

Cystatin C and cardiovascular disease

An Example Mendelian Randomization study

Sander W. van der Laan • s.w.vanderlaan-2@umcutrecht.nl • @swvanderlaan



UMC Utrecht
Center for Circulatory Health



The concept of Mendelian Randomization

APOLIPOPROTEIN E ISOFORMS, SERUM CHOLESTEROL, AND CANCER

SIR,—It is unclear whether the relation between low serum cholesterol levels and cancer¹ is causal. In many studies occult tumour may have depressed cholesterol levels though in others the relation was found when serum cholesterol had been measured many years before the cancer was diagnosed. The relation is probably not explained by diet, because in the Seven Countries Study² men with different diets had corresponding differences in mean cholesterol levels experienced similar mean cancer rates.^{3,4} On the other hand, within each region cancer incidence was higher in men with a serum cholesterol in the lowest part of the cholesterol distribution for that country.⁵ Thus, naturally low cholesterol levels are sometimes associated with increased cancer risk.^{1,3}

Differences in the aminoacid sequence of apolipoprotein E (apo E) are major determinants of differences in plasma cholesterol levels within a population. Apo E has a key role in the clearance of cholesterol from plasma.⁶ The synthesis of apo E is under the control of several independent loci, each containing genes coding for the three major isoforms, E-2, E-3, and E-4, with respective population frequencies of about 8, 77, and 15%.⁵ The homozygous E-3/E-3 is the most common phenotype encountered and E-2/E-2 is the least common. From apo E-2 to apo E-3, one cysteine residue is replaced by arginine, and from apo E-3 to apo E-4 another cysteine residue is replaced. As a result the avidity of apo E containing lipoproteins for lipoprotein receptors increases from apo E-2 to apo E-3 to apo E-4. In several populations,^{6,8} including the Finns and the Japanese (Dr G. Utermann, personal communication), the gradient in serum cholesterol within the population is associated with a gradient in apo E phenotype. E-2 being associated with lower serum low-density lipoprotein and total cholesterol levels than E-3 and E-4. Thus, if a naturally low cholesterol favours tumour

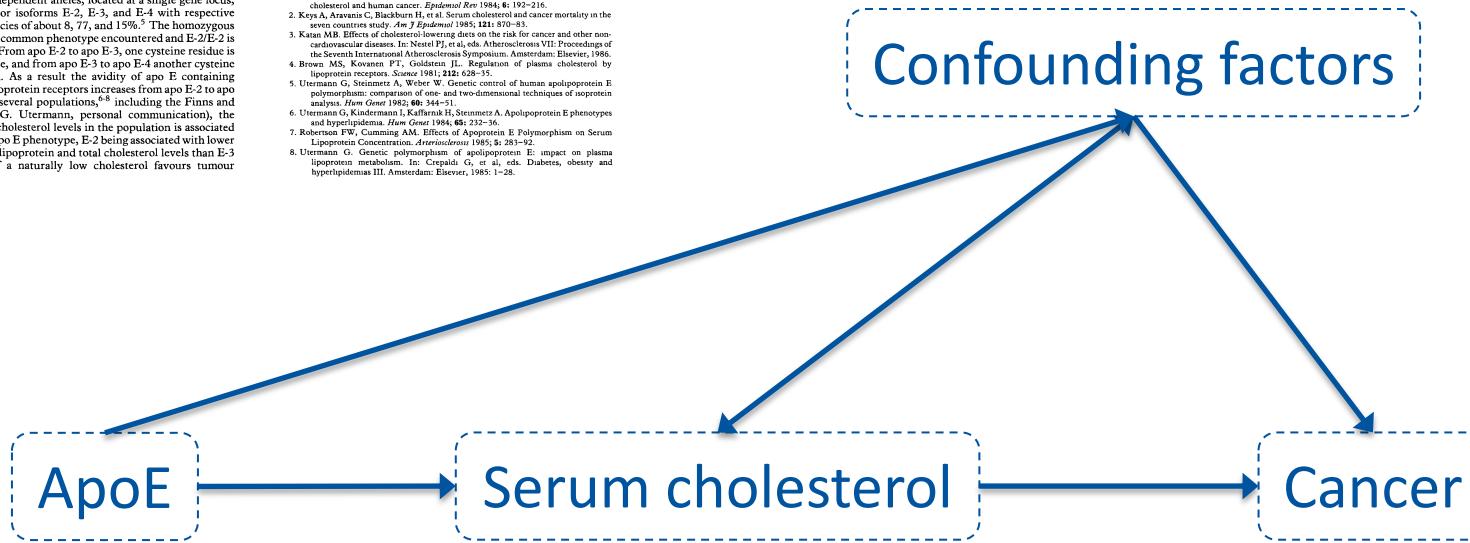
growth, then subjects with the E-2/E-2 or E-2/E-3 phenotype should have an increased risk of cancer.

Unlike most other indices of lipid metabolism, apolipoprotein aminoacid sequences are not disturbed by disease, and the apo E phenotype found in a patient will have been present since birth. A comparison of apo E phenotypes in cancer patients with those in matched controls might thus shed light on the relation between low cholesterol and cancer. If it is causal then the E-2 allele should be more common among patients and E-3 and E-4 more common among controls. On the other hand, equal distribution of apo E phenotypes among cases and controls would suggest that the association between low cholesterol and cancer is spurious. Measurement of apo E phenotype by isoelectric focusing of plasma is a routine determination in lipid laboratories; epidemiologists interested in cholesterol and cancer should include it in their studies.

Department of Human Nutrition,
Agricultural University,
6703 BC Wageningen, Netherlands

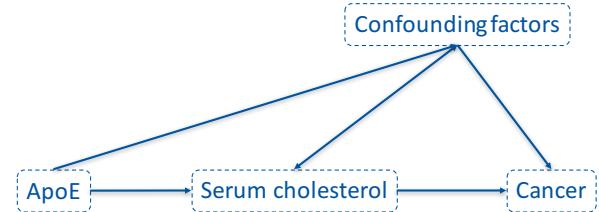
MARTIJN B. KATAN

1. McMichael AJ, Jensen OM, Parkin DM, Zaridze DG. Dietary and endogenous cholesterol and human cancer. *Epidemiol Rev* 1984; **6**: 193–216.
2. Keys A, Aravanis C, Blackburn H, et al. Serum cholesterol and cancer mortality in the Seven Countries Study. *N Engl J Med* 1985; **312**: 870–83.
3. Katan MB. Effects of cholesterol-lowering diets on the risk for cancer and other non-cardiovascular diseases. In: Nestel PJ, et al., eds. *Atherosclerosis VII: Proceedings of the Second International Conference*. Amsterdam: Elsevier, 1986.
4. Brown MS, Kovacs PT, Goldstein JL. Regulation of plasma cholesterol by lipoprotein receptors. *Science* 1981; **212**: 628–35.
5. Utermann G. Human apolipoprotein E polymorphism: comparison of one- and two-dimensional techniques of isoprotein analysis. *Hum Genet* 1982; **60**: 344–51.
6. Utermann G, Kühlmann T, Kahlffer S, Steinonen A. Apolipoprotein E phenotypes and hyperlipidemia. *New Eng J Med* 1984; **300**: 232–36.
7. Robertson FW, Cumming AM. Effects of Apolipoprotein E Polymerisation on Serum Lipoprotein Concentrations. *Atherosclerosis* 1985; **51**: 285–92.
8. Utermann G. Human apolipoprotein E polymorphism: its impact on plasma lipoprotein metabolism. In: Crepaldi G, et al., eds. *Diabetes, obesity and hyperlipidemias III*. Amsterdam: Elsevier, 1985: 1–28.



The concept of Mendelian Randomization

- Use genetic variants *independent of any confounder*
- Include *prospective studies* to reduce reverse causation
- Include *deeply phenotyped studies* to correct for potential confounders in the observational analysis





Cystatin C and Cardiovascular Disease **AN EXAMPLE MR STUDY**



Cystatin C and Cardiovascular Disease

BACKGROUND

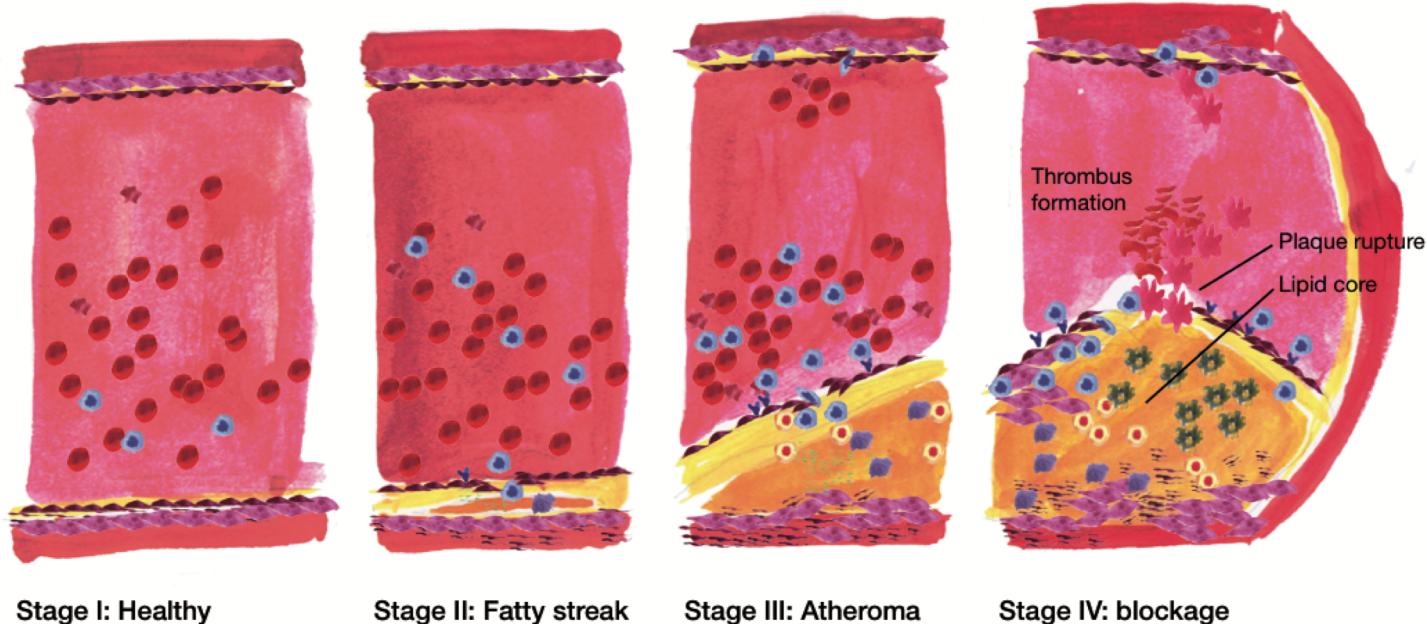


UMC Utrecht
Center for Circulatory Health

Atherosclerosis is the primary underlying cause of cardiovascular disease (CVD)

Legend

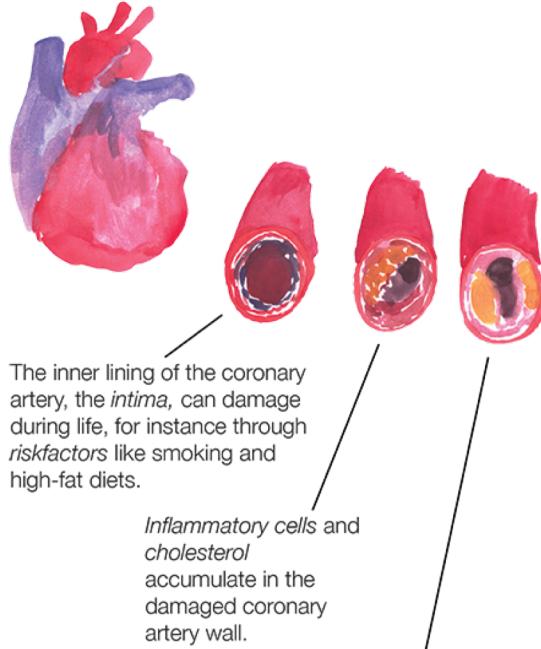
- Monocyte
- Macrophage
- Foam cell
- Apoptotic foam cell
- Red blood cell
- Platelet
- Activated platelet
- Smooth muscle cell
- Endothelial cell
- Extracellular matrix
- Matrix metalloproteases
- Receptor



CVD: myocardial infarction (“heart attack”) or cerebrovascular accident (“stroke”)

Heart attacks

The average human heart beats at 60 bpm; just like the beat of a slow song. Energy is necessary for that beat in the form of oxygen, which the heart gets through the coronary arteries ('the coronaries').



The inner lining of the coronary artery, the *intima*, can damage during life, for instance through *riskfactors* like smoking and high-fat diets.

Inflammatory cells and cholesterol accumulate in the damaged coronary artery wall.

Gradually, over the course of many decades, an *atherosclerotic plaque* forms and progresses to the point that it blocks the coronary artery to cause a **heart attack**, or ruptures releasing its contents to block another artery.

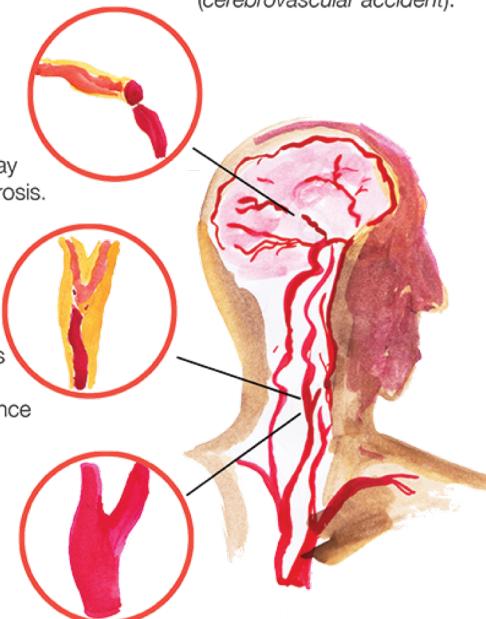
Strokes

The free floating thrombus, or *embolus*, can clog up arteries downstream, causing a **stroke** (cerebrovascular accident).

A *thrombus* can break away from the site of atherosclerosis.

Like in the heart riskfactors can cause *atherosclerosis* in for instance the carotid artery.

The average human brain uses about 20% of the total energy a body requires per day. The brain receives this energy, like oxygen and nutrients, via the healthy left and right *carotid* and *vertebral arteries*.



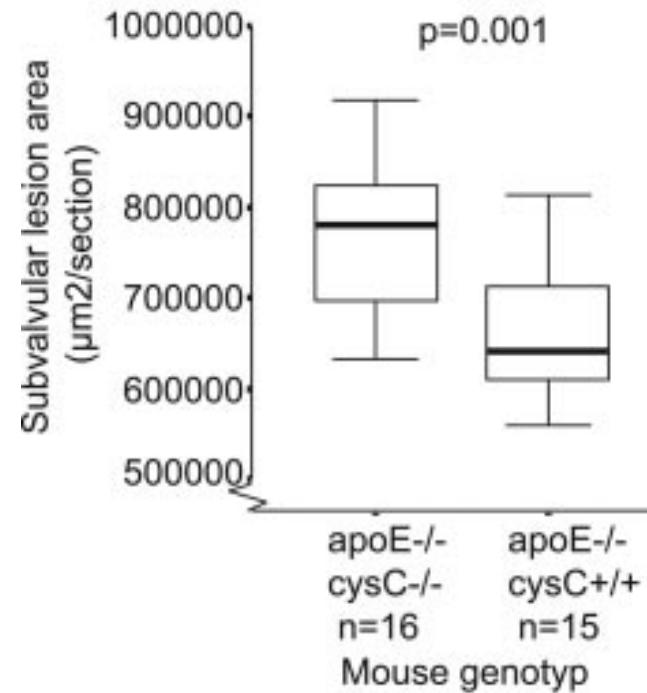
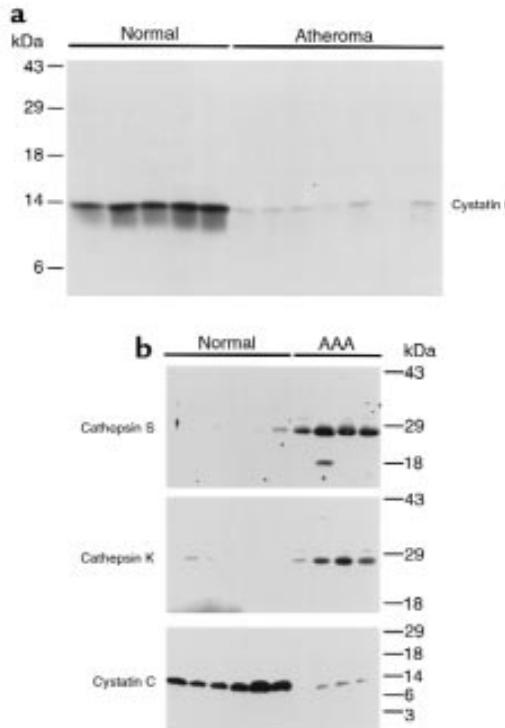
Cystatin C (CystC)

- Encoded by *CST3* on 20p11.21
- “Protease inhibitor”
 - It inhibits the activity of cathepsins (cysteine proteases) in vascular smooth muscle cells
 - Prevention of plaque formation by binding amyloid-beta precursor protein (APP) in Alzheimer’s disease



A role in human and murine atherosclerosis

- Reduction of CystC in abdominal aortic aneurysm (in smooth muscle cells)
- Lack of CystC in ApoE^{-/-} mice promotes atherosclerosis



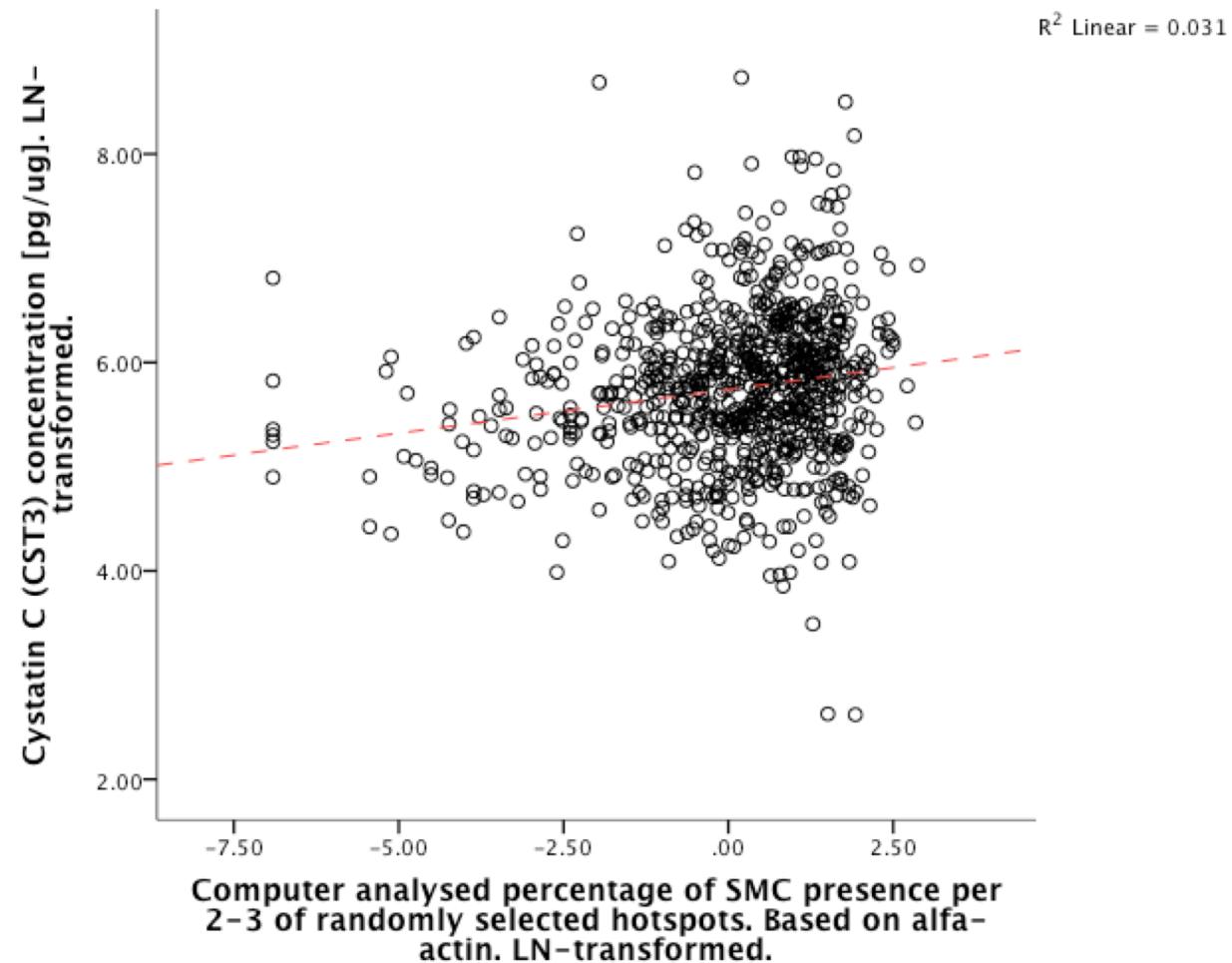
Shi *et al.* JCI 1999; Bengtsson *et al.* ATVB 2005



In carotid plaques CystC positively correlates with the percentage of smooth muscle cells

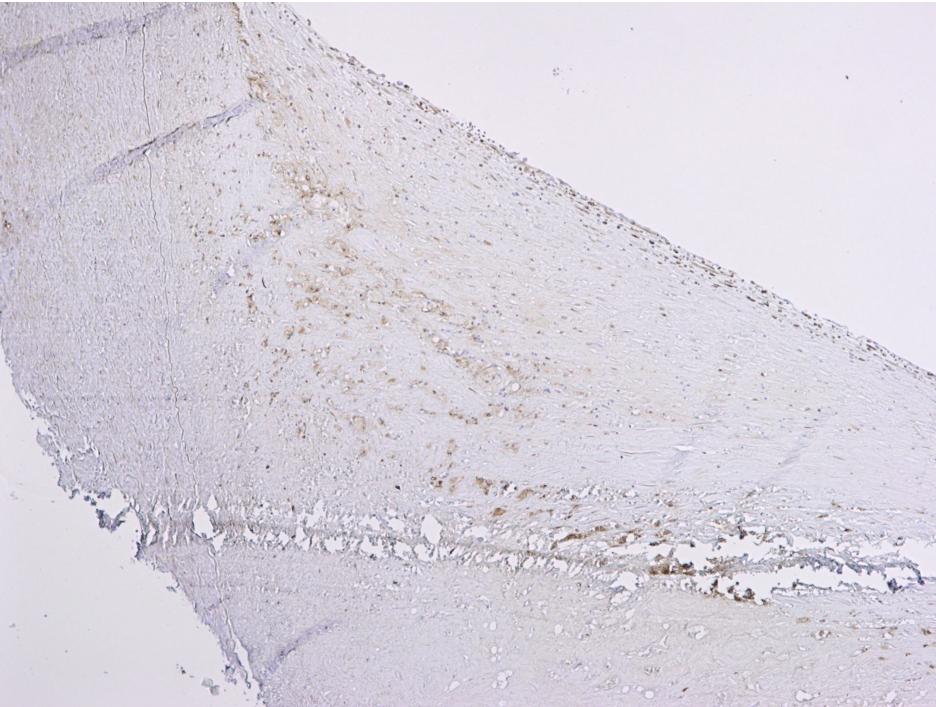
Rho = 0.185 (R = 0.176)

p = 1.10x10⁻⁷

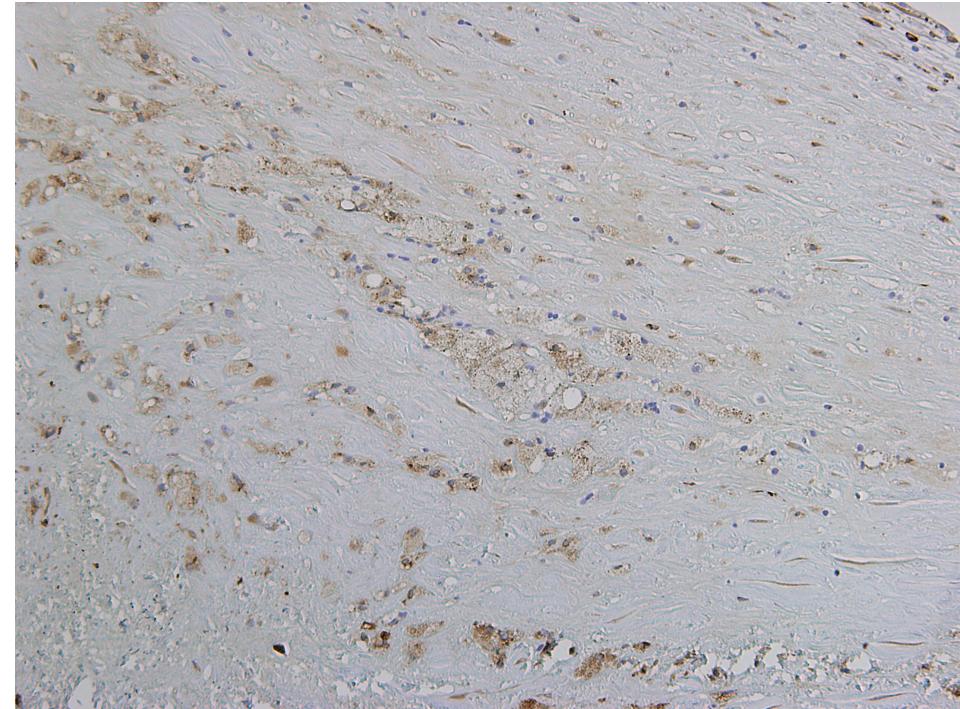


Carotid endarterectomy sample

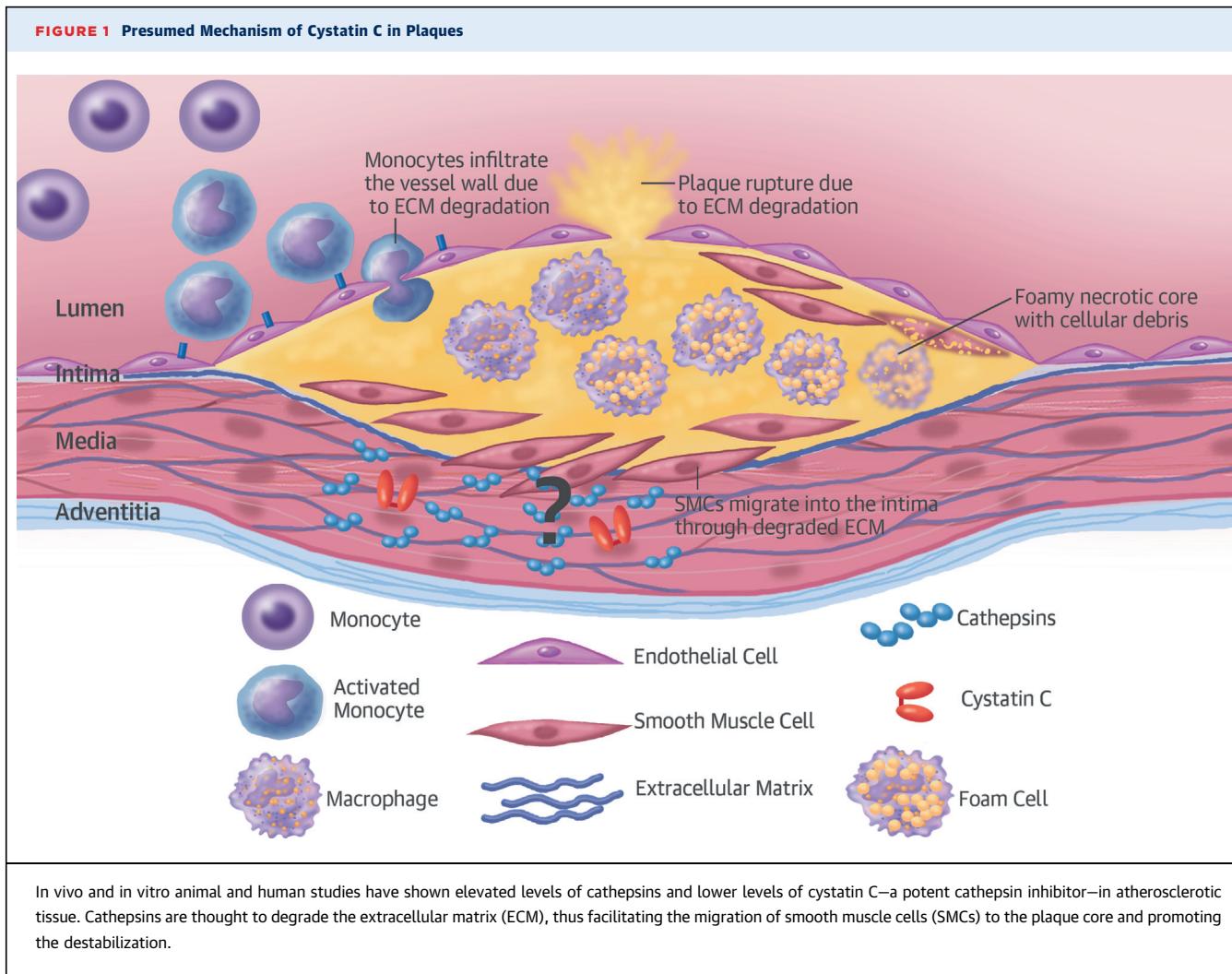
- Shoulder region, clear staining for CystC
- Appears to be stronger in foamy cells, (foamy) smooth muscle cells



Van der Laan *et al.* unpublished

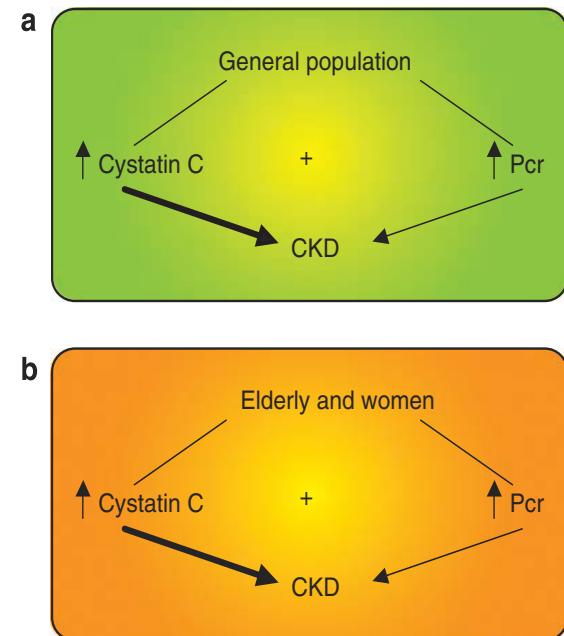
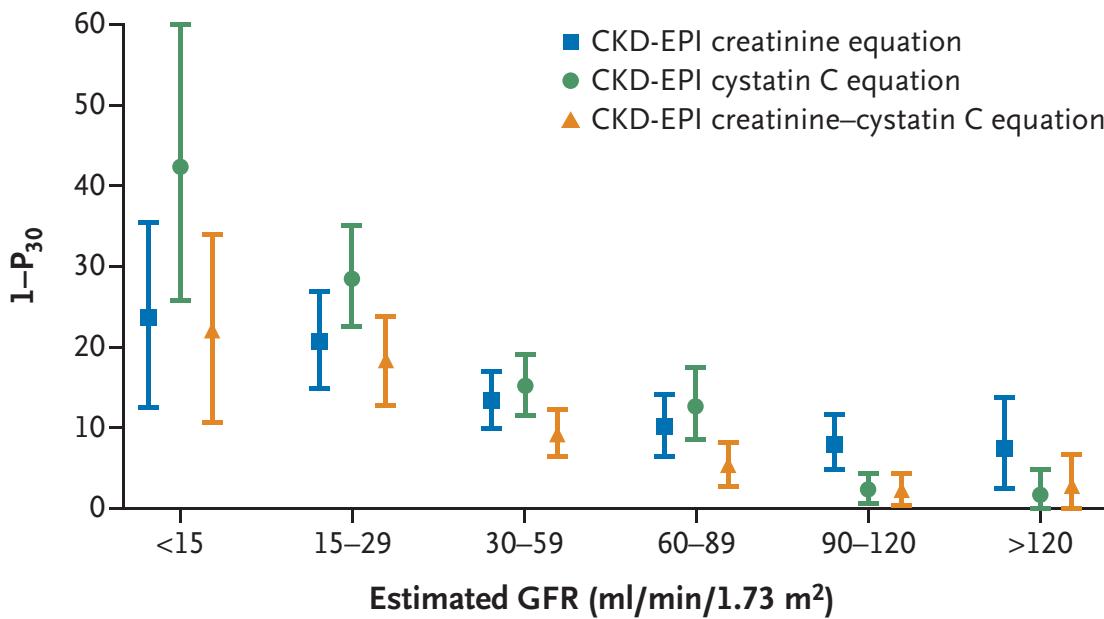


CystC in atherosclerosis



CystC as a biomarker for CKD

- One of the best naturally occurring biomarker for human estimated glomerular filtration rate (eGFR)
- It is a better biomarker than creatinine when eGFR>90 (mL/min/1.73m³)



Fried *et al.* Kid Intern 2009; Inker *et al.* NEJM 2012

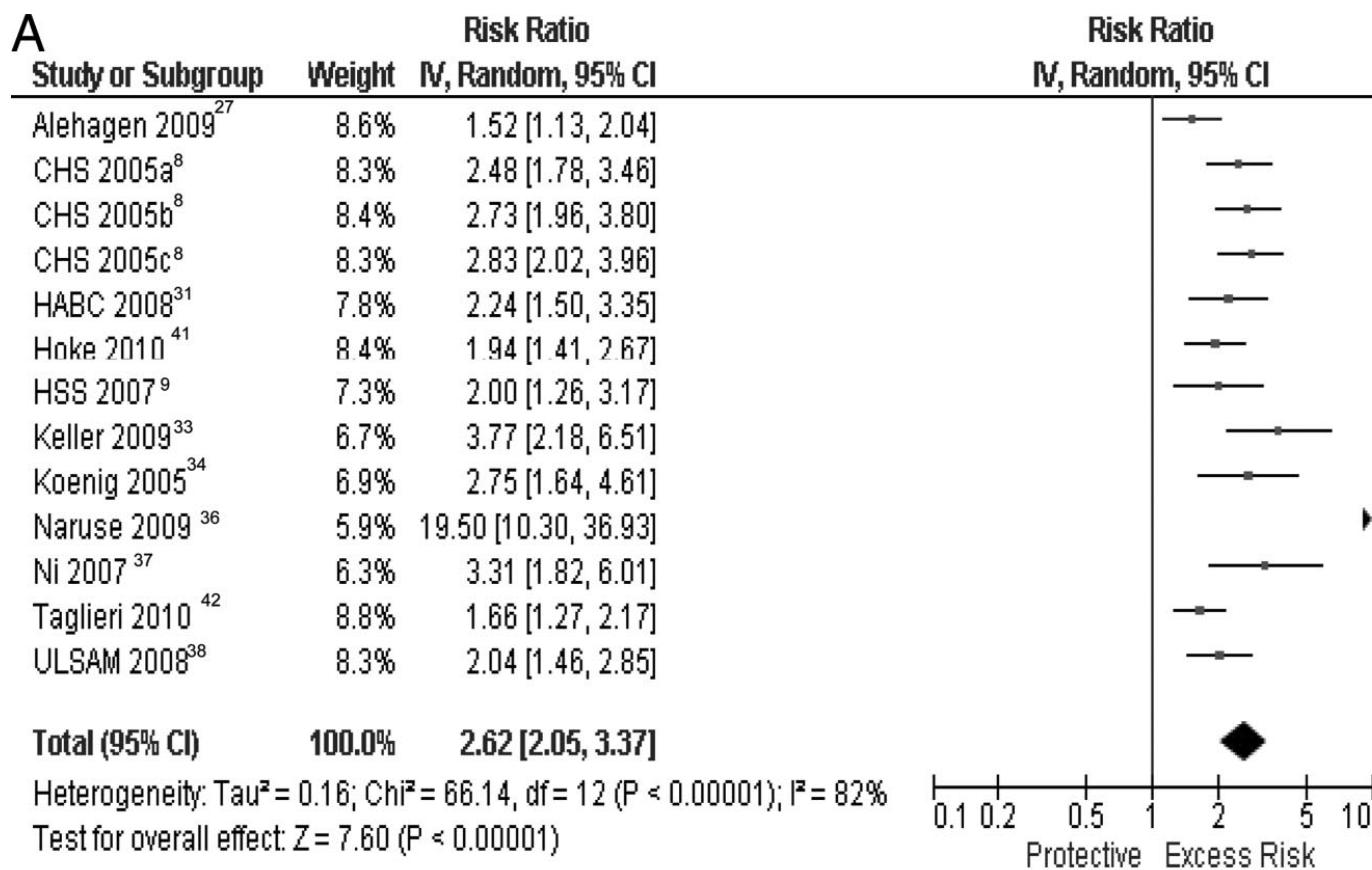
Epidemiological evidence

- In elderly without chronic kidney disease (CKD), CystC is prognostic of death, cardiovascular disease (CVD) and CKD
- Asymptomatic carotid atherosclerosis patients, CystC prognostic for CVD
- CystC associated with lesion size in acute stroke



Literature meta-analyzed

- CystC is associated with a higher risk of CVD, stroke and mortality in high-risk populations



Clinical applications

- 2 observational studies and 5 clinical trials focused on
 - Biomarker for diagnosis and prognosis
 - Marker of efficacy (treatment)
 - At least one drug in development: “RO4905417/INCLACUMAB” to reduce damage during PCI (Roche)
- Patents
 - Cavadis B.V.: exosome derived biomarker for adverse cardiovascular events (several patents [pending])
 - Biomarker for peripheral artery disease (2469279)
 - CystC as a “drug” for the treatment of atherosclerosis (WO/2001/030370)
 - Roche: selection of therapy in heart failure (WO/2014/086833)
 - Siemens: biomarker for pulmonary hypertension (WO/2014/052803)



Hypothesis...

- There is a large body of evidence from human and murine experiments and specimens that CystC is involved in atherosclerosis
- Epidemiological studies show that CystC is associated with increased risk CVD
- Plaque CystC and the percentage of SMCs in plaques are correlated

Is CystC causal to cardiovascular disease?





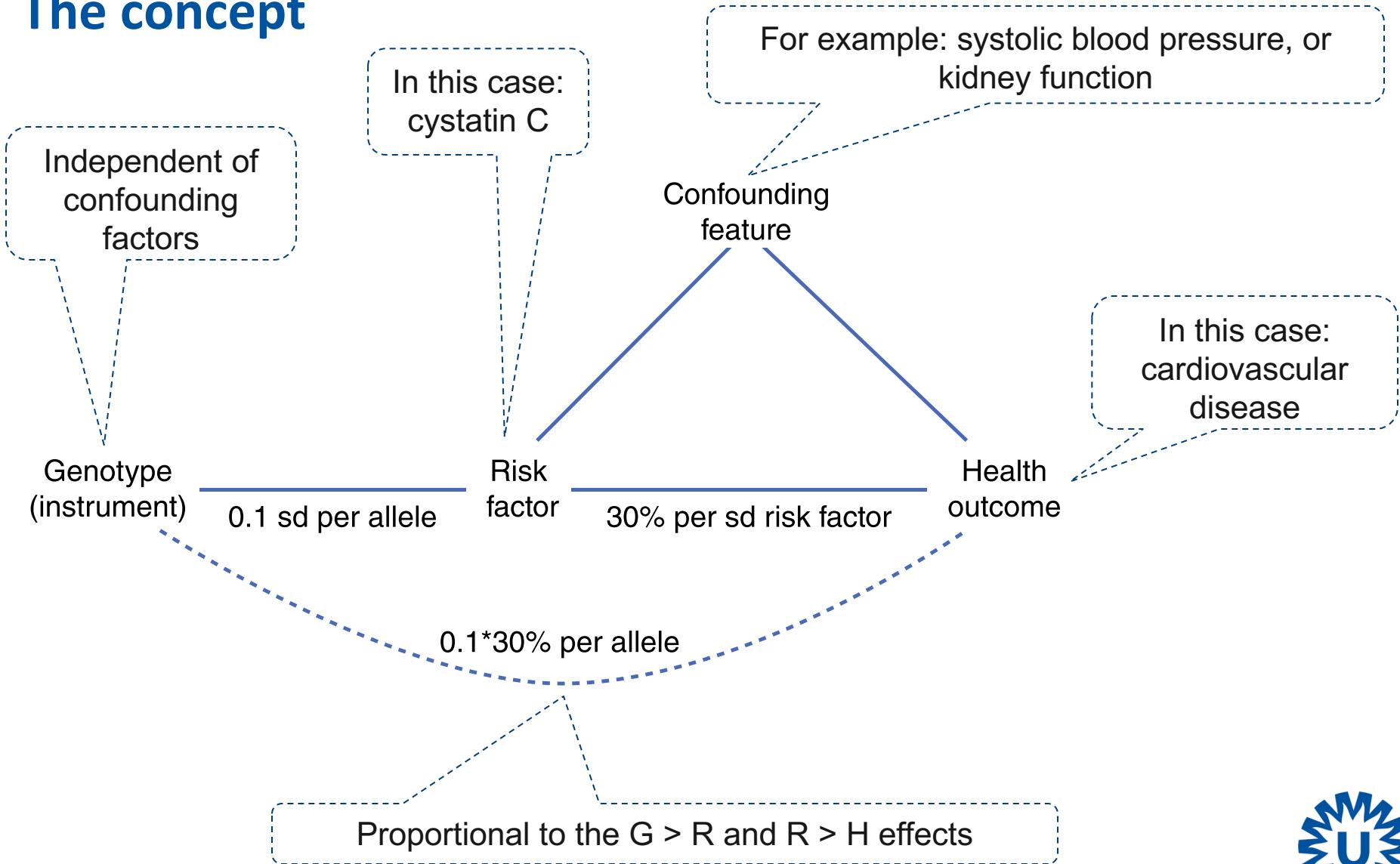
Cystatin C and Cardiovascular Disease

FULFILLING ASSUMPTIONS



UMC Utrecht
Center for Circulatory Health

The concept



Fulfilling assumptions

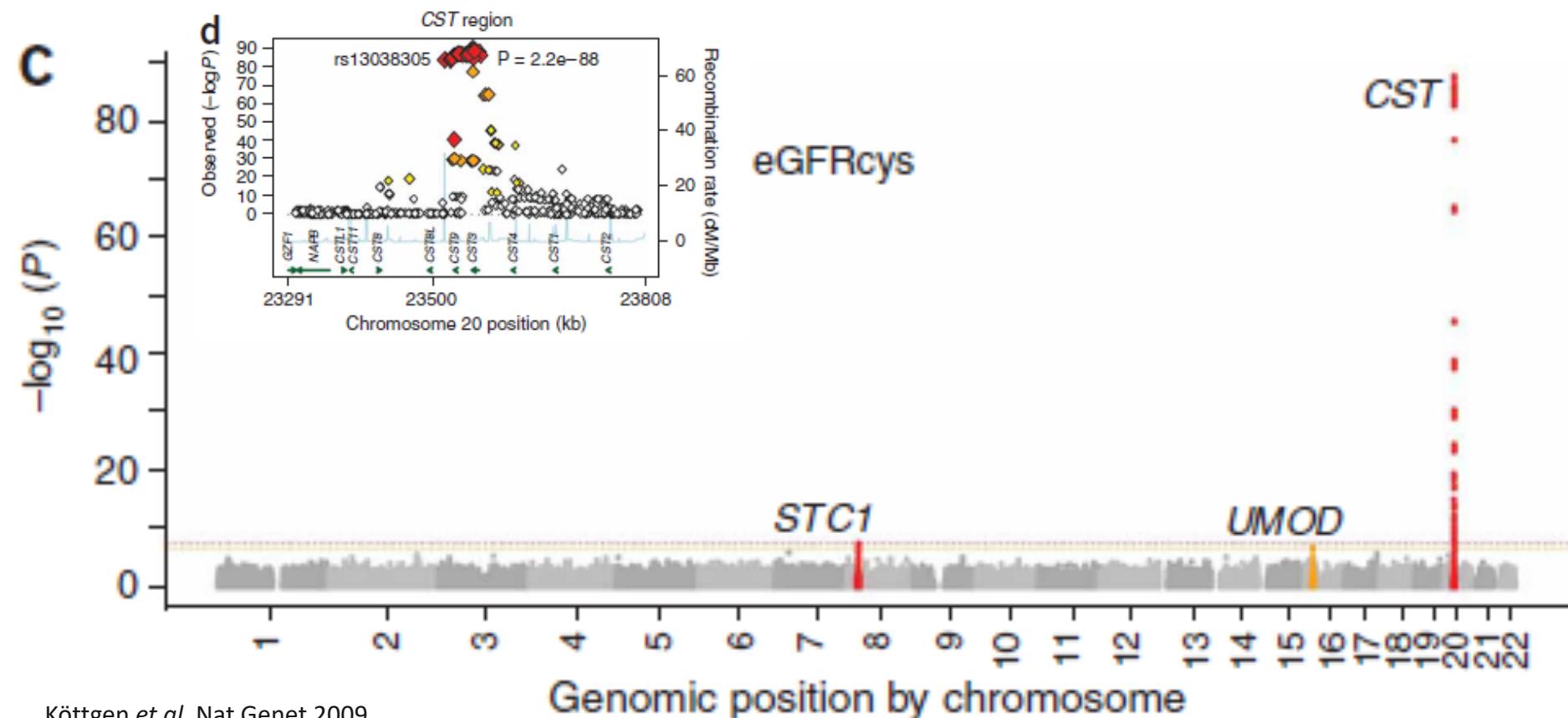
SNP SELECTION



UMC Utrecht
Center for Circulatory Health

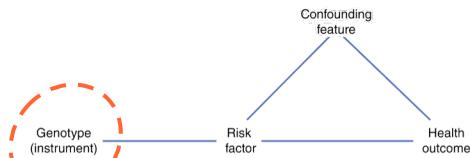
GWAS: locus with *CST3* pops up (naturally)

- 4 subsequent Genome–Wide Association Study Meta-Analyses of eGFR and CKD revealed a region containing *CST3* (Cystatin C) as significantly associated
 - Identification of three loci associated with eGFRcys aka CystC expression: *STC1*, *UMOD*, *CST3*



Köttgen et al. Nat Genet 2009

SNP selection



- One GWAS on CystC, rs1158167 in Framingham Heart Study
- Three (meta-analyses of) GWAS on eGFR_{CystC}
 - rs911119, rs13038305
 - $eGFR_{CystC} = 76.7 \times (\text{serum CystC})^{-1.19}$
 - Log linear relation between serum CystC and eGFR

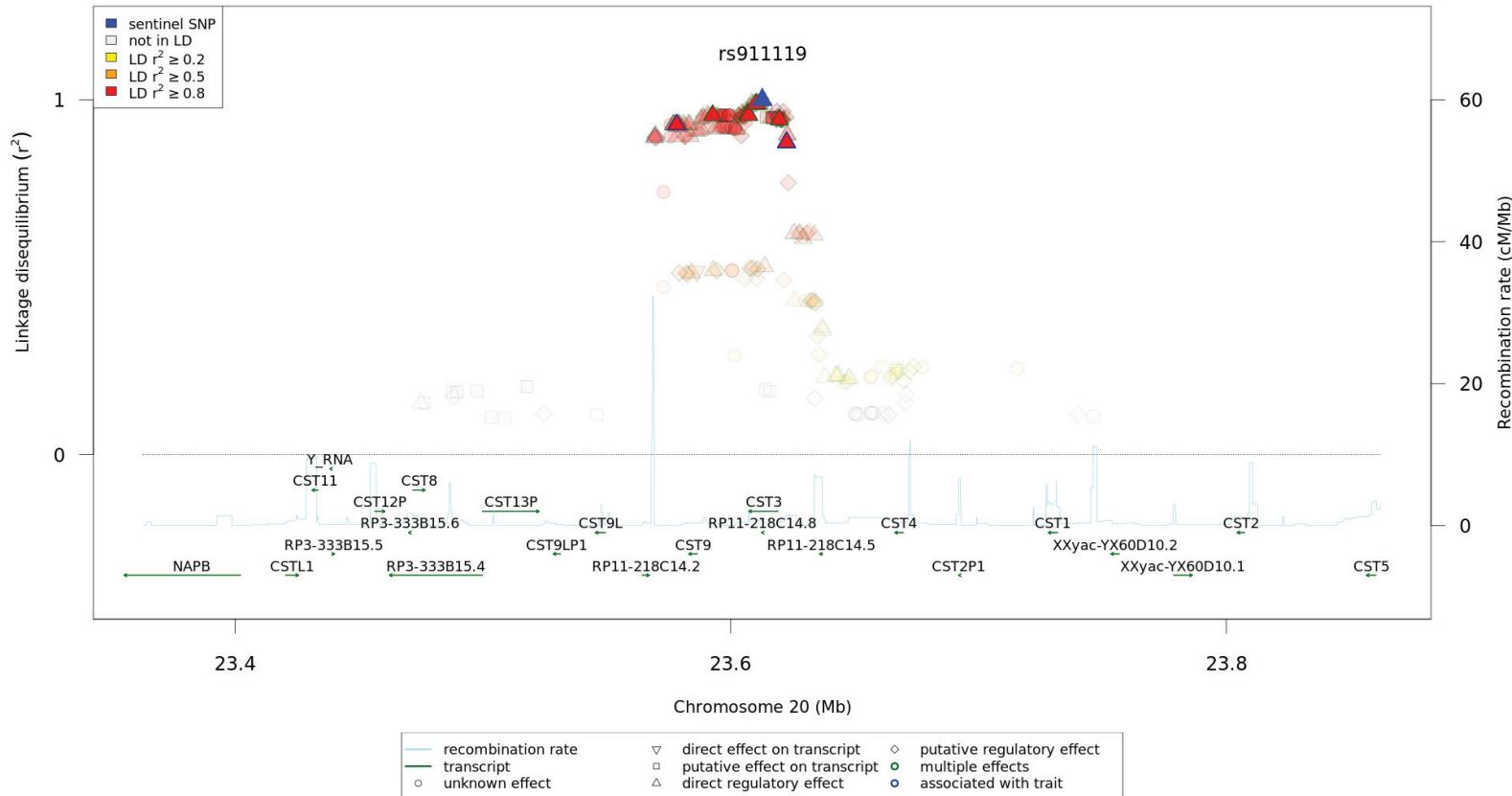
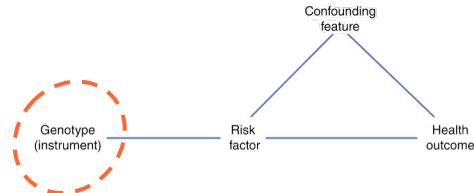
SUPPLEMENTAL TABLE 4: THE SELECTED SNP (RS911119, IN BOLD) AND ITS PROXIES USED IN THIS STUDY.

PROXY	DISTANCE	R ²	CHR	POSITION	MINOR	MAJOR	MAF	VARIANT TYPE	GENE
rs1158167	-34,548	0.913	20	23,526,189	G	A	0.240	downstream	-
rs17751897	-20,023	0.955	20	23,540,714	C	T	0.237	downstream	-
rs12625716	-5,892	0.955	20	23,554,845	A	G	0.236	downstream	-
rs6048952	-5,480	0.955	20	23,555,257	G	A	0.237	downstream	-
rs13038305	-2,475	1.000	20	23,558,262	T	C	0.233	intronic	CST3
rs911119	-	-	20	23,560,737	C	T	0.233	intronic	CST3
rs3827143	6,880	1.000	20	23,567,617	G	A	0.225	upstream	-
rs6114208	8,997	1.000	20	23,569,734	G	C	0.233	upstream	-
rs2424582	9,813	0.954	20	23,570,550	G	A	0.219	upstream	-

Distance of the SNP (in base pairs) from rs911119, linkage disequilibrium correlation (r^2), alleles, minor allele frequencies (MAF), and chromosomal positions are from 1000G pilot data (based on a European ancestral population [CEU]) available in SNAP and based on genome build 37³⁸⁴. The Variant type was determined using 1000G data from SNiPA¹¹⁴.



SNP selection

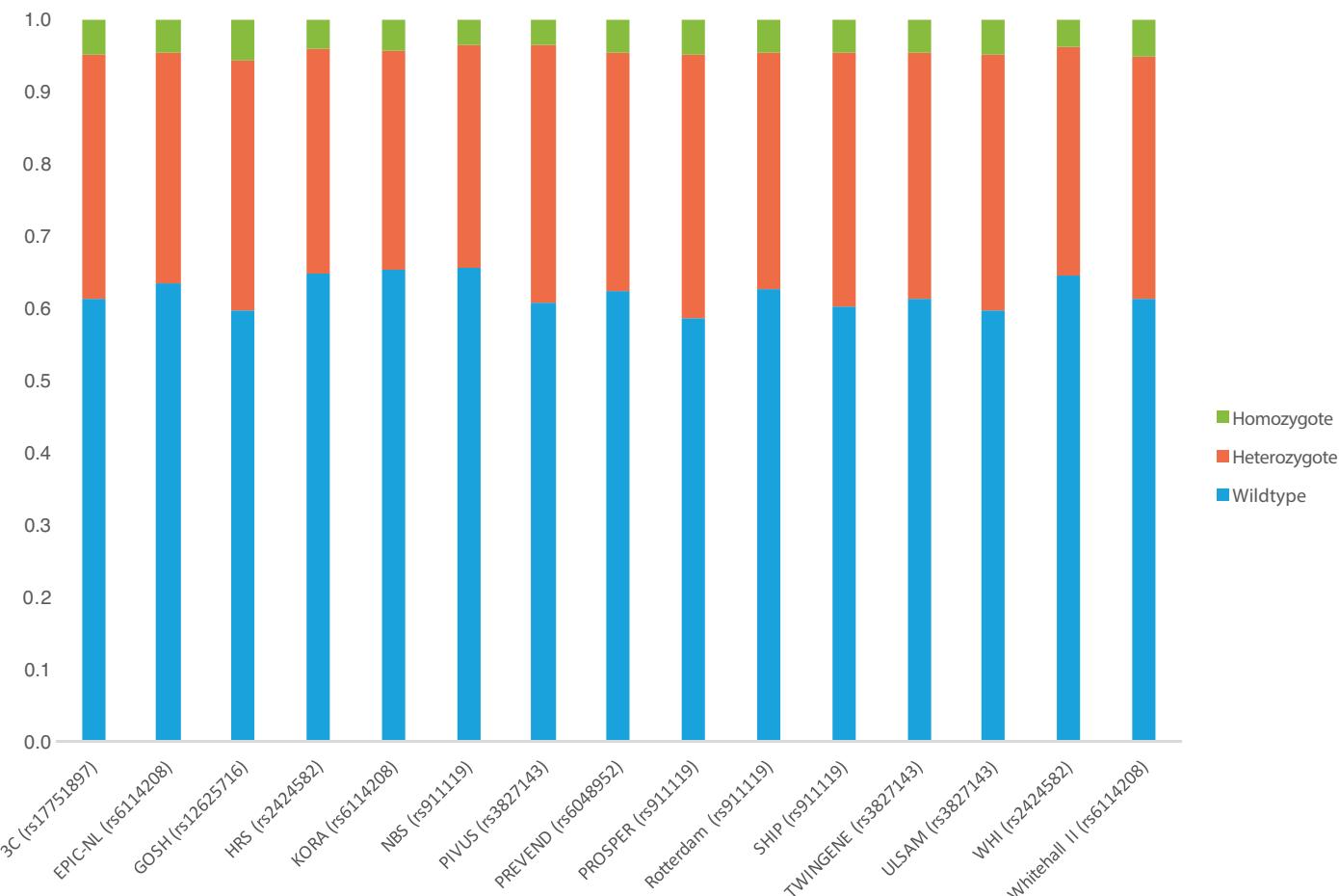
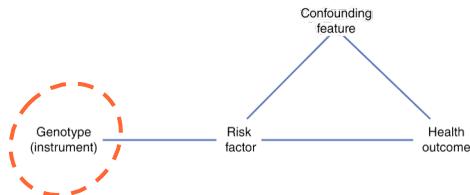


SUPPLEMENTAL FIGURE 1: OUR SENTINEL SNP (RS911119 IN BOLD) RELATIVE TO ITS GENOMIC LOCATION.

Each dot represents a SNP and is colored based on its linkage disequilibrium correlation (r^2) relative to rs911119 (y-axis). Chromosomal positions (x-axis) are from 1000G (phase 3, version 5, EUR population) available in SNiPA and based on genome build 37¹⁴. Functional SNP annotations are based on Ensembl77⁸⁵.



SNP selection



SUPPLEMENTAL FIGURE 2: GENOTYPE FREQUENCIES OF THE SNPs USED BY COHORTS IN THIS STUDY.

There was no SNP data available for the Tromsø Study.



Fulfilling assumptions

STUDIES & CONFOUNDING



UMC Utrecht
Center for Circulatory Health

Studies included

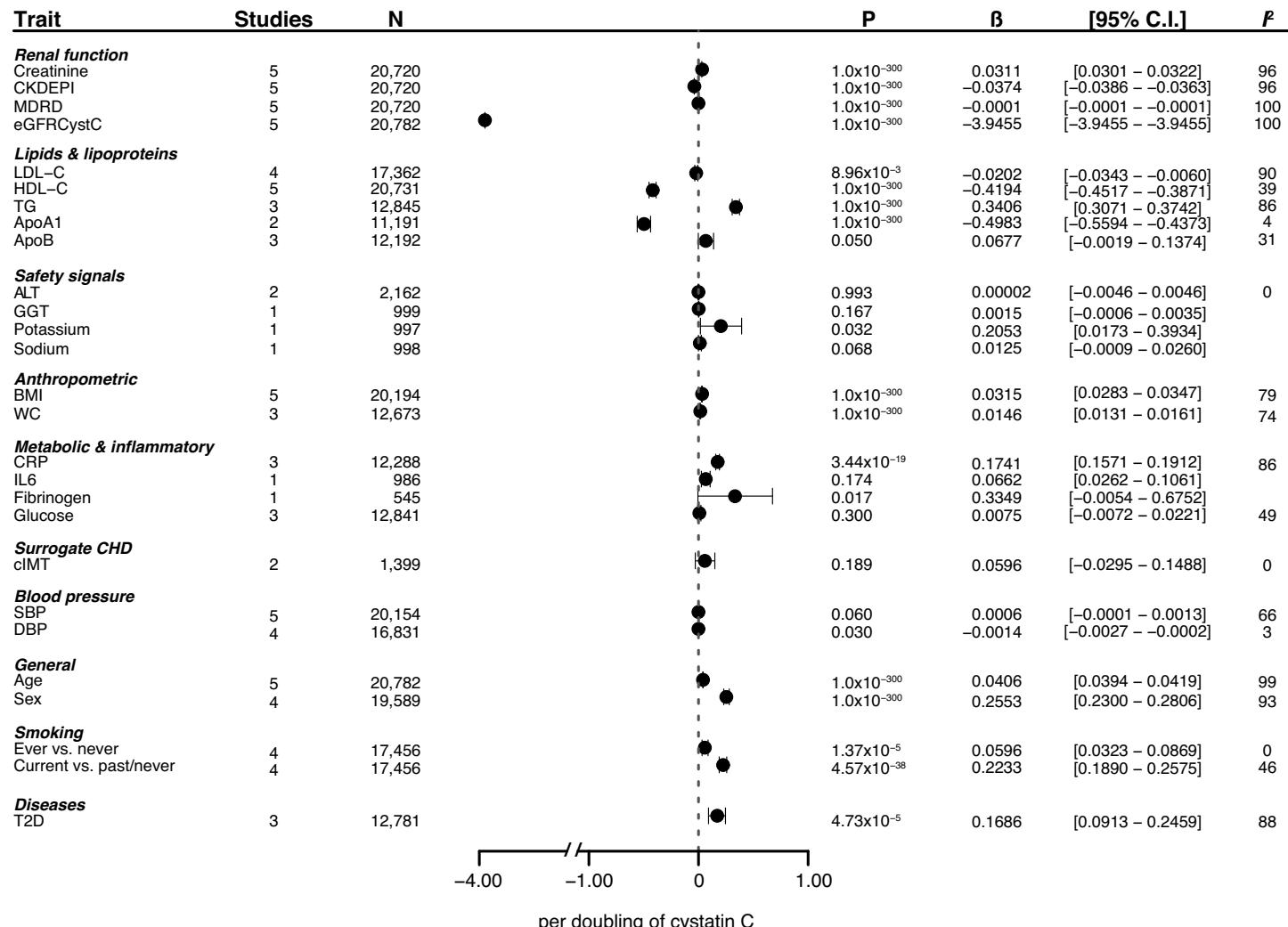
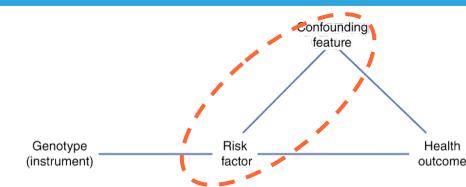
TABLE 1 Characteristics of Prospective Cohorts

Study	Total	SNP*	Cystatin C†	CVD‡	CHD‡	IS‡	HF‡	MI‡	Male	Age (yrs)	Cystatin C (mg/dl)
3C	6,440	6,435	1,244	1,717	1,235	459	439	486	39.19	74.30 ± 5.52	0.92 ± 0.24
EPIC-NL	6,265	5,192	—	1,967	1,430	537	—	1,430	22.39	53.80 ± 10.23	—
GOSH	1,478	1,479	—	493	111	235	233	—	42.08	51.08 ± 11.86	—
HRS	7,844	5,585	5,777	—	—	—	—	—	—	—	0.64 ± 0.34
KORA	4,856	1,867	4,676	540	341	255	—	341	49.53	49.75 ± 14.11	0.80 ± 0.21
NBS	1,819	1,297	—	66	—	66	—	170	49.48	61.05 ± 10.26	—
PIVUS	1,016	949	1,004	255	175	71	75	105	49.90	70.20 ± 0.17	0.90 ± 0.19
PREVEND	3,245	3,245	3,245	236	190	58	—	—	50.26	49.42 ± 12.25	0.87 ± 0.17
PROSPER§	5,244	5,150	—	2,561	2,034	779	211	762	48.13	75.34 ± 3.35	—
Rotterdam	7,983	5,974	3,906	3,579	1,934	1,328	1,625	1,176	38.90	73.06 ± 7.49	1.11 ± 0.28
SHIP	3,224	3,224	3,212	114	19	87	—	134	48.08	54.46 ± 15.26	0.88 ± 0.30
Tromsø	6,129	—	6,129	1,251	—	494	—	881	47.59	60.59 ± 10.25	0.86 ± 0.18
TWINGENE	6,902	6,902	6,740	932	610	287	206	—	47.23	64.83 ± 8.26	1.02 ± 0.30
ULSAM	1,221	1,107	1,193	503	285	175	220	—	100.00	71.00 ± 0.64	1.25 ± 0.27
WHI	7,854	7,844	—	4,831	2,934	2,115	—	2,934	0.00	67.97 ± 6.58	—
Whitehall II	4,961	5,011	—	349	254	111	—	254	74.58	49.19 ± 5.99	—
Overall	76,481	61,261	37,126	19,394	11,552	7,057	3,009	8,673	—	—	—

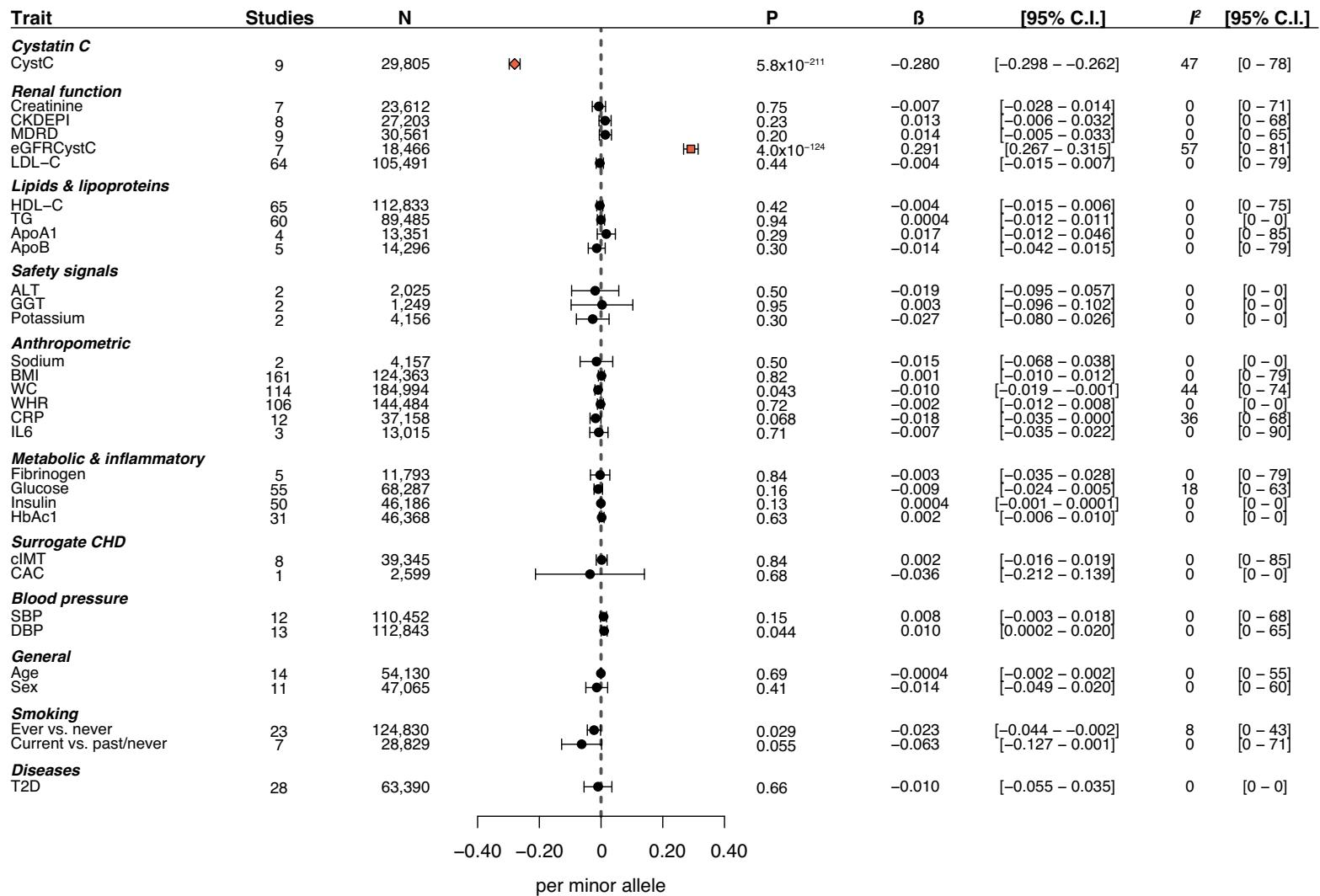
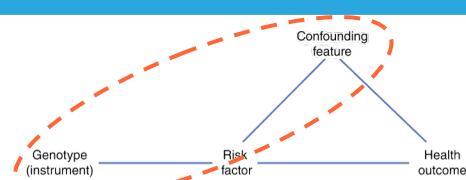
Values are n, %, or mean ± SD. *Total number of individuals with genotype data. †Genetic data were available in 29,805 of the 37,126 individuals that had values for cystatin C, which we used to associate rs911119 with circulating cystatin C. For the genetic analysis of CVD, CHD, IS, and HF, cohorts that contributed toward consortia were excluded. ‡Indicates total incident and prevalent cases of disease or composite diseases in the case of CVD. §PROSPER is a randomized clinical trial. ||For the association of SNP with cystatin C concentrations, 9,488 samples were available in TWINGENE.

CHD = coronary heart disease; CVD = cardiovascular disease; HF = heart failure; IS = ischemic stroke; MI = myocardial infarction; SNP = single-nucleotide polymorphism.

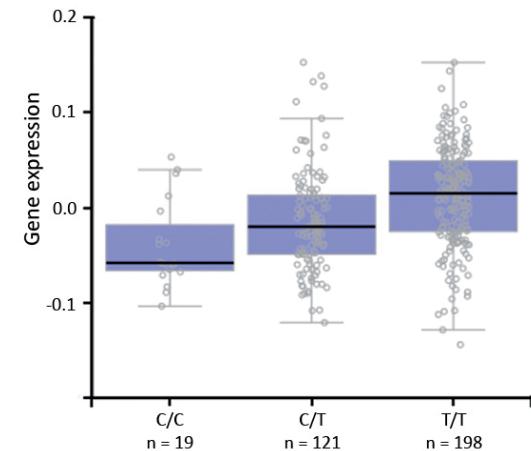
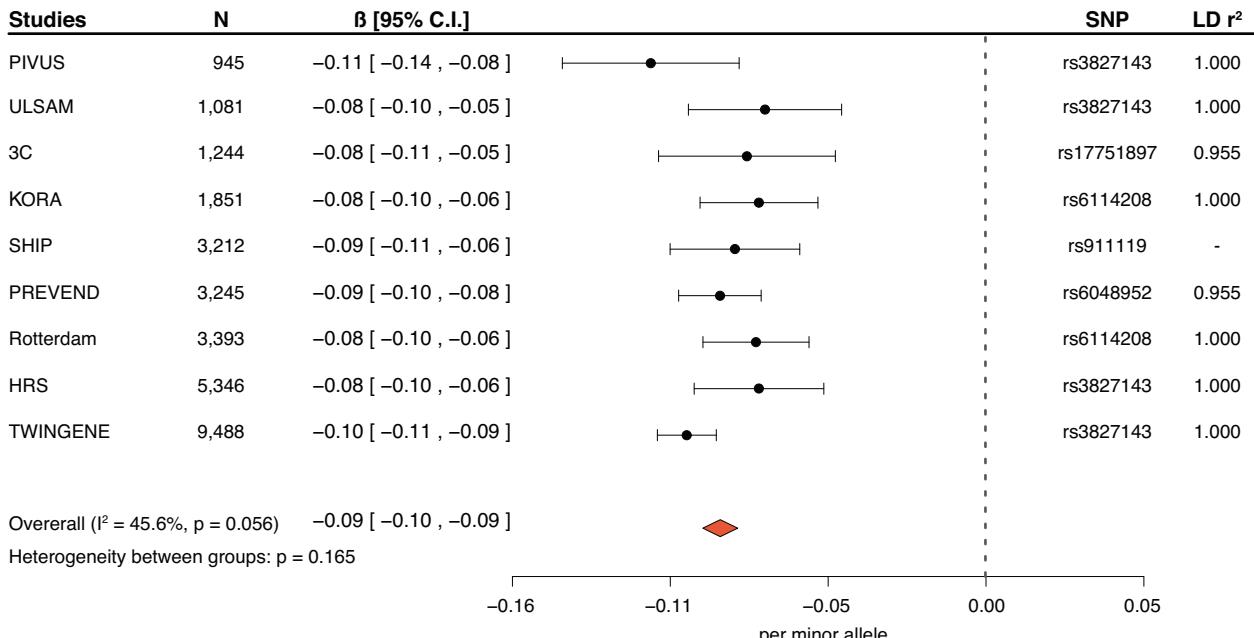
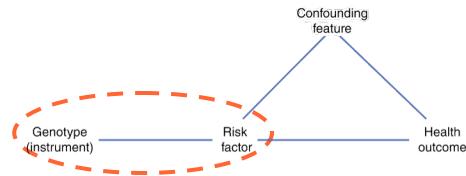
Serum cystatin C vs. risk factors



Cystatin C variant vs. risk factors



SNP vs. cystatin C



SUPPLEMENTAL FIGURE 4: ASSOCIATION OF GENETIC VARIANTS IN CST3 WITH CIRCULATING CYSTATIN C PER COHORT.

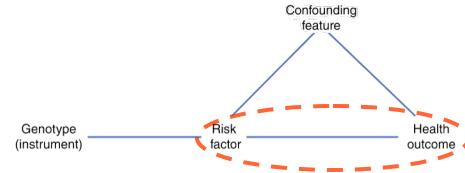
Effect sizes (β) are presented as log₂-transformed cystatin C concentrations in mg/L per minor allele. N indicates the total sample size. C.I.: confidence interval. SNP: the single-nucleotide polymorphism available for analysis. LD: the linkage disequilibrium r^2 relative to rs911119.

- Per minor allele there is 6.13% decrease in CystC [mg/L]
- $\beta = -0.09 [-0.10 - -0.09]$, $p = 5.95 \times 10^{-163}$, $N = 29,805$
- This explains $\approx 2.75\%$ of the phenotypic variation

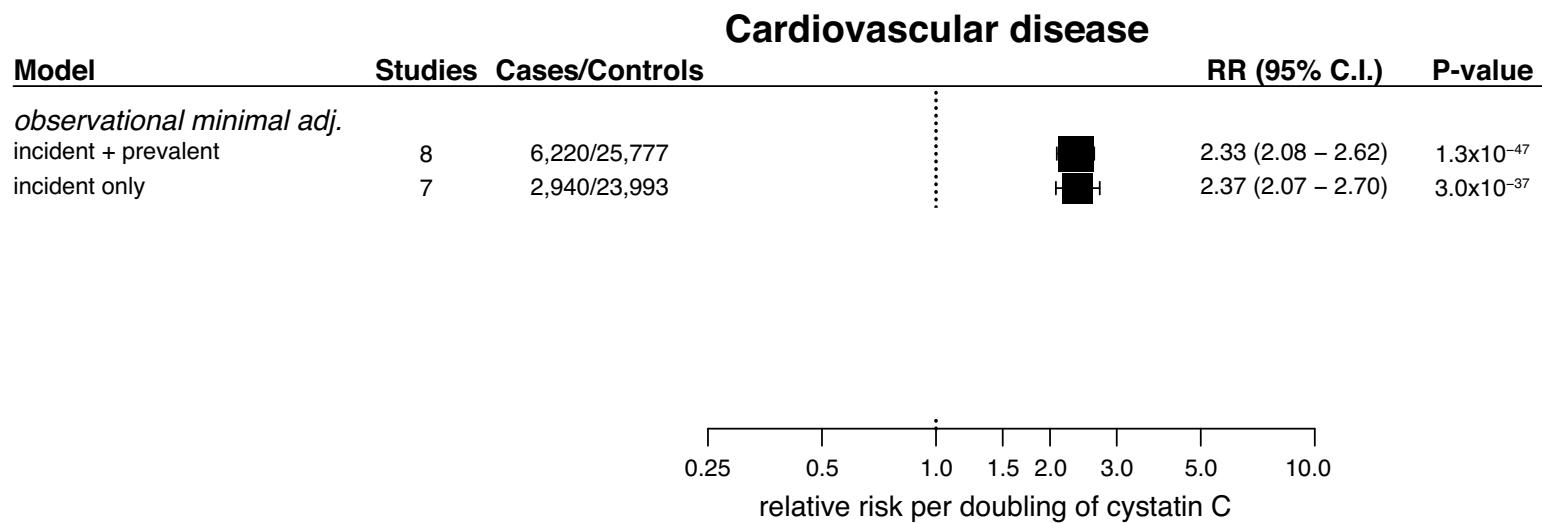
Van der Laan S.W. et al. JACC 2016



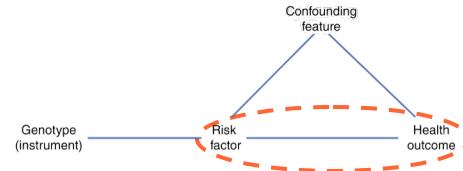
Observational analysis



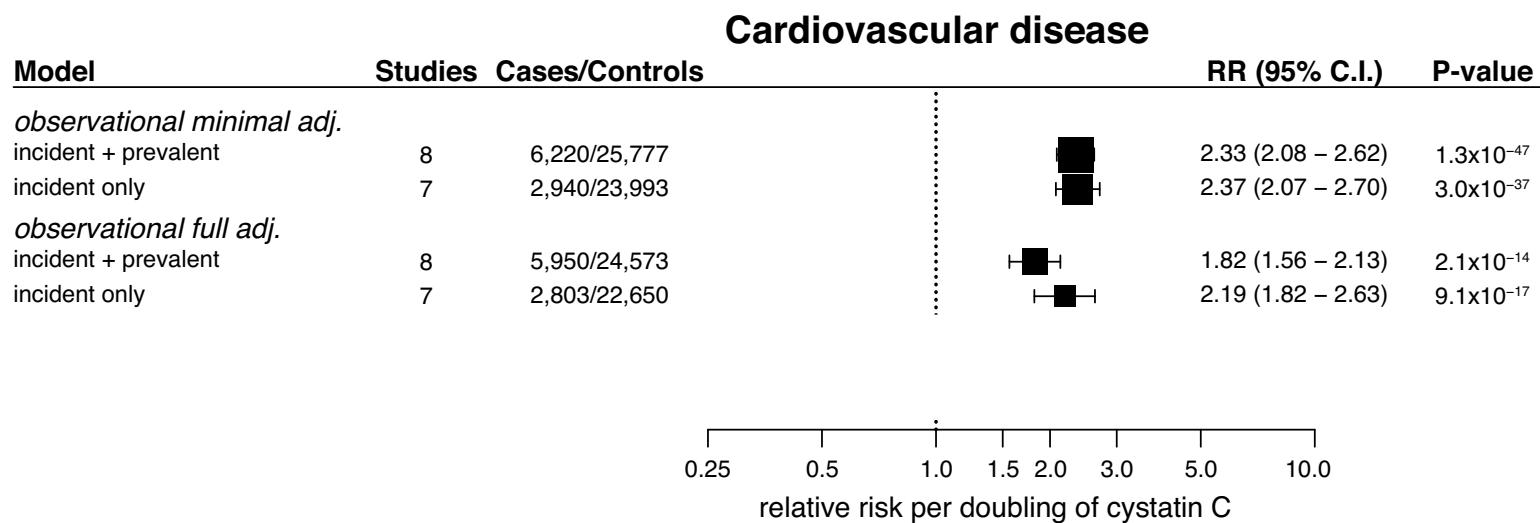
- Minimally adjusted model: age + sex



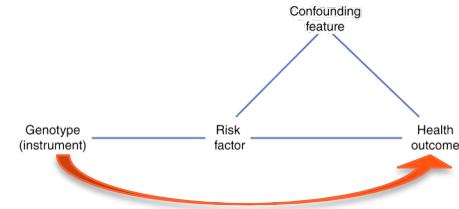
Observational analysis



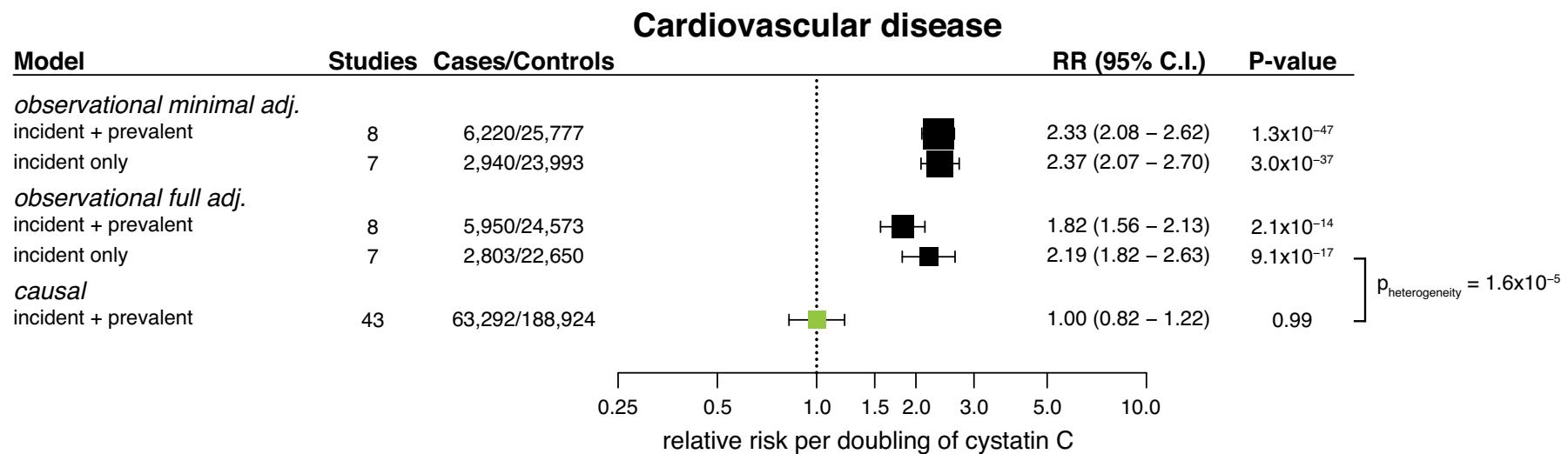
- Minimally adjusted model: age + sex
- Fully adjusted model: age + sex + HDL + systolic blood pressure + BMI + smoking status + CKDEPI



Mendelian randomization analysis



- Minimally adjusted model: age + sex
- Fully adjusted model: age + sex + HDL + systolic blood pressure + BMI + smoking status + CKDEPI
- Causal effect: *nada, nothing, niets, rien, nichts, ничего*
- Interaction analysis*: significant – models are different



Other outcomes

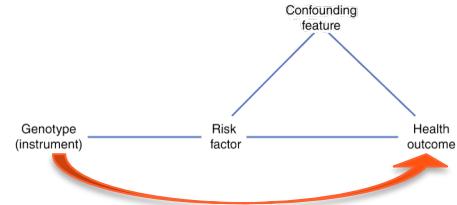
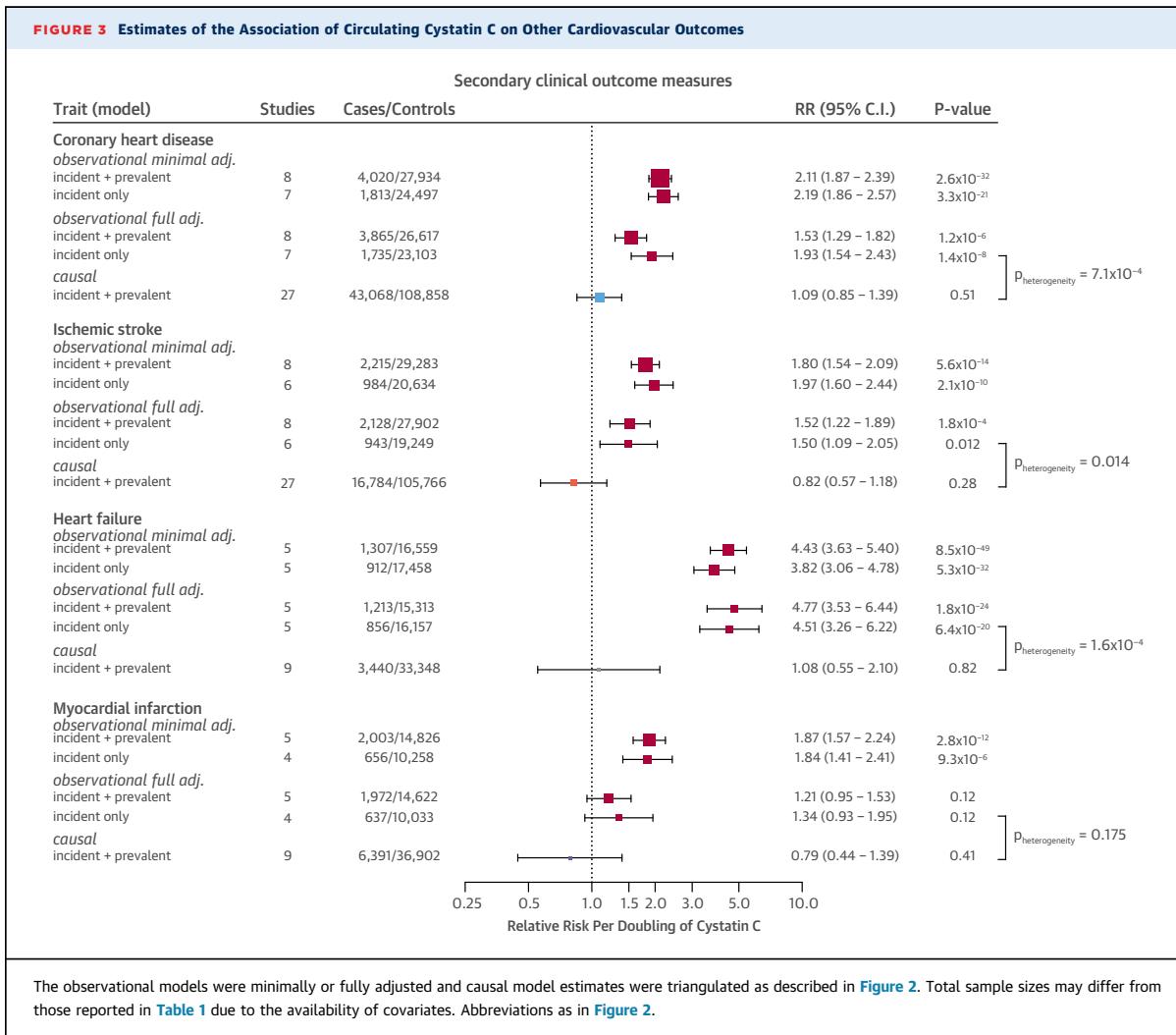


FIGURE 3 Estimates of the Association of Circulating Cystatin C on Other Cardiovascular Outcomes



Fulfilling assumptions
POWER?



UMC Utrecht
Center for Circulatory Health

Power

Power to detect association

LD with causal variant

$$E[\chi^2] \approx N\gamma^2 p(1-p)R^2$$

Effect size

Sample size

Allele frequency

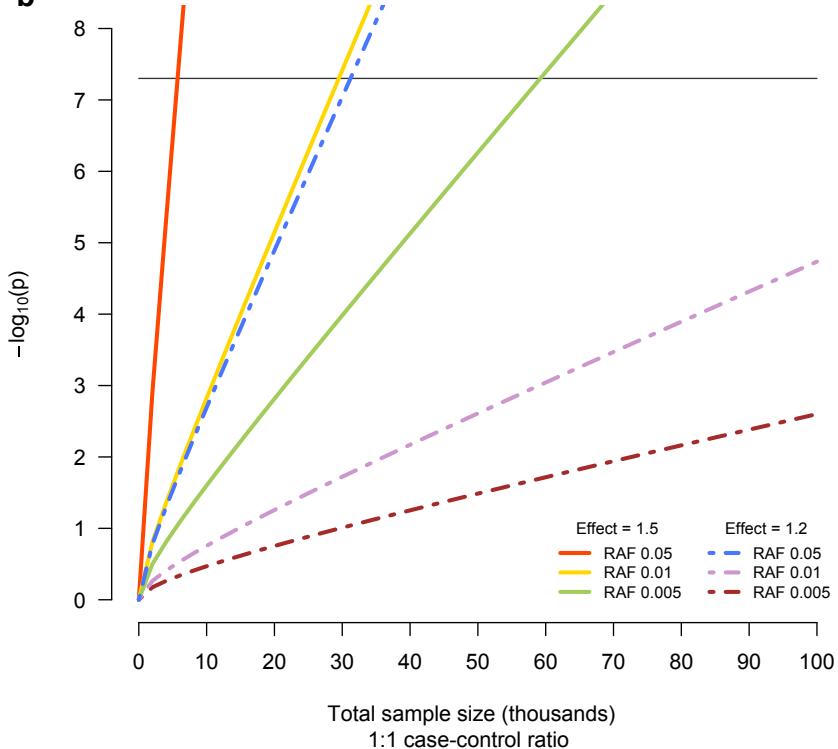


The power problem

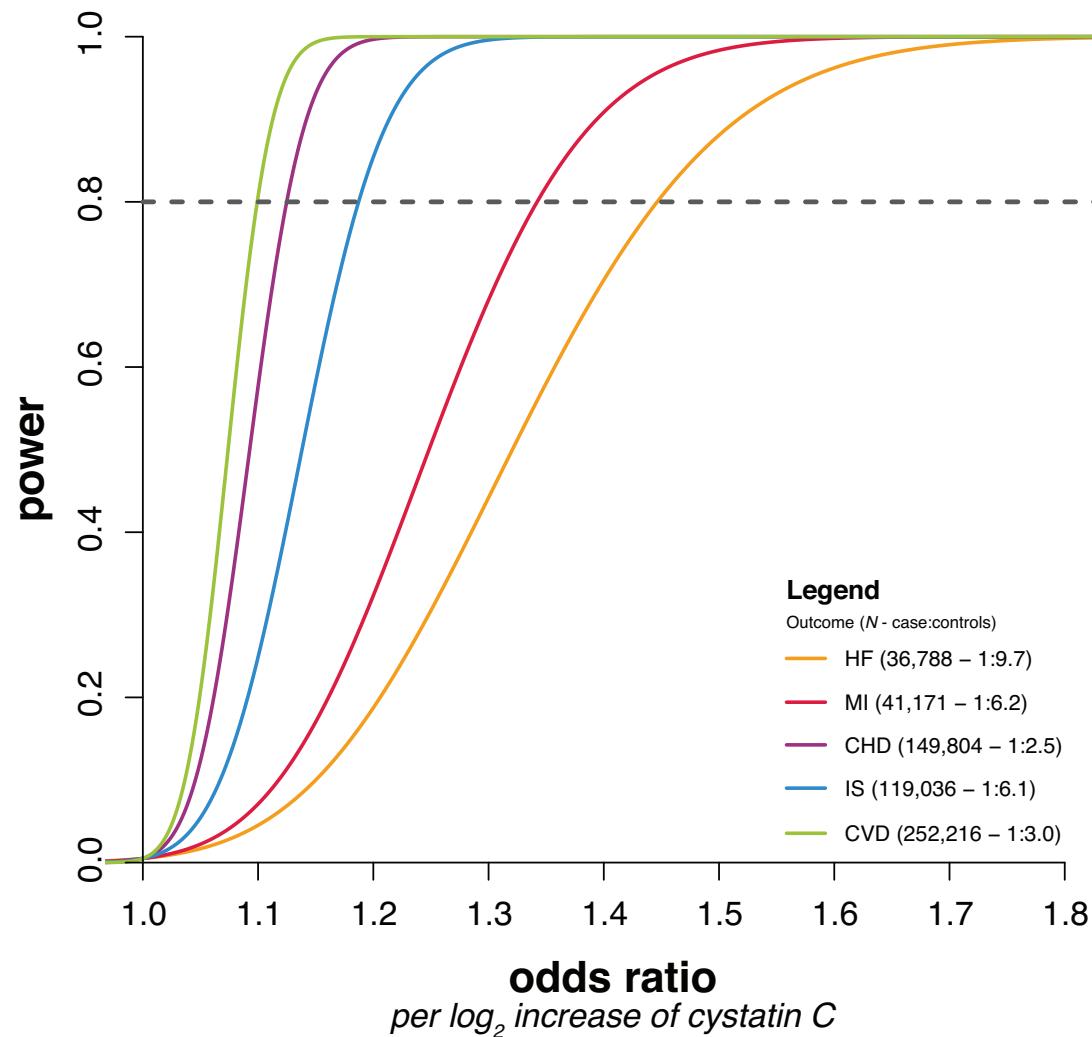
a

$$\begin{array}{c} \text{Power} \downarrow \\ E[X^2] \approx \text{Sample size} * \text{Effect size} * \text{Allele frequency} * \text{LD to causal variant} \\ N * \gamma^2 * p(1-p) * R^2 \end{array}$$

b



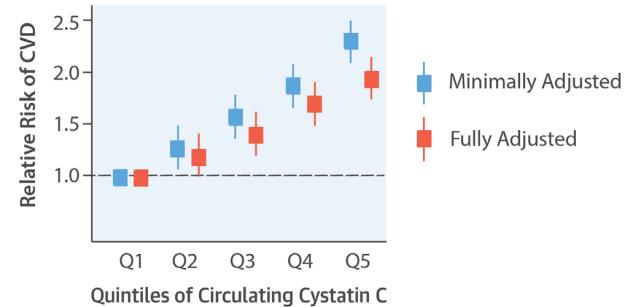
Did we have enough power?



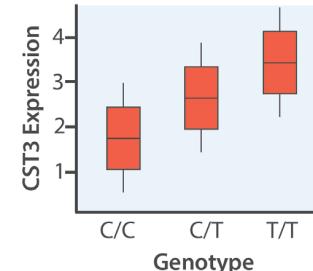
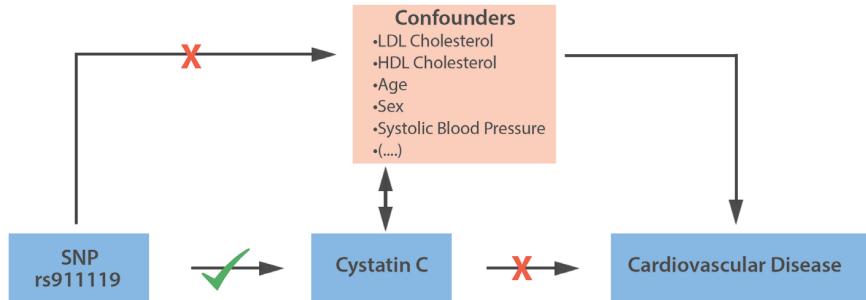
In summary

- There is a correlation between the percentage of SMCs and CystC in plaques, and CystC levels appear to be lower in SMCs from AAAs
- Strong epidemiological evidence exists for a role of CystC in CVD
- Variants in CystC associated with serum CystC levels, eGFR_{Cystc}
- MR analysis: we find no evidence for a causal effect of CystC on CVD

A. Observational Epidemiology



B. Mendelian Randomization



Acknowledgements

JOURNAL OF THE AMERICAN COLLEGE OF CARDIOLOGY

© 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/>).

VOL. 68, NO. 9, 2016

ISSN 0735-1097

<http://dx.doi.org/10.1016/j.jacc.2016.05.092>

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 Johannes Arpegård, MD,^{l,m} Philippe Amouyel, MD, PhD,^{n,o} Stella Trompet, PhD,^{p,q}
Melanie Waldenberger, PhD, MPH,^{g,r} Marcus Dörr, MD,^{e,s} Patrik K. Magnusson, PhD,^t Vilmantas Giedraitis, PhD,^u
Anders Larsson, MD, PhD,^v Andrew P. Morris, PhD,^{w,x} Janine F. Felix, PhD,^f Alanna C. Morrison, PhD,^y
Nora Franceschini, MD, MPH,^z Joshua C. Bis, PhD,^{aa} Maryam Kavousi, MD, PhD,^f
Christopher O'Donnell, MD, MPH,^{bb,cc} Fotios Drenos, PhD,^{dd,ee} Vinicius Tragante, PhD,^{ff} Patricia B. Munroe, PhD,^{gg}
Rainer Malik, PhD,^{hh} Martin Dichgans, MD, PhD,^{hh,ii} Bradford B. Worrall, MD, PhD,^{jj} Jeanette Erdmann, PhD,^{kk}
Christopher P. Nelson, PhD,^{ll,mm} Nilesh J. Samani, PhD,^{ll,mm} Heribert Schunkert, MD, PhD,^{nn,oo}
Jonathan Marchini, PhD,^{pp} Riyaz S. Patel, MD, PhD,^{qq,rr,ss} Aroon D. Hingorani, MD, PhD,^{qq} Lars Lind, MD, PhD,^v
Nancy L. Pedersen, PhD,^t Jacqueline de Graaf, MD, PhD,^{k,tt} Lambertus A.L.M. Kiemeney, PhD,^k
Sebastian E. Baumeister, PhD,^{d,uu} Oscar H. Franco, MD, PhD,^f Albert Hofman, MD, PhD,^f André G. Uitterlinden, PhD,^{vv}
Wolfgang Koenig, MD, PhD,^{e,nn,ww} Christa Meisinger, MD, MPH,^g Annette Peters, PhD, MSc,^{e,g}
Barbara Thorand, PhD, MPH,^g J. Wouter Jukema, MD, PhD,^{p,xx} Bjørn Odvar Eriksen, MD, PhD,^{yy,zz}
Ingrid Toft, MD, PhD,^{zz,†} Tom Wilsgaard, PhD,^{aaa} N. Charlotte Onland-Moret, PhD,^{bbb}
Yvonne T. van der Schouw, PhD,^{bbb} Stéphanie Debette, MD, PhD,^c Meena Kumari, PhD,^{ccc} Per Svensson, MD, PhD,^{l,m}
Pim van der Harst, MD, PhD,^{xx,ddd,eee} Mika Kivimaki, MD, MA,^{fff} Brendan J. Keating, PhD,^{ggg}
Naveed Sattar, MD, PhD,^{hhh} Abbas Dehghan, MD, PhD,^f Alex P. Reiner, MD, MSc,ⁱⁱⁱ Erik Ingelsson, MD, PhD,^{jjj,kkk}
Hester M. den Ruijter, PhD,^a Paul I.W. de Bakker, PhD,^{bbb,lli} Gerard Pasterkamp, MD, PhD,^{a,mmm} Johan Ärnlöv, PhD,^{b,*}
Michael V. Holmes, MD, PhD,^{nnn,*} Folkert W. Asselbergs, MD, PhD,^{ff,xx,ooo,*}



Acknowledgements

Financial support

ICIN (09.001), GENIUS (CVON2011-19), FP7 EU project CVgenes@target HEALTH-F2-2013-601456, BBMRI CP2013-63, Cavadis B.V.



UMC Utrecht



Disclosure: this work is partly financed by Cavadis

