin C. For the genetic analysis of CVD, CHD, IS, and HF, cohorts that contributed toward consortia were excluded.

CHARGE-HF. For all analyses, we excluded overlapping cohorts where appropriate (Online Table 3).

SNP SELECTION AND GENOTYPING. We searched PubMed and identified 5 publications reporting GWAS conducted for cystatin C or its clinical derivative (i.e., estimated glomerular filtration rate [eGFR] on the basis of cystatin C) (17-20). From these publications, 3 SNPs were identified (rs1158167 [20], rs13038305 [19], and rs911119 [18]), with rs911119 showing the strongest independent association with cystatin C. We therefore used rs911119 as our primary SNP of choice. When this SNP was not available, we used suitable proxies in linkage disequilibrium with rs911119 ($r^2$ >0.90) (Online Table 4, Online Figure 1).

The genotyping platforms used by the cohorts are outlined in Online Table 2. All SNPs were in Hardy-Weinberg Equilibrium (p > 0.067) (Online Table 5) with a call rate >95% or imputation quality >0.95, and comparable allele frequencies (Online Figure 2). Online Tables 6 and 7 describe the SNP characteristics from the individual study data of the CARDIoGRAM consortium and METAStROKE collaboration used in our study (21,22). The genotyping, imputation and quality control procedures of these and other consortia are described in Online Table 3.

Cystatin C (mg/l) was measured in 10 of the 16 prospective cohorts in a total of 37,126 individuals, of whom 29,805 had genotype data available. The assays used to quantify serum cystatin C in each study together with the assay QC parameters are outlined in Online Table 8. As cystatin C concentrations were not normally distributed, we log$^2$ transformed these prior to analysis, enabling us to express associations as “per doubling of cystatin C” in observational and Mendelian randomization analyses.

We queried data from the Genotype-Tissue Expression Project (GTEx) through the GTex Portal for rs911119 and its proxies for an effect on Cyst3 expression in whole blood (25). Details of the study design, tissue collection, sample preparation, ribonucleic acid sequencing, genotyping, quality control, and imputation have been described elsewhere (25).

Other expression quantitative trait locus (eQTL) datasets we queried have been described before and pertain to expression in monocytes (26), lymphoblastoid cell lines (27), fibroblasts, adipocytes, and lymphoblastoid cell lines from the MuTHER (Multiple Tissue Human Expression Resource) project (28).

Details on the cardiovascular risk factors and traits that we assessed are given in the Online Appendix.

CLINICAL OUTCOMES. Our primary outcome was CVD, a composite of CHD, IS, and HF. We defined CHD as morbidity or mortality from myocardial infarction (MI) (International Classification of Disease, 10th Revision [ICD-10] codes I21 and I22), acute coronary syndrome, unstable angina, >50% coronary
artery stenosis on angiography, and/or having an intervention by percutaneous coronary angioplasty or coronary artery bypass graft (ICD-10 codes: I20.0, I21, and I22; surgical codes: FNG02, FNG05, FNC, FND, and FNE). IS was defined as morbidity or mortality originating from occlusion and stenosis of cerebral and pre-cerebral arteries; this includes large artery stroke, small vessel disease, and cardioembolic stroke (ICD-10: I63). HF was defined as left ventricular failure, (combined) diastolic or systolic HF, and unspecified HF, excluding cardiac arrest (ICD-10 code I50).

We further defined secondary outcomes as CHD, IS, HF, and MI. Clinical outcome data were obtained from the patient and from cause of death registries or validated events. An overview of outcome definitions for each study is provided in Online Table 9.

**STATISTICAL ANALYSIS.** To standardize the analysis procedure, a pre-specified script was used in every study with access to participant data. We conducted observational analysis, genetic analysis, and Mendelian randomization analysis. Detailed information is included in the Online Appendix.

Meta-analyses estimates were pooled using a fixed-effects model with between-study heterogeneity quantified using the I² statistic (29). Random effects modeling was used as a sensitivity analysis. The total sample size used in each analysis depended on the covariates available and the type of case (incident-only or incident plus prevalent) (Online Table 10). Effect estimates from logistic and Cox-regression analyses are referred to as relative risks (RRs).

We applied Bonferroni correction for multiple testing in the genetic association analyses, and we thus set a p-value threshold of 0.05/(5 outcomes + 32 cardiovascular traits) = 0.0014. When appropriate, we adjusted for the relatedness among samples. For Mendelian randomization analyses of clinical events, we estimated the post hoc power as described previously (30). We used the genetic sample size and case/control ratios for each outcome trait in this study, together with the proportion of variance of cystatin C explained by the genetic variant (r² = 0.0275). We calculated the existing power to detect an effect using a Bonferroni-adjusted 2-sided type 1 error (α) of 0.05/5 = 0.01 (corrected for testing 5 clinical outcomes) (Online Figure 3).

Analyses were conducted in Stata Statistical Software Release 13, version 13.1 (StataCorp LP, College Station, Texas) and R version 3.2.3 “Wooden Christmas-Tree” (R Foundation for Statistical Computing, Vienna, Austria) with R Studio version 0.99.983 (RStudio, Inc., Boston, Massachusetts).

**RESULTS**

The Cystatin C Mendelian Randomization Consortium comprises 15 general population-based prospective cohorts and 1 randomized clinical trial including up to 76,481 individuals from European descent (Table 1, Online Tables 1 and 2). In total, 19,394 cardiovascular events were recorded comprising 11,552 CHD events, 7,057 IS cases, 3,009 HF events, and 8,673 MIs (Table 1). A total of 37,126 individuals had measures of serum cystatin C (Table 1, Online Table 8). To maximize power (Online Figure 3) for the genetic analyses of risk factors and clinical outcomes, we added data from relevant consortia, while excluding overlapping data from the 16 participating studies (Online Table 3). The baseline characteristics of the consortia were published previously (21-24, 31-43).

**ASSOCIATION AND SPECIFICITY OF THE GENETIC INSTRUMENT FOR CYSTATIN C CONCENTRATIONS.** The genetic instrument (rs911119, or its proxies) (Online Table 4, Online Figure 1) had similar allele frequencies among the cohorts (Online Figure 2) and showed a strong association with circulating cystatin C. In data from 29,805 individuals (who were genotyped of the 37,126 in whom cystatin C was measured), each additional copy of the minor allele was associated with a 6.13% reduction in cystatin C (95% confidence interval [CI]: 5.75 to 6.50; p = 5.95 × 10⁻²²) and explained 2.75% (95% CI: 0.75 to 4.76) of the phenotypic variation (F-statistic = 961) (Online Appendix, Online Figure 4). We queried various eQTL sources and confirmed that rs911119 only associated with expression of CST3 and not with that of other genes in the region ±500 kb surrounding rs911119 (Online Appendix, Online Figure 5, Online Table 12).

We replicated the association of rs911119 (or its proxies) with cystatin C-based eGFR (0.08 SD per allele; 95% CI: 0.07 to 0.08; p = 4.00 × 10⁻¹⁰⁴) (Online Figure 6) (17-20). We further confirmed a lack of association with creatinine-based eGFR (0.21 SD per allele; 95% CI: -0.11 to 0.52; p = 0.21) (Online Figure 6)(17-20).

**OBSERVATIONAL ASSOCIATIONS OF CIRCULATING CYSTATIN C.** In linear regression analyses adjusted for age and sex, higher serum cystatin C concentrations were associated with several cardiovascular risk factors and traits (Online Figure 7). In contrast, rs911119 showed no significant association with these traits after corrections for multiple testing (Online Figure 6). Use of fixed or random effects modeling did not alter summary estimates derived from meta-analysis (Online Figure 8).

An observational meta-analysis adjusted for age and sex showed a strong dose-dependent relation...
between cystatin C concentrations and CVD (Figure 2, Online Figure 9). Per doubling of cystatin C concentrations, the risk of CVD increased (RR: 2.33; 95% CI: 2.08 to 2.62; \( p = 1.28 \times 10^{-47} \); 6,220 cases and 25,777 control subjects), with the relationship being log-linear (Online Figure 9). Although adjustment for additional confounders diminished the association, an independent relation between cystatin C and CVD persisted (RR: 1.82; 95% CI: 1.56 to 2.13; \( p = 2.12 \times 10^{-14} \)) after adjustment for age, sex, high-density lipoprotein cholesterol, body mass index, systolic blood pressure, eGFR, and smoking status (Figure 2, Online Figure 10, Online Table 11). Adjusting for additional potential confounders (high-sensitivity C-reactive protein, total cholesterol, and glucose) did not further diminish the association (Online Table 11), nor did confining the analysis to incident-only cases (Figure 2, Online Figure 10). In the fully adjusted observational analysis, cystatin C was also associated with an increased risk of CHD, IS, and HF, but not with MI (Figure 3, Online Figure 11, Online Table 11).

We meta-analyzed genetic data from 43 studies with 63,292 CVD cases (including 20,251 CHD cases from CARDIoGRAM, 15,388 CHD cases from C4D, 12,389 IS cases from METASTROKE, and 2,526 HF cases from CHARGE) and a total of 188,924 control subjects (Online Table 10), but found no association of rs911119 with CVD (RR per minor allele: 1.00; 95% CI: 0.98 to 1.02; \( p = 0.994 \) (Online Figure 12). Likewise, we found no association of the genetic variant with CHD, IS, HF, or MI (Online Figure 12).

**MENDELIAN RANDOMIZATION ANALYSIS.** In Mendelian randomization analysis, taking into account both the genetic association with cystatin C (Online Figure 4) and CVD (Online Figure 12) to triangulate the underlying causal effect, we detected no evidence for a causal relation between circulating cystatin C and CVD (odds ratio [OR]: 1.00 per doubling of cystatin C; 95% CI: 0.82 to 1.22; \( p = 0.994 \)) (Figure 2). This was statistically different from the observational estimate obtained from the fully-adjusted model using incident-only events (\( p \) for heterogeneity = \( 1.6 \times 10^{-5} \)). Likewise, no causal association of cystatin C was detected for any individual subtype of vascular disease (Figure 3).

**POWER.** With a combined sample size of 63,292 CVD events, 43,068 CHD events, 16,784 IS events, and 3,440 HF cases (Online Figure 12), we estimated to have >80% power to detect an OR >1.10 per doubling cystatin C for CVD, 1.13 for CHD, 1.19 for IS, and 1.45 for HF (Online Figure 3).

**DISCUSSION**

In this first, large-scale Mendelian randomization analysis, we investigated whether the previously reported robust association between circulating cystatin C and risk of CHD and ischemic stroke (5,6) was likely to be causal. In our model, adjusted for traditional risk factors, cystatin C indeed was strongly associated with CVD risk (Figure 2) in a dose-dependent manner (Online Figures 9 and 11). Even when limited to incident-only cases and in

---

**FIGURE 2** Estimates of the Association of Circulating Cystatin C With CVD Risk

<table>
<thead>
<tr>
<th>Model</th>
<th>Studies</th>
<th>Cases/Controls</th>
<th>Cardiovascular disease</th>
<th>RR (95% C.I.)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>observational minimal adj.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>incident only</td>
<td>7</td>
<td>2,940/23,993</td>
<td></td>
<td>2.37 (2.07 – 2.70)</td>
<td>3.0x10^{-12}</td>
</tr>
<tr>
<td>incident + prevalent</td>
<td>8</td>
<td>5,950/24,573</td>
<td></td>
<td>1.82 (1.56 – 2.13)</td>
<td>2.1x10^{-14}</td>
</tr>
<tr>
<td><strong>observational full adj.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>incident only</td>
<td>7</td>
<td>2,803/22,650</td>
<td></td>
<td>2.19 (1.82 – 2.63)</td>
<td>9.1x10^{-12}</td>
</tr>
<tr>
<td>incident + prevalent</td>
<td>8</td>
<td>5,950/24,573</td>
<td></td>
<td>2.37 (2.07 – 2.70)</td>
<td>3.0x10^{-12}</td>
</tr>
<tr>
<td><strong>casual</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>incident only</td>
<td>7</td>
<td>2,803/22,650</td>
<td></td>
<td>2.19 (1.82 – 2.63)</td>
<td>9.1x10^{-12}</td>
</tr>
<tr>
<td>incident + prevalent</td>
<td>43</td>
<td>63,292/188,924</td>
<td></td>
<td>1.00 (0.82 – 1.22)</td>
<td>0.99</td>
</tr>
</tbody>
</table>

The observational models were minimally adjusted for age and sex (minimal), or fully adjusted for age, sex, body mass index, smoking, high-density lipoprotein cholesterol, estimated glomerular filtration rate, and systolic blood pressure (full). The causal estimates were triangulated using effect estimates of the association of the genetic instrument with cystatin C concentrations (reported in Online Figure 4) and cardiovascular disease (CVD) (Online Figure 12). Total sample sizes may differ from those reported in Table 1 due to the availability of covariates. adj. = adjusted; CI = confidence interval; RR = relative risk.
a fully adjusted analysis, cystatin C had an independent association with clinical events. However, in an adequately powered Mendelian randomization approach, we did not identify evidence of a causal relationship between circulating cystatin C and CVD or any individual cardiovascular component.

Our Mendelian randomization analyses confirmed and extended findings from a recent report analyzing data from the population-based Malmö Diet and Cancer study as well as the CARDIOGRAM meta-analysis, suggesting a lack of association between an SNP (rs13038305, linkage disequilibrium $r^2 = 0.99$ with rs911119) (Online Table 3) in CST3 and the risk of CHD (44). However, in that large analysis, a formal instrumental variable estimate was not synthesized, nor was the association of the SNP with IS or HF investigated. Our meta-analysis, on the basis of data from 43 cohort studies including more than 250,000 individuals with more than 63,000 cardiovascular events, is by far the largest and most comprehensive study to date to examine these associations.

For Mendelian randomization to generate a valid causal estimate, several assumptions needed to be fulfilled. One such assumption was sufficient statistical power. We estimated to have $>80\%$ power to detect ORs smaller than the lower limit of the observed association of cystatin C with CVD from multivariate analyses (Online Figure 3).

The observational models were minimally or fully adjusted and causal model estimates were triangulated as described in Figure 2. Total sample sizes may differ from those reported in Table 1 due to the availability of covariates. Abbreviations as in Figure 2.
Epidemiological evidence shows that increased levels of circulating cystatin C are associated with increased risk of disease. Whether this relation is truly causal or is a consequence of confounding or reverse causality is hard to determine. Our study replicated the strong observational associations between circulating concentrations of cystatin C and risk of cardiovascular diseases (CVDs), but also showed that cystatin C was associated with many potential confounders. We used a genetic variant (rs911119) in the gene CST3, which associates with CST3 gene expression and directly encodes cystatin C. The genetic variant showed a very strong association with circulating cystatin C concentrations, but not with potential confounders. In Mendelian randomization analysis, no evidence for a causal association with CVD was identified. Thus, our study provides no evidence in support of a causal role for circulating cystatin C in the etiology of atherosclerotic vascular disease. HDL = high-density lipoprotein; LDL = low-density lipoprotein; SNP = single nucleotide polymorphism.
Another assumption was that the instrument is strongly associated with the biomarker of interest. Indeed, common variation in the CST3 locus almost exclusively associated with cystatin C (and thus eGFR on the basis of cystatin C) in both previous studies (18) and ours (Online Figures 4 and 6). Convincingly, eQTL analyses confirmed that rs911119 was strongly associated with CST3 expression, but not with the nearby gene CST9, arguing against a potential pleotropic effect (Online Appendix, Online Figure 5). Although we found nominally significant associations with diastolic blood pressure, waist circumference, and smoking, these associations did not persist after correction for multiple testing.

**STUDY LIMITATIONS.** In any Mendelian randomization study, the genetic instrument (in this case rs911119) should not experience “weak instrument bias” (43). In our study, this seemed very unlikely, given the strong association with cystatin C (F-statistic of 961). Furthermore, weak instrument bias would bias the causal estimate toward the observational estimate; in contrast, the causal estimates that we reported were statistically different from the observed estimates and consistently null.

Our study relied on the ability of the assay to quantify serum concentrations of cystatin C with sufficient accuracy and precision. Recent studies have shown that genetic variants can change the epitope measured by the assay (44,45). We cannot rule out the possibility that our instrument (rs911119) or its proxies altered the epitope (versus actually changing the quantity of circulating cystatin C), nor can we be certain to what extent such a change would affect the ability to detect an association with cystatin C concentrations. Last, in principle, the assay type and the time period of measurement could have influenced our findings, although in our studies, the mean cystatin C concentrations were comparable (Table 1) and we found consistent associations between our genetic variant and cystatin C (Online Figure 4) and between cystatin C and risk of CVD across studies.

Although we fitted a multivariate model that extensively adjusted for confounders for observational analyses, residual confounding may still exist, which is a classic challenge for conventional observational epidemiology. Specifically, as no gold standard measurements of renal function (such as inulin-based GFR measurements) were quantified in studies contributing to this analysis, it remains possible that residual confounding by impaired kidney function remained and was not fully accounted for by adjustments in our observational analyses. As a biomarker for kidney function, cystatin C has proven its value and represents a stronger predictor for CVD risk than does creatinine (4). Thus, although our analyses provided no evidence for a causal association between cystatin C and CVD, it did not preclude the use of cystatin C in disease prediction.

We should note that considerable heterogeneity (I²) existed in our observational analysis (Online Figure 7). This might have been due to the number of studies included (up to 8) in our observational analysis (as compared with the genetic analysis). Conversely, little heterogeneity existed in our genetic analysis (Online Figure 6). Adding more studies to the observational analysis (46) or stratifying on the basis of these subgroups (29) might reduce heterogeneity and/or identify potential characteristics that account for heterogeneity. Also, a more uniform definition of clinical outcomes across studies contributing toward the observational analysis of cystatin C and event risk might reduce the heterogeneity further.

**CONCLUSIONS**

We conducted a comprehensive Mendelian randomization of circulating cystatin C in the development of CVD in the general population. Our findings suggest that residual confounding (e.g., by impaired renal function) and/or reverse causality, rather than a causal effect of cystatin C per se, likely explained the observational relationship between cystatin C and clinical events (Central Illustration). As such, interventions aimed at lowering circulating cystatin C are unlikely to represent an effective means to prevent CVD.
COMPETENCY IN MEDICAL KNOWLEDGE: Epidemiological studies show a strong association between circulating cystatin C concentrations and cardiovascular risk, independent of renal function, but the results of a large Mendelian randomization study do not support a causal relationship.

PERSPECTIVES

REFERENCES


KEY WORDS coronary heart disease, genetics, heart failure, ischemic stroke

APPENDIX For an expanded Methods section and supplemental figures and tables, please see the online version of this article.