HAROKOPIO UNIVERSITY

SCHOOL OF HEALTH SCIENCE AND EDUCATION

DEPARTMENT OF NUTRITION AND DIETETICS

DIET AND GENETIC SUSCEPTIBILITY IN PHENOTYPES RELATED TO CORONARY ARTERY DISEASE

PhD Thesis

MARIA DIMITRIOU

ATHENS, 2016

DIET AND GENETIC SUSCEPTIBILITY IN PHENOTYPES RELATED TO CORONARY ARTERY DISEASE

MARIA (MELINA) DIMITRIOU

HAROKOPIO UNIVERSITY, SCHOOL OF HEALTH SCIENCE AND EDUCATION,
DEPARTMENT OF NUTRITION AND DIETETICS

DIET AND GENETIC SUSCEPTIBILITY IN PHENOTYPES RELATED TO CORONARY ARTERY DISEASE

MARIA (MELINA) DIMITRIOU

HAROKOPIO UNIVERSITY, SCHOOL OF HEALTH SCIENCE AND EDUCATION
DEPARTMENT OF NUTRITION AND DIETETICS

ΔΙΑΤΡΟΦΗ ΚΑΙ ΓΕΝΕΤΙΚΗ ΠΡΟΔΙΑΘΕΣΗ ΣΕ ΦΑΙΝΟΤΥΠΟΥΣ ΠΟΥ ΣΧΕΤΙΖΟΝΤΑΙ ΜΕ ΤΗ ΣΤΕΦΑΝΙΑΙΑ ΝΟΣΟ

ΜΑΡΙΑ (ΜΕΛΙΝΑ) ΔΗΜΗΤΡΙΟΥ

ΧΑΡΟΚΟΠΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ, ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ ΥΓΕΙΑΣ ΚΑΙ ΑΓΩΓΗΣ

ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΔΙΑΙΤΟΛΟΓΙΑΣ-ΔΙΑΤΡΟΦΗΣ

To my father

Στον πατέρα μου

SUPERVISOR

GEORGE V. DEDOUSSIS, PhD

Professor in Human Cellular and Molecular Biology, School of Health Science and Education,

Department of Nutrition and Dietetics,

Harokopio University of Athens

SUPERVISORY BOARD

GEORGE V. DEDOUSSIS, PhD

Professor in Human Cellular and Molecular Biology, School of Health Science and Education,

Department of Nutrition and Dietetics,

Harokopio University of Athens

DEMOSTHENES B. PANAGIOTAKOS, PhD

Professor in Biostatistics- Epidemiology of Nutrition

School of Health Science and Education, Department of Nutrition and Dietetics, Harokopio University of Athens

LOUKIANOS S. RALLIDIS, MD, PhD

Associate Professor of Cardiology,

School of Medicine, University of Athens

Second Department of Cardiology,

ATTIKON University Hospital

NIKOS YIANNAKOURIS, PhD

Associate Professor in Biology and Physiology,

School of Health Science and Education, Department of Nutrition and Dietetics, Harokopio University of Athens

IOANNIS MANIOS, PhD

Associate Professor in Nutrition Education and Assessment,

School of Health Science and Education, Department of Nutrition and Dietetics, Harokopio University of Athens

ANDRIANA KALIORA, PhD

Assistant Professor in Human Nutrition and Food,

School of Health Science and Education, Department of Nutrition and Dietetics, Harokopio University of Athens

GEORGE PAPANIKOLAOU, MD, PhD

Lecturer in Pathophysiology and Human Genetics,

School of Health Science and Education, Department of Nutrition and Dietetics, Harokopio University of Athens

I Maria Dimitriou declare that:

a) I am the copyright owner of this thesis and to my knowledge, I do not defame others, nor infringe the copyright of third parties.

b) I accept that the Library and Informational Center is granted- through the deposition form - with the right to publish the thesis and make it available through the Internet, via the institutional repository.

c) The thesis is written according to the guidelines from the Harokopio University of Athens, School of Health Science and Education, Department of Nutrition and Dietetics.

Acknowledgments

This study was originally prepared and developed by Dr. Eirini Theodoraki, under the direction and supervision of Professor George V. Dedoussis. This work was partially supported by a research grant (PENED 2003) from the Greek General Secretary of Research and Technology.

Special acknowledgment is made to Professor George V. Dedoussis for his guidance and consultation throughout the implementation of the doctoral thesis. I also express deep appreciation to Professor George V. Dedoussis, Professor Demosthenes B. Panagiotakos and Associate Professor Loukianos S. Rallidis for comments that greatly improved the manuscripts, derived from this work. I would also like to thank the rest of the supervisory board members: Associate Professor Nikos Yiannakouris, Associate Professor Ioannis Manios, Assistant Professor Andriana Kaliora and Lecturer Professor George Papanikolaou. My sincere gratitude also goes to Ioanna-Panagiota Kalafati, PhD Candidate, for her friendship, along with her advice and technical assistance with genetic toolsets. I am also thankful to Vassiliki Eirini Kariakli, Charalampia Amerikanou and Evangelia Karaglani for their contribution during the first years of this project. The present laboratory group members have contributed immensely to my time at Harokopio University and have been the source of friendship and good collaboration. Therefore, I sincerely thank Aliki-Eleni Farmaki, Efthimia Katsareli, Ioanna-Panagiota Kalafati and Charalampia Amerikanou for working with or alongside during these years.

My sincere thanks to the research team at Wellcome Trust Sanger Institute, Hinxton, UK for genotyping the DNA samples in the laboratory facilities. I also thank all the dietitians and clinicians for their contribution to this project.

Last but not least, I would like to extend my deepest gratitude to my family: my mother Georgia Dimitriou and my brother Andreas Dimitriou. Without their encouragement, trust, love and support, it would not be possible to conduct this research and reach academic credentials this far.

Finally, I thank every reader for reading this PhD thesis.

Thank you.

Abbreviations and acronyms

ACE	Angiotensin-converting enzyme
ADA	American Diabetes Association
ADDITION-Europe	AngloDanish-Dutch Study of Intensive Treatment In People with Screen Detected Diabetes in Primary Care
AHA	American Heart Association
APOE	Apolipoprotein E
ARIC	Atherosclerosis Risk in Communities Study
BMI	Body mass index
BP	Blood pressure
CAC	Coronary artery calcium
CAD	Coronary artery disease
CARDIoGRAM	Coronary Artery Disease Genome-wide Replication and Meta-analysis
CCLS	Cooper Center Longitudinal Study
CHA	Chicago Heart Association
CI	Confidence interval
CVD	Cardiovascular disease
CHD	Coronary heart disease
DASH	Dietary Approaches to Stop Hypertension
DBP	Diastolic blood pressure
DM	Diabetes mellitus
DP	Dietary pattern
FA	Factor Analysis
FG	Fasting glucose
FH	Family History
FHS	Framingham Heart Study
FFQ	Food-frequency questionnaire
FI	Fasting insulin
GCTA	Genome-wide complex trait analysis
GRS	Genetic risk score
GPS	Glucose preventive score
GWAs	Genome wide association study(studies)
HC	Hip circumference
HDL-C	High density lipoprotein
HEI	Healthy eating index
HOMA-IR	Homeostasis model assessment
HR	Hazard ratio
HWE	Hardy-Weinberg equilibrium
IHD	Ischemic heart disease
LD	Linkage disequilibrium
LDL-C	Low density lipoprotein
LIPA	Lipase
MAF	Minor allele frequency
MD	Mediterranean diet
MedDietScore	Mediterranean diet score
MESA	Multi-Ethnic Study of Atherosclerosis
MET	Metabolic equivalent
MI	Myocardial infarction
MONICA	Monitoring Trends and Determinants in Cardiovascular Disease
MUFAs	Monounsaturated fatty acids
NCEP	National Cholesterol Education Program

NHANES	National Health and Nutrition Examination Surveys
PA	Physical activity
PAL	Physical activity level
PCA	Principal component analysis
PREDIMED	PREvención con DIeta MEDiterránea study
PROCAM	Prospective Cardiovascular Münster
PUFAs	Polyunsaturated fatty acids
QC	Quality control
RCT	Randomized controlled trial
RH	Relative hazard
RR	Relative risk
OHGS	Ottawa Heart Genomics Study
OR	Odds ratio
SD	Standard devation
SCS	Seven Countries Study
SFAs	Saturated fatty acids
SNPs	Singe nucleotide polymorphisms
SPB	Systolic blood pressure
THISEAS	The Hellenic study of Interactions between Single nucleotide polymorphisms and Eating in Atherosclerosis Susceptibility
TGs	Triglycerides
TC	Total cholesterol
T1DM	Type 1 diabetes mellitus
T2DM	Type 2 diabetes mellitus
OR	Odds ratio
PCA	Principal component analysis
PUFA	Polyunsaturated fatty acids
SHS	Second-hand smoke
SSB	Sugar-sweetened beverages
US	United States
WC	Waist circumference
WHO	World Health Organization
WHR	Waist-to-hip ratio
WTCCC	Wellcome Trust Consortium

Gene Index

Gene ID	Official full name / Description
ABCG5	ATP binding cassette subfamily G member 5
ABCG8	ATP binding cassette subfamily G member 8
ABHD2	Abhydrolase domain containing 2
ABO	ABO blood group (transferase A, alpha 1-3-N-acetylgalactosaminyltransferase; transferase B, alpha 1-3-galactosyltransferase)
ADAMTS7	ADAM metallopeptidase with thrombospondin type 1 motif 7
ADCY5	Adenylate cyclase 5
ADORA2A	Adenosine A2a receptor
ADRA2A	Adrenoceptor alpha 2A
ANKS1A	Ankyrin repeat and sterile alpha motif domain containing 1A
APOA1	Apolipoprotein A1
APOA5	Apolipoprotein A5
APOB	Apolipoprotein B
APOE	Apolipoprotein E
APOC1	Apolipoprotein C1
BCAS3	BCAS3, microtubule associated cell migration factor
CDKAL1	CDK5 regulatory subunit associated protein 1 like 1
CDKN2A	Cyclin dependent kinase inhibitor 2A
CDKN2BAS1	CDKN2B antisense RNA 1
CENTD2	This name gene is an alias for ARAP1 (ArfGAP with RhoGAP domain, ankyrin repeat and PH domain 1)
CNNM2	Cyclin and CBS domain divalent metal cation transport mediator 2
COL4A1	Collagen type IV alpha 1 chain
COL4A2	Collagen type IV alpha 2 chain
CRY2	Cryptochrome circadian clock 2
CTNNAL1	Catenin alpha like 1
CXCL12	C-X-C motif chemokine ligand 12
CYP17A1	Cytochrome P450 family 17 subfamily A member 1
C14orf68	This name gene is an alias for SLC25A47 (solute carrier family 25 member 47)
GCKR	Glucokinase regulator
DGKB	Diacylglycerol kinase beta
EDNRA	Endothelin receptor type A
FADS1	Fatty acid desaturase 1
FAM148B	This name gene is an alias for C2CD4B (C2 calcium dependent domain containing 4B)
FES	FES proto-oncogene, tyrosine kinase
FLT1	Fms (proto-oncogene) related tyrosine kinase 1
FOXA2	Forkhead box A2
FURIN	Furin, paired basic amino acid cleaving enzyme

GCKR	Glucokinase regulator
GGCX	Gamma-glutamyl carboxylase
GIPR	Gastric inhibitory polypeptide receptor
GLIS3	GLIS family zinc finger 3
GLS2	Glutaminase 2
GRB10	Growth factor receptor bound protein 10
GUCY1A3	Guanylate cyclase 1 soluble subunit alpha
G6PC2	Glucose-6-phosphatase catalytic subunit 2
HDAC9	Histone deacetylase 9
HHIPL1	HHIP like 1
IGF1	Insulin like growth factor 1
IGF2BP2	Insulin like growth factor 2 mRNA binding protein 2
IL6R	Interleukin 6 receptor
KCNE2	Potassium voltage-gated channel subfamily E regulatory subunit 2
KCNK5	Potassium two pore domain channel subfamily K member 5
KIIA1462	KIIA1462
KL	Klotho
LDLR	Low density lipoprotein receptor
LHX3	LIM homeobox 3
LIPA	Lipase A, lysosomal acid type
LPA	Lipoprotein(a)
LPAL2	Lipoprotein(a) like 2, pseudogene
LPIN3	Lipin 3
LPL	Lipoprotein lipase
MADD	MAP kinase activating death domain
MC4R	Melanocortin 4 receptor
MFGE8	Milk fat globule-EGF factor 8 protein
MIA3	MIA family member 3, ER export factor
MRAS	Muscle RAS oncogene homolog
MTRN1B	Melatonin receptor 1B
NOA1	Nitric oxide associated 1
NOS3	Nitric oxide synthase 3
NT5C2	5'-nucleotidase, cytosolic II
PCSK1	Proprotein convertase subtilisin/kexin type 1
PCSK9	Proprotein convertase subtilisin / kexin type 9
PDGFD	Platelet derived growth factor D
PDX1	Pancreatic and duodenal homeobox 1
PEMT	Phosphatidylethanolamine N-methyltransferase
PHACTR1	Phosphatase and actin regulator 1
PMAIP1	Phorbol-12-myristate-13-acetate-induced protein 1
PLG	Plasminogen
POM121L9P	POM121 transmembrane nucleoporin like 9, pseudogene
PPAP2B	Phosphatidic acid phosphatase type 2B
PPP1R3B	Protein phosphatase 1 regulatory subunit 3B
PRKAR2A	Protein kinase cAMP-dependent type II regulatory subunit alpha
PROX1	Prospero homeobox 1
P2RX2	Purinergic receptor P2X 2

RAI1	Retinoic acid induced 1
RASD1	Ras related dexamethasone induced 1
REST	RE1 silencing transcription factor
RREB1	Ras responsive element binding protein 1
SH2B3	SH2B adaptor protein 3
SMAD3	SMAD family member 3
SMG6	SMG6, nonsense mediated mRNA decay factor
SLC2A2	Solute carrier family 2 member 2
SLC22A3	Solute carrier family 22 member 3
SLC22A4	Solute carrier family 22 member 4
SLC22A5	Solute carrier family 22 member 5
SLC30A8	Solute carrier family 30 member 8
SORT1	Sortilin 1
SWAP70	SWAP switching B-cell complex 70kDa subunit
TCF7L2	Transcription factor 7 like 2
TCF21	Transcription factor 21
TMEM195	This name gene is an alias for AGMO (alkylglycerol monooxygenase)
TRIB1	Tribbles pseudokinase 1
UBE2Z	Ubiquitin conjugating enzyme E2 Z
VAMP5	Vesicle associated membrane protein 5
VAMP8	Vesicle associated membrane protein 8
WDR12	WD repeat domain 12
ZBED3	Zinc finger BED-type containing 3
ZC3HC1	Zinc finger C3HC-type containing 1
ZEB2	Zinc finger E-box binding homeobox 2
ZNF259	This name gene is an alias for ZPR1 zinc finger. The protein encoded by this gene is found in the cytoplasm of quiescent cells but translocates to the nucleolus in proliferating cells

Contents

Abstract

Introduction: Cardiovascular disease (CVD) is the leading cause of mortality and disability worldwide. Coronary artery disease (CAD) accounts for the most common type of CVDs. Epidemiological studies have provided new insights into CAD related risk factors and have contributed to the evolution of the field of preventive cardiology. Genetic susceptibility and diet play a key role in CAD risk.

Aims and Objectives: The objectives of the current PhD thesis were a) to record a wide range of markers between individuals with established CAD and individuals free of the disease, b) to identify novel CAD-associated loci, c) to evaluate the effect of a genetic risk score (GRS) on CAD risk d) to identify dietary patterns (DPs) associated with CAD risk and e) to examine the joint effect of genetic predisposition and diet/lifestyle on CAD risk and on intermediate phenotypes.

Methods: This is a medical center-based case-control study conducted in the region of Athens. Demographic, anthropometric, biochemical and lifestyle variables were recorded. Genetic analysis was performed using micro-arrays and a GRS was constructed based on a panel of 53 CAD associated loci (wGRS-53). Also, a glucose preventive score (GPS) and a GRS based on 20 glucose raising loci were modeled (wGRS). DPs were identified and associated with the likelihood of having CAD. For statistical analysis, IBM SPSS Statistics 13.0 and 21.0, PLINK v.1.07 and SNPTEST v.2.5.2 were used.

Results: Exclusive olive oil consumption was associated with 37% lower odds of CAD, after taking into account the Mediterranean diet adherence (OR=0.63; 95% CI 0·42, 0·93; P=0·02). The GPS was associated with lower glucose concentrations (b ±SE:-0.083 ± 0.021mmol/L, $p=1.6\times10^{-04}$). The association of the wGRS with glucose levels was attenuated, after interaction with the GPS. The wGRS-53 was associated with increased CAD risk (OR=1.03, 95% CI=1.01-1.05, *p*=0.11), after confounding. Regarding DPs, the Western type diet pattern was positively associated with CAD risk (OR=1.20; 95% CI=1.09-1.32, *p*=0.000). The study also participated in two Consortia, supporting important outcomes with data from the present Greek sample.

Discussion: Exclusive olive oil consumption was associated with lower CAD risk, after adjusting for adoption of an overall healthy DP such as the Mediterranean diet. The study of interactions indicated that the effect of adherence to a healthier lifestyle or to an unhealthy DP is not similar among individuals with different genetic backgrounds. In addition, insights

in the genetic basis of CAD were provided. In the near future, GRSs could be useful in clinical practice for identifying individuals at high CAD risk and/or predicting CAD incidence.

Key words: case-control, coronary artery disease, diet, olive oil, single-nucleotide polymorphism

Περίληψη

Εισαγωγή: Τα καρδιαγγειακά νοσήματα (ΚΝ) αποτελούν την πρώτη αιτία θνησιμότητας και θνητότητας σε παγκόσμιο επίπεδο. Η στεφανιαία νόσος (ΣΝ) αποτελεί την πιο συχνή κλινική εκδήλωση των ΚΝ. Οι επιδημιολογικές μελέτες βοήθησαν στην κατανόηση των παραγόντων κινδύνου της ΣΝ κι έχουν συνεισφέρει στην εξέλιξη της πρόληψης στον τομέα της καρδιολογίας. Η γενετική προδιάθεση και η διατροφή διαδραματίζουν σημαντικό ρόλο στον κίνδυνο εμφάνισης ΣΝ.

Σκοποί: Σκοποί της παρούσας διατριβής αποτελούν α) η καταγραφή διαφόρων δεικτών υγείας μεταξύ ασθενών με διεγνωσμένη ΣΝ και μαρτύρων, β) η αναγνώριση νέων γενετικών τόπων, γ) η αξιολόγηση ενός γενετικού σκορ (ΓΣ), δ) η αναγνώριση διατροφικών προτύπων (ΔΠ) και ε) η αξιολόγηση της επίδρασης της γενετικής προδιάθεσης σε συνδυασμό με τη διατροφή ή τον τρόπο ζωής, ως προς την πιθανότητα παρουσίας ΣΝ και ως προς ενδιάμεσους φαινότυπους.

Μεθοδολογία: Η παρούσα διατριβή αποτελεί μελέτη ασθενών μαρτύρων και πραγματοποιήθηκε σε ελληνικό δείγμα στην περιοχή της Αθήνας. Στο πλαίσιο της μελέτης πραγματοποιήθηκαν μετρήσεις για τη συλλογή δημογραφικών, ανθρωπομετρικών, βιοχημικών και συμπεριφοριστικών δεδομένων. Έπειτα από την απομόνωση γενετικού υλικού, πραγματοποιήθηκε γενετική ανάλυση με τη χρήση μικροσυστοιχιών. Κατασκευάστηκε ένα ΓΣ βασισμένο σε 53 πολυμορφισμούς (ΓΣ-53) που έχουν συσχετισθεί με τη ΣΝ. Επίσης, δημιουργήθηκαν ένα ΓΣ βασισμένο σε 20 πολυμορφισμούς (ΓΣ-20) που έχουν συσχετιστεί με τα επίπεδα γλυκόζης αίματος κι ένα συμπεριφοριστικό σκορ (ΣΣ) για την πρόληψη αυξημένων επιπέδων γλυκόζης. Επιπρόσθετα, αναγνωρίστηκαν ΔΠ και ελέγχθηκε η συσχέτισή τους ως προς την πιθανότητα παρουσίας ΣΝ. Για τη στατιστική ανάλυση των δεδομένων χρησιμοποιήθηκαν τα στατιστικά πακέτα, IBM SPSS Statistics 13.0 και 21.0, PLINK v.1.07 και SNPTEST v.2.5.2.

Αποτελέσματα: Η αποκλειστική κατανάλωση ελαιολάδου σχετίστηκε με 37% μείωση του σχετικού κίνδυνου ΣΝ, ύστερα από στάθμιση για την προσκόλληση στη μεσογειακή διατροφή (OR=0·63; 95% CI=0.42-0.93; P=0.02). Το ΣΣ συσχετίστηκε με χαμηλότερα επίπεδα γλυκόζης αίματος (b ±SE: -0.083±0.021mmol/L, $p=1.6\times10^{-04}$). Η επίδραση του ΓΣ-20 στα επίπεδα γλυκόζης εξασθένισε, μετά από αλληλεπίδραση με το ΣΣ. Το ΓΣ-53 συσχετίστηκε με αυξημένο κίνδυνο για ΣΝ (OR=1.03, 95% CI=1.01-1.05, p=0.021),

σταθμισμένο για ηλικία και φύλο. Αναφορικά με τα ΔΠ, το Δυτικό τύπου πρότυπο συσχετίστηκε θετικά με την πιθανότητα παρουσίας ΣΝ (OR=1.20; 95% CI=1.09-1.32, *p*=0.000). Η μελέτη συμμετείχε σε δυο πολυκεντρικές μελέτες συνεισφέροντας στα ευρήματα αυτών με δεδομένα από το συγκεκριμένο ελληνικό δείγμα.

Συμπεράσματα: Η αποκλειστική χρήση ελαιολάδου φαίνεται να δρα προστατευτικά ως προς την εκδήλωση ΣΝ. Η μελέτη των αλληλεπιδράσεων υποδεικνύει ότι η επίδραση ενός υγιεινού τρόπου ζωής ή ενός ανθυγιεινού ΔΠ είναι διαφορετική μεταξύ ατόμων με διαφορετική γενετική προδιάθεση. Τα αποτελέσματα παρέχουν περαιτέρω πληροφορίες για την γενετική βάση της ΣΝ. Τα ΓΣ μπορούν στο εγγύς μέλλον να παρέχουν κλινική χρησιμότητα για την αναγνώριση ατόμων σε υψηλό κίνδυνο εμφάνισης ΣΝ ή/και να έχουν προγνωστική αξία για την εμφάνιση ΣΝ.

Λέξεις-Κλειδιά: ασθενείς-μάρτυρες, στεφανιαία νόσος, διατροφή, ελαιόλαδο, πολυμορφισμός

Doctoral Dissertation Outlook

The present doctoral dissertation is organized in 6 chapters followed by 7 included reports. The thesis is based on **T**he **H**ellenic study of **I**nteractions between **S**ingle nucleotide polymorphisms and **E**ating in **A**therosclerosis **S**usceptibility **(THISEAS)** cohort, which is a medical-based case-control study conducted in a Greek sample. Chapter 1 provides an introduction (literature review) regarding coronary artery disease. In the same chapter and after the introduction, the aims are stated. The design and methods used during this thesis work are presented in Chapter 2. Chapter 3 presents four published articles in peer reviewed journals (results from the THISEAS study and results from two meta-analyses generated from Consortia participation) and two unpublished reports. Chapter 4 includes the conclusions and future perspectives. The list of references and supplementary material are presented in chapters 5 and 6, respectively. Finally, the appendix in chapter 7 presents a review of the literature paper relevant to the theme framework of the PhD thesis.

Chapter 1

INTRODUCTION

1 INTRODUCTION

1.1. Cardiovascular Disease

The cardiovascular system consists of the heart, blood and vessels (Mendis, Puska and Norrving, 2011). It is responsible for the transportation of gases, nutrients, hormones and blood cells to all parts of the body. In addition, the cardiovascular system regulates the body temperature and acid-base balance. This is a rather broad description of the cardiovascular system, since the cardiovascular mechanics is chaotic, complex and not completely understood. The presence of multiple regulatory mechanisms, which are contemporaneously active and interact with many cardiovascular variables, play a key role to the complexity of the cardiovascular network (Porta et al., 2009).

Cardiovascular disease (CVD) includes disorders of the heart and blood vessels and vascular diseases of the brain. CVD due to atherosclerosis include ischaemic heart disease (IHD) and coronary artery disease (CAD) (heart attack), cerebrovascular disease (stroke), disease of the aorta and arteries, including hypertension and peripheral artery disease. Other CVDs, include congenital heart disease (heart structure malformations), rheumatic heart disease, cardiomyopathies and cardiac arrhythmias (Mendis, Puska and Norrving, 2011).

CVD is the leading cause of mortality and disability worldwide, with more than 80% of deaths taking place in low-income and middle-income countries. In 2011, World Health Organization (WHO) (WHO, 2011) estimated that more 17 million people died from CVD in the year 2008, thus accounting for 30% of all global deaths. By the year 2030, it is estimated that CVD will account for 23.6 million deaths worldwide (Mendis, Puska and Norrving, 2011). Figures 1.1 and 1.2 illustrate the mortality rates of CVD in high-income and low-income countries, in the year 2008.

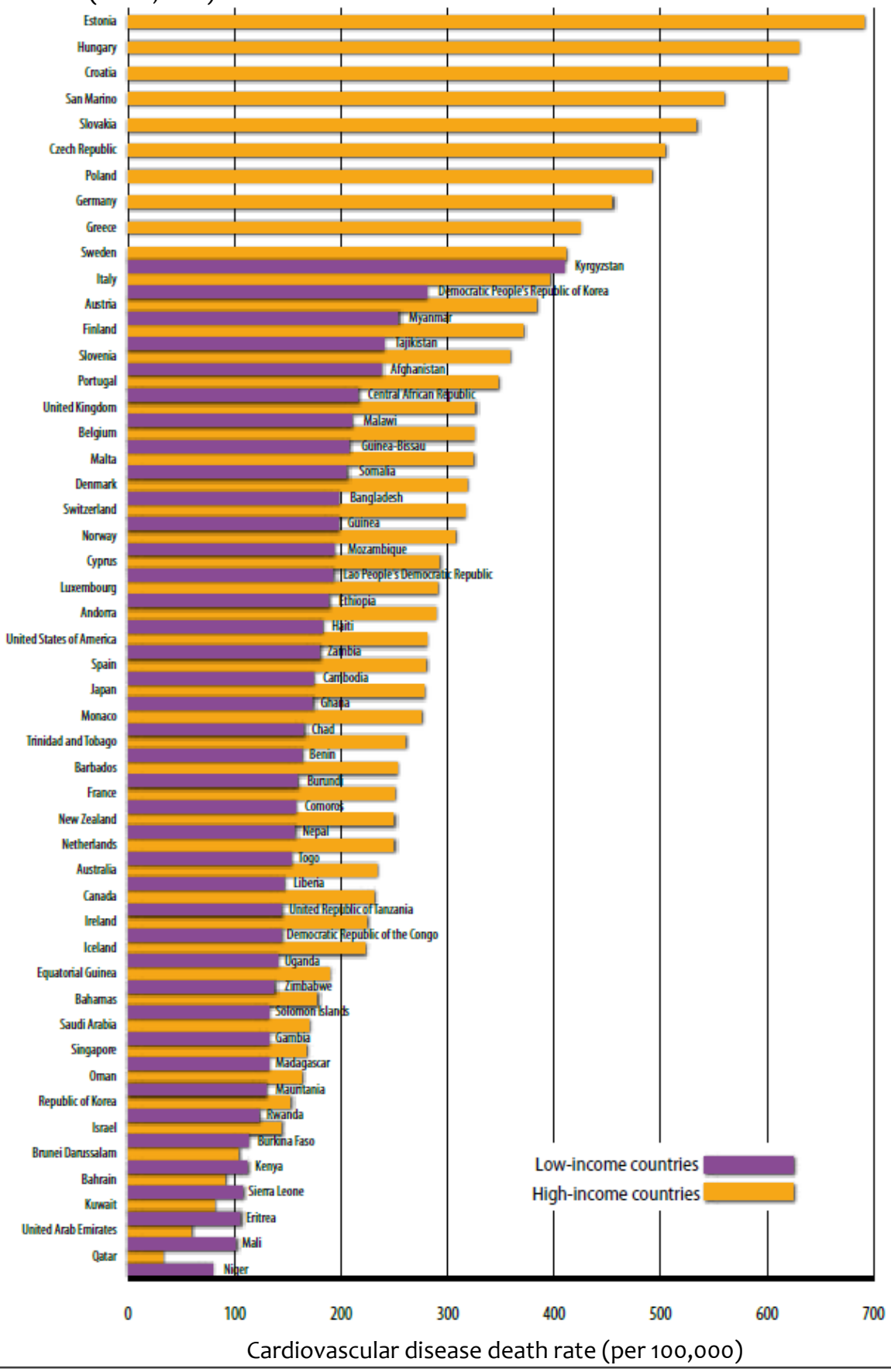
According to the latest report of the American Heart Association (AHA), death rates from CVD have declined 30.8%, from 2001 to 2011. However in 2011, 1 CVD death occurs in every 40 seconds (Mozaffarian, 2016). During the same year, the direct and indirect economic burden of CVD is enormous and accounts for 15% of total health expenditure. In absolute numbers, 320.1 billion dollars is the estimated cost of CVD and stroke in the United States (US) (Medical Expenditure Panel Survey, 2011; Shanthi et al, 2009).

The public health burden of CVD and stoke challenged the AHA to implement goals for CVD reduction and health promotion. Therefore, the AHA impact goals are stated in the following sentence.

"By 2020, to improve the cardiovascular health of all Americans by 20% while reducing deaths from cardiovascular diseases and stroke by 20%." (Lloyd-Jones et al., 2010)

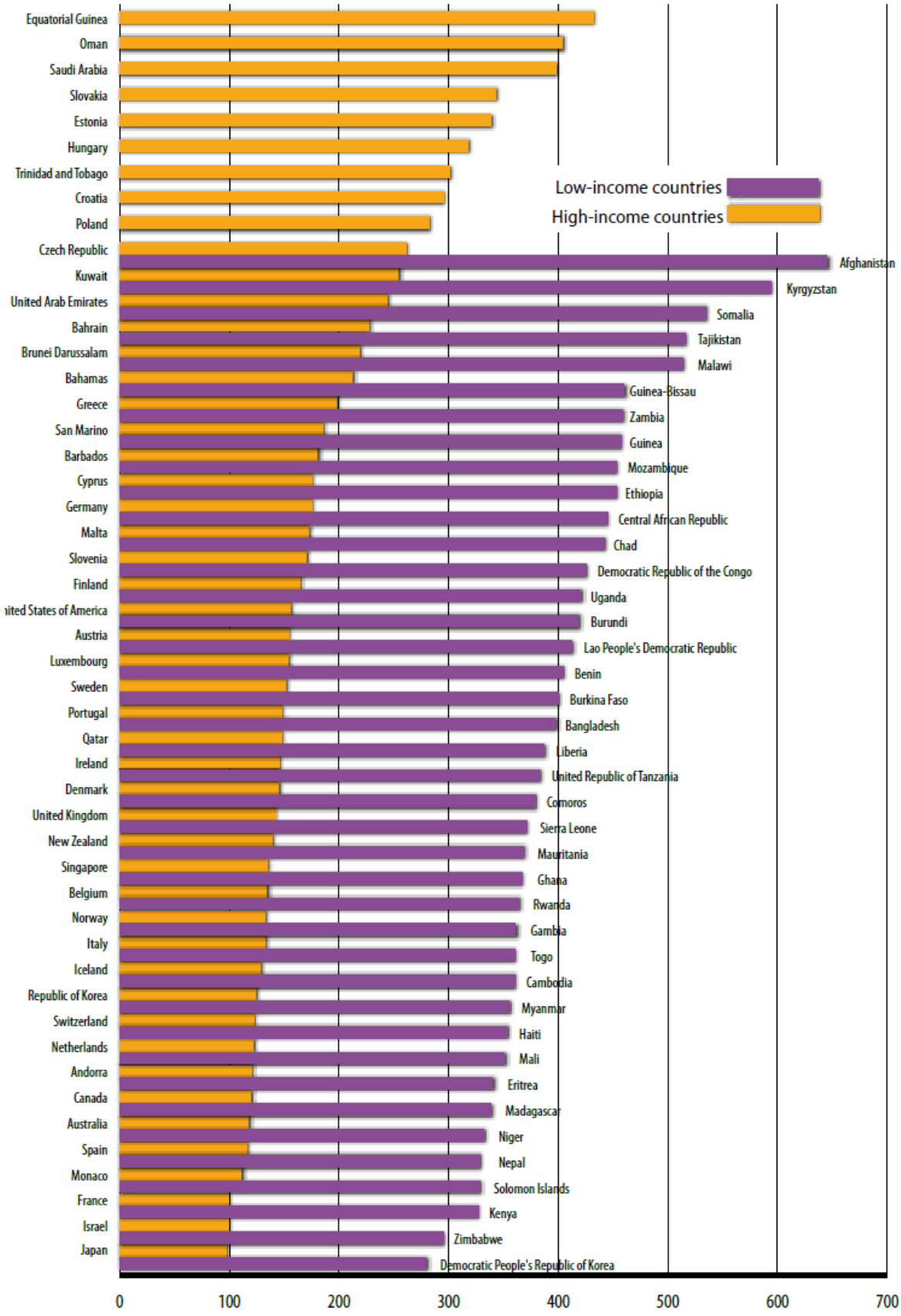
It is not to the intent of this doctoral thesis to include the epidemiology and pathophysiology of the CVD continuum; instead this part of the thesis will present a comprehensive update of the current evidence of CAD.

Figure 1.1 | Mortality rates of CVDs in high-income and low-income countries, 2008 (WHO, 2011)



Source: Global Atlas on cardiovascular disease prevention and control, World Health Organization in collaboration with the World Heart Federation and the World Stroke Organization, 2011, p.112

Figure 1.2 | Mortality rates of CVDs in high-income and low-income countries, 2008 (age-adjusted) (WHO, 2011)



Source: Global Atlas on cardiovascular disease prevention and control, World Health Organization in collaboration with the World Heart Federation and the World Stroke Organization, 2011, p.111

1.2. Coronary Artery Disease Definition and Epidemiology

1.2.1. Epidemiological Studies of Coronary Artery Disease globally

CAD accounts for the most common type of CVDs. The term is often interchangeable with coronary heart disease (CHD), atherosclerotic heart disease or ischemic heart disease. Clinical manifestations of CAD include stable and unstable angina, myocardial infarction (MI) and coronary death (Wong, 2014). The underlying pathology of CAD is atherosclerosis.

It is estimated that 17.3 million people died from CVD in 2008 and this number contributes to 30% of all deaths worldwide. Of these deaths, 7.3 million were due to CAD (WHO, 2011). This number is expected to rise up to 23.6 million by 2030.

Epidemiological studies have provided new insights into CAD related risk factors and have contributed to the evolution of the field of preventive cardiology. The first epidemiological study was conducted in Minnesota in 1946 and recruited 281 men under the age of 55. This study paved the way for the **Seven Countries Study (SCS)** to come in 1958. The SCS established the contribution of diet and lifestyle as important risk factors to CAD risk (Keys et al, 1963; Keys et al. 1984).

The **Framingham Heart Study (FHS)** was initiated in 1947 and the study design was first described in 1951. The FHS is a hallmark study in the identification of many risk factors that are associated with increased CAD risk, namely age, sex, blood cholesterol, smoking, hypertension and diabetes. In addition, the FHS highlighted the contribution of the genetic susceptibility to CAD risk providing new insights into genetic epidemiology. Importantly, the fact that the co-existence of risk factors have an accumulative effect on CAD risk assessment was first introduced by this study. Therefore, the FHS was a pioneer in establishing that individuals with multiple risk factors were in greater CAD risk than those with a single risk factor (Kannel et al., 1961). One important contribution of the FHS was the development of a coronary prediction model, known as the Framingham score (Wilsion et al., 1998).

The **National Health and Nutrition Examination Survey (NHANES)** begun in the 1960s. In 1990, the survey became a series of programs focused on assessing the overall health along with the nutritional status of adults and children in the US. NHANES findings are used to determine the prevalence of major diseases , including CVD. Also, risk factors of a disease are examined. Such information is often used by epidemiological studies and science research.

One of the most important study conducted globally is the **Monitoring Trends and Determinants in Cardiovascular Disease (MONICA)** project. The MONICA project was developed by WHO in 1979 and was built on the FHS, the SCS, the Cardiovascular Survey Methods and European Myocardial Infarction Registers. The project was focused on a 10-year period monitoring CAD mortality and morbidity and trends of classic risk factors across different populations and within the two genders.

Another important study in CAD epidemiology is the **INTERHEART** study, which is a case-control study of acute MI in 52 countries (Yusuf et al., 2004). The main aim of this cohort was to investigate the effect of risk factors (namely blood lipids, obesity, hypertension, diabetes, diet, physical activity (PA), alcohol consumption and psychosocial factors) in different ethnic groups and geographical regions. The main results of the study showed that the risk factors under study were associated with MI across all ethnic groups, age groups and different regions. The findings of the study were published in Lancet (Rosengren et al., 2004, Yusuf et al., 2004).

National epidemiological studies have provided important knowledge in preventive cardiology. The **Prospective Cardiovascular Münster (PROCAM)** study was initiated in 1979 and was conducted on the German population (Assmann and Schulte, 1998). The main aims of the study were to determine the prevalence of CAD risk factors and to improve the prediction and early detection of CAD. After a 10-year follow-up, the PROCAM study investigators developed a scoring system in an attempt to predict CAD risk in clinical practice (Assmann and Schulte, 2002).

The **Atherosclerosis Risk in Communities** study **(ARIC)** is a prospective epidemiological study, which was initiated in 1987 in four communities in the US. The study was designed to elucidate the pathology of atherosclerosis and atherosclerotic disease and to investigate variation in CVD risk factors. The reports from the ARIC study are important in CVD research.

The **Multi-Ethnic Study of Atherosclerosis (MESA)** enrollment begun in 2000 and people from four ethnic groups (28% African-Americans, 23% Hispanic, 11% Asian and 38% white) were enrolled (Bild et al., 2002). The MESA study was the first study to demonstrate that coronary calcium is a predictor of CAD incidence, beyond the traditional risk factors. The risk prediction of calcium scores was similar among the ethnic groups (Detrano et al., 2008). Other important findings from MESA study include reports on air-pollution, diet, aortic calcification and genetics (Budoff et al., 2011; Park et al., 2010; McGeachie et al., 2009; Nettleton et al., 2008).

The **PREvención con DIeta MEDiterránea study (PREDIMED)** was initiated in Spain at 2003 and was designed to examine the impact of Mediterranean diet (MD) on the primary prevention of CVD. The study was registered as a large randomized trial but it was scheduled to continue as an observational cohort of high-risk participants included in follow-up (Martínez-González et al, 2012). The PREDIMED study demonstrated that the MD could be a useful tool in CAD prevention (Ros et al, 2014) and extra virgin olive oil was associated with reduced CVD risk and mortality in high risk individuals (Guasch-Ferré et al, 2014).

1.2.2. Epidemiological Studies of Coronary Artery Disease in Greece

The first endeavor to illustrate CAD rates in a Greek sample was the SCS. The SCS included Crete and Corfu in Greece, representing one of the four Mediterranean cohorts. Men habitants of rural villages in the two Greek islands were enrolled in the study. The participants' diet of the two Greek islands was characterized by high olive oil and vegetable intake, reflecting an eating pattern, called the MD. The SCS demonstrated that the MD was associated with low CAD incidence and mortality rates. Furthermore, the SCS was the first study that demonstrated the prevalence of CAD risk factors in Greece, highlighting Greece among the cohorts with low CAD prevalence (Menotti et al, 1989). However, the Greek sample was not representative of the whole population in Greece.

After the SCS, two local studies attempted to investigate CVD risk factors in an Athenian sample in Greece (Moulopoulos et al., 1987; Kalandidi et al., 1992). The published findings of these studies, although essential, were not representative of the population. In the year 2000, a multi-centre case-control study was conducted, the **CARDIO2000 study**, in an attempt to fill in the gap in the scientific literature and further evaluate CVD risk factors in the Greek population. The CARDIO2000 study was the first study in Greece to explore the association between gender, demographic, education, diet, lifestyle (eg, smoking habits, tobacco exposure, PA habits, alcohol consumption), stress, depression and biochemical risk factors with the risk of acute coronary syndromes development (Panagiotakos et al., 2001; Chrysohoou et al., 2003). The most important outcome of the CARDIO2000 study demonstrated that hyperlipidemia, hypertension, diabetes mellitus type 2 (DMT2), family history (FH) and sedentary lifestyle increase acute coronary syndromes incidence. On the other hand, adherence to the MD and high educational level play a protective role in the reduction of coronary risk.

The subsequent years, Greece experienced great socio-economic development, resulting to lifestyle changes within the population. A population-based health and nutrition survey, the **ATTICA study**, was initiated at the beginning of 2001. The study was conducted in the province of Athens, including 3042 adults from urban and rural areas (78% and 22%, respectively). The main aims of the ATTICA study were a) to measure the prevalence of CVD risk factors in a representative Greek sample and b) to assess their prognostic significance through follow-up examinations in order to be able to successfully identify individuals at high CAD risk and/or predict CAD incidence (Pitsavos et al., 2003). One of the major findings of the study was the high prevalence of CVD risk factors in the Greek population, placing a remarkable proportion of Greek adults at high risk for developing cardiovascular events in the future. This outcome was established during the 5-year follow up, which revealed a considerable increase in the prevalence of CVD risk factors (independently of gender and age) and an increase in CVD incidence rates. In absolute numbers, approximately 550,000 adults developed a fatal or non-fatal CVD event, which represents 8.5% of the study's participants (Panagiotakos et al. 2009). The burden of high CVD incidence rates and CVD risk factors, was further confirmed after the 10-year follow- up completion (Panagiotakos et al., 2015). The findings of the study, although cannot be generalized to the whole country, possibly suggest a shift of Greece from a country with "low" to a country with "high" CAD incidence and mortality rates. Up to date, the ATTICA study remains one of the most important health and nutrition surveys in CVD conducted in the Greek population (in the province of ATTICA), with many published papers in the scientific literature.

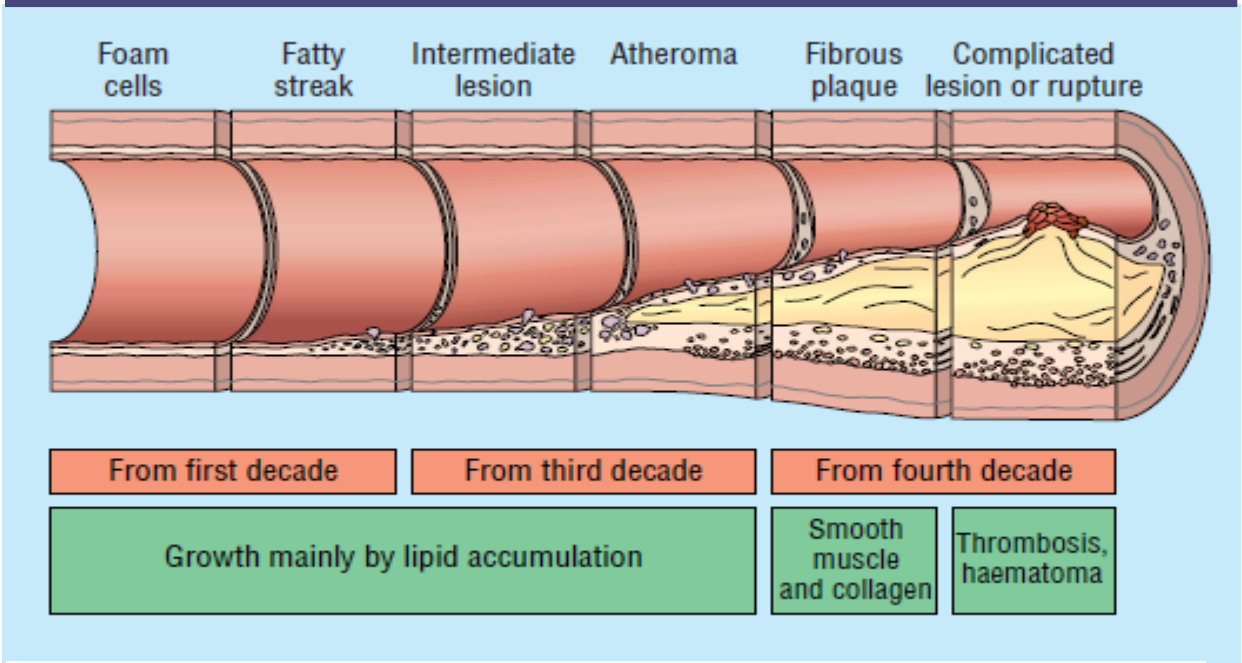
1.3. Pathophysiology and Clinical manifestations of Coronary Artery Disease

During the last decade, our understanding regarding the pathophysiology of CAD has remarkably evolved. The main manifestation of CAD is atherosclerosis. Atherosclerosis is a complex disorder, which can be described as a gradual thickening of the inner layer (intima) of the coronary arteries, causing narrowing of the lumen of the artery to various degrees and subsequent total or near total occlusion. Inflammation and interaction of metabolic risk factors with immune mechanisms are implicated to the initiation and progression of atherosclerotic lesions (atheromata, from the Greek words athera and oma). Atherosclerotic lesions may be present from young adulthood (Wong et al., 2012). Atherosclerosis is accelerated under the presence of bacterial products or traditional risk factors. Specifically, bacterial products, dyslipidemia, products of glycoxidation associated with hyperglycemia, vasoconstrictor hormones implicated in hypertension, proinflammatory cytokines from excess adipose tissue initiate the expression of adhesion molecules, which subsequently promote the migration of blood cells to the inner surface of the arterial wall. In all stages of atherosclerosis, inflammation plays an important role. The progression of atherosclerosis involves the formation of foam cells and fatty streaks to the establishment of atheromatous plaque as described in the following section.

Early lesions (Type I lesions) of the intima are the precursors of advanced atherosclerotic lesions. Early lesions include foam cell formation and fatty streak development. Foam cells represent the initial changes in the intima, where an increased number of intimal macrophages as well as, macrophages filled with lipid droplets appear. Fatty streaks (Type II lesions) are the first visible lesions and are composed of (macrophage) foam cells in layers, smooth muscle cells that now contain lipid droplets and extracellular lipid. T lymphocytes have also been detected. Lymphocytes and macrophages are often combined under the term "mononuclear cells" (Stary et al., 1994). Early atherosclerotic lesions can be present early in life, specifically in newborns and children (Napoli et al., 2007). In early lesions, the intimal architecture either consists of the usual components or is minimally disrupted. Lesion progression leads to the accumulation of the extracellular lipid among the layers of smooth muscle cells and below the layers of foam cells, thus disrupting the intimal smooth muscle cell architecture. These histological changes designate the intermediate lesions (transitional lesions) that precede the atheroma. Its histological feature is the additional extracellular lipid droplets, which are membrane-bound or free. The intermediate lesion are also known

as Type III lesion, the transitional lesion, and as preatheroma. The atheroma is considered to be an advanced lesion type since it predisposes to sudden lesion progression and cardiac events. Atheromatous lesions transform when atherogenic lipoproteins exceed critical levels and mechanical forces enhance lipoprotein deposition in the same regions. Advanced lesions may come along with intimal thickening (Stary et al., 1994). As lesion progresses, the extracellular lipid that accumulates and form the classic, lipid-rich core of the atherosclerotic plaque. Atherosclerotic plaques can be stable or unstable (vulnerable). Stable plaques are asymptomatic and are rich of extracellular matrix and smooth muscle cells. Unstable plaques are prone to rupture and are rich of macrophages, foam cells and extracellular matrix that is prone to rupture. When a plaque is ruptured, thrombotic substances enter the circulation and thrombus progression is induced (Fan and Watanabe, 2003). **Figure 1.3.** illustrates the progression of atheromatous plaque from initial lesion to complex and ruptured plaque. Our understanding of the pathophysiology of CAD and atherogenesis has undergone remarkable evolution during the last decades.

Figure 1.3 Progression of atheromatous plaque



Source: Even D Grech. BMJ 2003; 326:1027-1030.

When a plaque produces more than a 50% diameter reduced blood flow through the coronary may lead to stable angina. Acute coronary events usually arise when thrombus formation follows disruption of a plaque. Intimal injury causes exposure and susceptiveness

of the thrombogenic matrix or lipid pool and triggers thrombus formation. Acute coronary events include unstable angina and MI. In acute MI, occlusion is more complete than in unstable angina, where arterial occlusion is usually subtotal (>90%). Downstream embolism of thrombus may also produce microinfarcts.

Angina is the most common symptom of CAD and approximately, 50% of CAD patients appear with retrosternal chest pain discomfort and dyspnea. Chest pain arises from nerve stimuli near the endocardium. The pain is typically described as heavy, tight pressure on the chest or feels like indigestion. There is great variation between patients regarding pain localization (chest, neck, shoulders, arms, back or jaw). Musculoskeletal and gastro-oesophageal pains may co-exist. Also, patients present with atypical pain, which results to diagnosis difficulty in many cases. Triggering factors include emotion, extreme temperature and exercise. Angina may be a) stable and symptoms relent with rest or nitroglycerin or b) unstable and it is a disturbance of the usual pattern of stable angina. Unstable angina is considered to be an emergency (Dawis, 2001).

Acute MI is due to the formation of an occlusive thrombus triggered by the rupture of an atheromatic plaque. Obstruction of a coronary artery lasting more than 30 minutes in the absence of collateral circulation results in ischemic myocardial necrosis. Chest pain may be present prior to infarction, however infarction may occur without signs or symptoms (silent). Most often, chest discomfort appears for more than 30 minutes and the pain is reflected in the centre or the left side of the chest. The symptoms of MI can be similar to angina and include breath shortness, nausea, cold sweat, fatigue, vomiting and sleep disturbance. In MI, the pain is severe and constant, while in angina the pain goes away with rest. Unstable angina and MI are considered to be acute coronary events (Pepine and Nichols, 2007).

1.4. Risk factors of Coronary Artery Disease

Several risk factors have been recognized and associated with increased risk of CAD development. CAD risk factors were initially demonstrated by the findings of the FHS (Kanel et al., 1961). To date, epidemiologic information has demonstrated that CAD risk factors cluster and interact in a consistent way for CAD promotion. This information allows the clinicians to identify individuals at high risk for CAD development. There are risk factors that

cannot be changed, risk factors that can be modified and others under investigation (Grech ED, 2003).

Non-modifiable risk factors

- Age
- Family history (genetics)
- Gender (male sex)

Modifiable risk factors

- Hyperlipidaemia
- Diabetes Mellitus
- Overweight and obesity
- Diet (atherogenic diet)
- Sedentary lifestyle (inactivity)
- Excessive alcohol intake
- Smoking
- Excessive stress?

Uncertain risk factors

- Hypertriglyceridaemia
- Mircroalbuminuria
- Hyperhomocysteinaemia
- Lp(a)lipoprotein
- Fibrinogen
- C-reactive protein
- Uric acid
- Renin

Non-modifiable and modifiable risk factors will be further discussed in the subsequent sections of the current thesis. Uncertain risk factors are addressed in this thesis.

1.4.1. Non-Modifiable Risk Factors

Age

Increasing age is one of the most significant predictors of CAD. It is estimated that the majority of people who die of CAD are above the age of 65. Statistics from WHO and United Nations, demonstrated that CAD/IHD rates increase by a 2.3 to 2.7-fold for every decade for life in men and by 2.9 to 3.7-fold for women. Data during 2000 and 2009, demonstrated that age-standardized IHD mortality rates are flat (Eastern Europe) or increasing (Central Asia countries) in low and middle income countries. On the other hand, IHD mortality rates have a declining trend in most high income countries (Western Europe) (Finegold JA, et al., 2013). According to the British Heart Foundation, heart attack morbidity rates of 55-59 years old women declined by 50%, during 1968 and 1998, while morbidity rates of 60-64 years old men declined by a third (Scarborough et al., 2011).

In Greece, data from the ATTICA study, demonstrated that CVD risk development was higher by 10 times in men of 65-75 years than men of 35-45 years old. Likewise, the relative risk (RR) was 18 times higher in older women (Panagiotakos et al., 2008).

The joint effect of ageing and population growth contributes to increased absolute numbers of CAD deaths worldwide, underlying the burden of the disease.

Family **H**istory

FH is an independent risk factor for CAD, unrelated to conventional risk factors (Andresdottir et al., 2002; ten Kate et al., 1982). Therefore, it remains an important consideration when assessing the risk for CAD development (Snowden et al., 1982) in clinical settings. A recent meta-analysis, which included case-control and cross-sectional cohorts, demonstrated that positive maternal and parental history of CAD should be considered equally important for CAD transmission to the offspring (Weijmans et al., 2015).

A prospective cohort, the Cooper Center Longitudinal Study (CCLS), examined the association between premature FH and CAD mortality in 49,255 men. The results of the study demonstrated that men with premature FH had approximately 50% higher risk for CAD mortality, compared with men with negative FM (Bachmann et al., 2012).

According to the guidelines of the National Heart, Lung and Blood Institute, a positive FH of CAD is defined as a first-degree relative (a parent and/or a sibling) with a history of treated CAD before the age of 55 in men and 65 years in women. The Expert Panel suggests the inclusion of positive FH in identifying those at risk, since evidence from epidemiological studies have strongly demonstrated that FH is an independent risk factor that infers future CAD incidence (Murabito et al., 2005; Lloyd et al., 2004; Myers et al, 1990).

The fact that the clustering of conventional risk factors could not entirely predict CAD occurrence indicate that genetic susceptibility and gene-environment interactions are also associated with the familial aggregation of the disease. Overview of the evidence for association of genetic variants with increased CAD risk and CAD incidence is described in details in the "Genetic Susceptibility of Coronary Artery Disease" chapter section (section 1.5.).

Gender

Women lag 7 to 10 years behind men in the incidence of CAD. This has been mainly attributed to the protective effect of endogenous oestrogens that is lost after menopause.

Therefore, by middle age, women have a lower risk for CAD development than men. This gap closes, after age 55 and CAD risk is escalated in both men and women (Ouyang et al. 2006).

Despite gender differences in CAD risk between middle-aged men and women, CAD prevalence in women should not be underestimated. The Women's Ischemia Syndrome Evaluation (WISE) study has demonstrated that young women with oestrogen dysfunction have a 7-fold increase in CAD risk (Maas and Appelman, 2010). In 2009, the NHANES demonstrated that the prevalence of MI has increased in middle-aged women (35 to 54 years) and declined in men of the same age (Towfighi, Zheng and Ovbiagele, 2009). Furthermore, subclinical atherosclerosis may be present in premenopausal women with CAD risk factors, which changes to more vulnerable lesions after menopause (Maas and Appelman, 2010).

Men and women still share the same risk factors for CVD development and therefore, CAD incidence. Gender-related differences are yet to be clarified and gender bias in treatment should be eliminated.

1.4.2. Modifiable Risk Factors

Hyperlipidaemia

Hyperlipidaemia is a disorder of lipid metabolism and regulation of the body, clinically manifested by abnormal levels of plasma total cholesterol (TC), low-density lipoprotein cholesterol (LDL-C), triglycerides (TGs) or high density lipoprotein cholesterol (HDL-C).

High TC is a risk factor for CAD (Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults). TC and LDL-C levels were positively associated with CAD death in men in a 12-year follow-up study (Neaton et al., 1992). This also applies for women (Piepoli et al., 2016). In a meta-analysis from 22 trials, it was demonstrated that every 1mmol/l decrease in LDL-C levels (with statin therapy) was associated with a 20-25% reduction in CVD mortality and non-fatal MI (Cholesterol Treatment Trialists' (CTT) Collaborators et al., 2012). Up to date, the association of TC and LDL-C with CVD risk is well documented in many studies. In addition, low levels of HDL-C is an independent CAD risk factor, although attempts to increase HDL-C levels did not decrease CAD risk. Elevated levels of TGs are also considered to be a significant marker for CAD, although the effect on CAD risk is much smaller than hypercholesterolaemia (Piepoli et al., 2016).

Hypertension

Hypertension is defined as systolic blood pressure (SBP) of ≥140mmHg or a diastolic blood pressure (DBP) of ≥90mmHg and/or the use of antihypertensive treatment (Seventh Report of the Joint National Committee of Prevention, Detection, Evaluation and Treatment of High Blood Pressure) (Chobanian et al., 2003). With this definition and based on 2009 to 2012 data, about one third of US adults have hypertension (Mozaffarian et al., 2015). Blood pressure (BP) elevation is proportional to the age of the population. Worldwide, hypertension is estimated to cause 7.8 million deaths, which account for approximately 13% of all annual deaths (WHO, 2011). According to the most recent European Guidelines on CVD prevention in clinical practice, the classifications of BP levels are summarized in Table 1.4.1 (Piepoli et al., 2016).

Table 1.4.1. Classification of blood pressure levels

	Systolic BP[a] (mmHg)		Diastolic BP (mmHg)
Optimal	<120	and	<80
Normal	120-129	and/or	80-84
High-normal	130-139	and/or	85-89
Grade 1 hypertension	140-159	and/or	90-99
Grade 2 hypertension	160-199	and/or	100-109
Grade 3 hypertension	≥180	and/or	≥110
Isolated systolic hypertension	≥140	and	<90

[a]BP= blood pressure

The FHS was the first study that indicated hypertension as an important independent risk factor for CVD occurrence, among other risk factors such as dyslipidemia, glucose intolerance and cigarette smoking (Kannel 1976; McGee and Gordon, 1976). Up to date, a large body of scientific evidence has demonstrated a consistent positive and progressive association between hypertension and CAD. Although hypertension is not the cause of CAD, it remains an important predictor of CAD risk. In addition, the importance of DBP and SBP as CAD risk indicators changes with age progression. DBP was shown to be a stronger predictor for CAD risk, before 50 years of age. On the other hand, after 60 years of age, SBP was shown to be a more important predictor for CAD risk (Franklin et al., 2001). It is noteworthy, that in the same age group (>60 years of age), pulse pressure was the strongest CAD predictor, while DBP was inversely associated with CAD risk. The Chicago Heart Association Detection Project in Industry (CHA) study is a prospective study, on 11,000

men aged from 18 to 39 years followed up for 25 years. One important finding of the study was that men with high-normal BP (130-139/85-89mmHg) or stage 1 hypertension (140-159/90-99mmHg) accounted for 60% of CAD, CVD, and all-cause mortality. In addition, their life expectancy was shortened by 2.2 and 4.1 years, respectively (Miura et al., 2001). High-normal BP is also associated with increased risk of CVD in both women and men. However, the hazard ratio (HR) was greater in women than in men (2.5 and 1.6, respectively), suggesting that borderline BP causes more endothelial dysfunction in women than men (Vasan et al., 2001). A meta-analysis of 61 studies revealed that each 20mm Hg increase of SBP of or 10mm Hg increase of DBP doubles the risk of a fatal coronary event (Lewington et al., 2002).

A common target for BP is <140/90 mm Hg. The AHA statement recommends a lower target goal of BP of <130/80 mm Hg in individuals with established CAD (Rosendorff et al., 2007). At present, an even lower BP goal for CAD prevention or treatment is the subject of debate.

Diabetes **M**ellitus

There are two main types of diabetes mellitus (DM) : Type 1 diabetes mellitus (T1DM) and T2DM. In 2010, along with the established criteria for the diagnosis DM, the American Diabetes Association (ADA) adopted the use of the A1C (hemoglobin A1c or HbA1c) threshold of >6.5%. Therefore, DM is defined accordingly to the criteria summarized in Table 1.4.2.

Table 1.4.2. Criteria for the diagnosis of diabetes mellitus

A1C ≥ 6.5%
or
fasting plasma glucose ≥ 126 mg/dl
or
2-h plasma glucose ≥ 200 mg/dl
or
a random plasma glucose ≥ 200 mg/dl in a patient with hyperglycemia or hyperglycemic crisis

Source: American Diabetes Association, 2013

In 2010, 1.7 million adults (≥ 20 years old) were diagnosed with DM in the US and the burden of the disease worldwide was estimated to be 6.4%. In the same year, 2.7 million deaths were

attributable to DM, which accounted for 5.2% of all deaths. By the year 2030, the prevalence of DM is estimated to reach 7.7%. Although T2DM is considered rare in children and adolescents, the prevalence of the T2DM has reached approximately 50% of childhood DM within 8 years (2001 to 2009). This accounts for a 30.5% increase in the pediatric population (Mozaffarian et al., 2016).

DM is an independent risk factor for CAD (Kannel and McGee, 1979; Kannel an McGee, 1979) and is considered to be a CAD risk equivalent (National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult TreatmentPanel III), 2002). Patients with DM have a 3-fold higher risk for CVD compared to those without DM (fox et al., 2007).

In a population based-study in Germany (MONICA/KORA Myocardial Infarction Registry), the proportion of CAD death, without prior MI, was higher in diabetic than in non-diabetic patients (odds ratio (OR) 1.26; 95% confidence interval (CI) 1.17-1.36) (Icks et al., 2014). The results of a randomized controlled trial (RCT) in Europe, the AngloDanish-Dutch Study of Intensive Treatment In People with Screen Detected Diabetes in Primary Care (ADDITION-Europe), highlighted the importance for early treatment of patients with T2DM. In this study, patients were screened for DM and subsequently assigned either to a routine care or an intensive treatment of the disease. The duration of the follow-up was 5.3 years. CVD risk factors (HbA1c, cholesterol and blood pressure) were significantly, although slightly, improved in the intensive care group. Incidence of first CVD event and mortality rates did not differ significantly between the two groups. The absence of difference could be attributed to the early and high-quality treatment of DM as a risk factor, even in the routine care group. CVD incidence was lower than expected in both groups (Griffin et al., 2011).

The increased risk for CAD attributed to DM per se is independent of other risk factors, namely overweight/obesity and dyslipidemia, commonly observed in patients with established DM. Therefore, screening is essential in children, adolescents and adults , who are overweight/obese and have one or more risk factors for developing T2DM (NCEP Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult TreatmentPanel III), 2002).

The target level for DM, as an important CAD/CVD risk factor is a HbA1c goal of <7% (Authors/Task Force Members et al., 2016).

Overweight and **o**besity

Body weight is defined in categories according to body mass index (BMI) cut-offs. BMI is a simple screening tool for body fatness. According to WHO, overweight is defined as BMI ≥25kg/m^2 and obesity as BMI ≥30kg/m^2 (WHO, 2016). A high value of BMI is an indicator of excessive fat accumulation. BMI cut-off points are depicted in Table 1.4.3.

Table 1.4.3. Body mass index cut-off points

	Body mass index (kg/m^2)	Disease risk
Underweight	<18.5	
Normal weight	18.5-24.9	
Overweight	25.0-29.9	Increased
Obesity class I	30.0-34.9	High
Obesity class II	35.0-39.9	Very high
Obesity class III (extreme obesity)	≥40.0	Extremely high

Source: World Health Organization, 2008

Body fat distribution is also important. Waist circumference (WC) measurement or waist-to-hip ratio (WHR) are also indicators of body fatness. Sex-specific thresholds of WC and WHR, as an example cited in WHO report, are demonstrated in Table 1.4.4.

Table 1.4.4. Body mass index cut-off points

	Cut-off points	Disease risk
Waist circumference	>94 cm (men); >80 cm (women)	Increased
Waist circumference	>102 cm (men); >88 cm (women)	Substantially increased
Waist-to-hip ratio	≥0.90 cm (men); ≥0.85 cm (women)	Substantially increased

Source: World Health Organization, 2008

BMI has substantially increased worldwide and has more than doubled over the last decades. In 2013, the proportion of overweight adults reached 36.9% in men and 38% in women globally. Therefore, the prevalence of overweight and obesity has reached epidemic proportions and has affected –although disproportionally- all age groups, both sexes and both the developed and the developing world (Ng et al., 2013). It is estimated that, by the year 2020, the increase in BMI will counterbalance the positive effects of the observed decrease in smoking rates (Steward et al., 2009).

Overweight and obesity is associated with increased CVD risk, including CAD and contribute to the increase of CVD burden worldwide (Klein et al., 2004). Excess body fat accumulation in the abdomen (intra-abdominal fat) can result in an increased level of BP, blood lipids, insulin resistance, inflammation, ventricular dysfunction and increased prevalence of CAD (Yatsuya et al., 2014; Peterson et al., 2004).

Even modest weight loss, approximately 5% of the initial weight, can improve metabolic features or prevent metabolic abnormalities that are risk factors for CAD (eg, dyslipidemia, hypertension, T2DM, inflammation) (Goldstein, 1992).

Epidemiological studies have demonstrated that weight reduction is associated with a favorable effect on CVD risk factors. Pooled data from 97 prospective studies demonstrated that the HR for CAD was 1.27 (95% CI 1·23–1·31) for every 5kg/m^2 increment for BMI. However, HR was reduced to 1.15 when the analysis was adjusted for risk factors. The authors concluded that the association between BMI and CAD was mediated by these risk factors to a proportion of 46% (Global Burden of Metabolic Risk Factors for Chronic Diseases Collaboration et al., 2014).

In the scientific literature, it is argued that more solid evidence is needed from RCTs to establish whether obesity is directly associated with CAD risk and CAD incidence or obesity risk for CAD is mediated though metabolic risk factors (Alexander, 2001). This task seems even more difficult given that weight loss maintenance for prolonged time is difficult to be achieved.

From public health perspective and given the direct or indirect association between increased BMI and CAD risk, it is important to maintain normal weight and to prevent or control the overweight/obesity.

Diet

Many researchers have tried to explore the relationship between dietary factors and CVD for almost half a century. The effects of many nutrients, foods, and dietary patterns (DPs) on CVD have been evaluated by numerous studies. The Mediterranean-type diet is the most studied and the impact of diet on CVD was first demonstrated by the SCS (Kromhout et al., 1989). Specifically, the Mediterranean-type diet was inversely associated with the incidence of CAD in Southern Europe when compared with the US and Northern Europe, after adjusting for confounding factors. The characteristics of a healthy diet with respect to CAD prevention are summarized in the subsequent paragraphs (Piepoli et al., 2016).

- The type of fatty acids consumed play a more important role than the total fat content of the diet. Therefore, saturated fatty acids (SFAs) consumption should account for no more than 10% and trans fatty acids should account to no more than 1% of total energy intake (preferably from natural origin and not processed foods). A 2-3% reduction in CAD incidence was demonstrated, when 1% of SFA energy intake was replaced by polyunsaturated fatty acids (PUFAs). However, there is a gap in evidence, whether there is a benefit from substituting carbohydrates or monounsaturated fatty acids (MUFAs) for SFAs (Astrup et al., 2011). In addition, an approximate 2% increase in energy intake of trans fatty acids leads to a 23% increase in CAD risk (Mozaffarian et al., 2006). The association of fatty acid intake with serum cholesterol levels is more robust compared with dietary cholesterol. Decreased intake of unsaturated fatty acids also leads to a decreased dietary cholesterol intake.
- Recommendations on fruit and vegetable intake support for 200 grams of fruits and 200 grams of vegetables per day, accounting for 2-3 servings per day. The INTERHEART study, demonstrated that a healthy diet characterized by increased consumption of fruits and vegetables has a protective role against MI (Iqbal et al., 2008). The protective effect of vegetable intake in CAD has been also demonstrated by the Physicians' Health Study in male subjects. Specifically, in this study, men who consumed at least 2.5 servings per day of vegetables had a RR of 0.77 (95%CI 0.60–0.98) for CAD compared with men with vegetable consumption of less than one serving per day (Liu et al., 2001). The WHO attributes about 11 % of ischemic heart disease deaths to insufficient intake of fruits and vegetables (WHO, 2009).
- Sodium intake should be reduced to a maximum of 5 grams per day, while the optimal intake levels are even lower reaching approximately 3 grams per day. Sodium intake reduction seems to be important in CAD and stroke prevention.
- Guidelines for CAD prevention recommend fish consumption 1-2 times per week, preferably oily fish. A meta-analysis of 17 cohorts with a follow-up period of 15.9 years, demonstrated that in comparison with the very low fish intake (<1 serving/month), CAD mortality was lower by 16% in individuals with low fish intake (1 serving/week) (RR=0.84, 95% CI=0.75-0.95), 21% in individuals with moderate fish intake (2-4 servings/week) (RR= 0.79, 95% CI=0.67-0.92) (Zheng et *al*., 2012).
- A recent meta-analysis has demonstrated that nut intake is inversely associated with CAD risk. Specifically, nut consumption of 30 grams per day was associated with 29% decrease in CAD risk (RR =0.71, 95%CI=0.59,0.85) (Luo et al., 2014).

- The consumption of beverages, namely alcoholic and sugar-sweetened beverages (SSB) should be limited or discouraged. Guidelines on healthy diet recommend alcoholic beverages to be limited to 2 glasses per day for men and 1 glass per day for women, that is 20 grams and 10 grams of alcohol per day, respectively. In a systematic review and meta-analysis, dose-response analysis demonstrated lower CAD mortality occurring with moderate alcohol consumption (1-2 drinks per day) (Ronksley, 2011).

A prospective study in women has demonstrated that SSB consumption of 1 serving per day and more than 2 servings per day, was associated with a 23% (RR=1.23, 95% CI=1.06- 1.43) and 35% (RR= 1.35, 95%=1.07-1.69) higher CAD risk, respectively, compared to low consumption (less than 1 serving per month) and after confounding (Fung et al., 2009). WHO recommends sugar intake to be limited to 10% of energy intake or less (WHO, 2015).

- A 30-45% grams of fibre intake is also recommended, since high fibre intake seems to reduce postprandial glucose, TC and LDL-C levels. A recent meta-analysis revealed an inverse association between fibre intake and CAD risk. Specifically, CAD risk was lower by 9% with every 7 grams per day of higher total fibre intake (RR=0.91, 95% CI=0.87-0.94) (Threapleton et al., 2013).

The favorable effect of olive on CAD has long been studied and has been associated with lower CAD risk and mortality (Guasch-Ferré et al., 2014, Buckland et al, 2012). In Greece, the CARDIO2000 study, a multicenter case-control study, showed that exclusive olive oil consumption was associated with a 0.55 times lower likelihood of having ACS among hypercholesterolemic subjects (Kontogianni et al., 2007).

Research on specific dietary nutrients of foods that may highlight the protective effects against CAD is still ongoing. However, apart from focusing to the protective potential of individual nutrients, it is also important to study the impact of DPs on CAD. A DP conceptualizes the nutritional intake, the quality and variety of the overall diet (section 1.6.).

Sedentary **L**ifestyle (inactivity)

PA is defined as "any bodily movement produced by skeletal muscles that results in energy expenditure beyond resting expenditure", according to the AHA scientific statement (Thompson et al., 2003). On the other hand, a sedentary lifestyle is defined as a lifestyle devoid of regular, habitual or leisure-time amounts of PA.

Sedentary lifestyle is a major risk factor for CAD (Artinian et al., 2010). There is solid scientific evidence that physical inactive individuals have lower levels of fitness, a higher risk for

developing medical health issues and higher rates for chronic diseases compared to physical active individuals (Physical Activity Guidelines Advisory Committee, 2008). It was estimated that a 2.3% decrease in physical inactivity within 20 years (from 1980 to 2000) had prevented approximately 5% CAD deaths in the US (Ford et al., 2007).

PA prevents CAD development, improves CAD risk factors and reduces symptoms in patients with established CAD (Thompson et al., 2003). In addition, the scientific evidence has demonstrated that PA amount and dose is inversely related with CAD and stroke (Carnethon et al., 2009). The PA guidelines for adults support for evidence that approximately 150 minutes per week of moderate-intensity aerobic activity, compared with inactivity, can reduce CVD risk (Schoenborn and Stommel, 2011). Recommendation guidelines for adults support for at least 150 minutes per week of moderate intensity aerobic activity and resistance activity (muscle–strengthening) at least two days per week (US Department of Health and Human Services, 2010). Health professionals should advise and encourage individuals to add PA in their everyday life and inform them of how physical inactivity can have deleterious effects on their overall health.

Smoking

Tobacco smoking, including second-hand smoke (SHS), is one of the leading risk factors for disease worldwide. In 2010, tobacco smoking (including SHS) was the leading risk factor for global disease burden among men and the fourth leading risk factor among women. In relative numbers, tobacco smoking accounted for 8.4% of global disease risk in men and 3.7% in women, according to the findings of the Global Burden of Disease Study 2010 (Lim et al, 2012). A 50-year prospective study in Britain demonstrated that non-smokers have 10 years more life expectancy than persistent smokers. In addition, about 50% of persistent smokers were killed by their behavioral habit (Doll, Peto and Boreham, 2004).

Up to date, the scientific evidence is suggestive and not sufficient to support whether smoking is directly and causally related to CAD or is indirectly related. A recent review argues the association of smoking and CAD (Stalones, 2015). Prospective studies have well demonstrated that smoking precedes the onset of CAD, however the evidence of a causal association between smoking and CAD remain contradictory. There is compelling evidence to support that smoking is associated with a 2-fold to 4-fold increase in CAD risk in smokers compared to non-smokers (Bailar, 1999; Doyle et al., 1964; Bronte-Steward, 1961). In

addition, the intensity of smoking further increases CAD risk (Centers for Disease Control and Prevention, 2010; Buechley, Drake and Breslow et al.,1958).

The consistency of association between smoking and CAD in different studies is compelling, despite the argument whether smoking causes CAD or is associated through common factors with CAD. Smoking cessation lowers CAD risk and provides short-term (from 20 minutes to few hours) and long-term (within years) benefits in overall health. Historically has been said that "smoking is now the most dangerous drug addiction", therefore smoking prevention and quitting should be considered to be one of the top priorities in CAD reduction strategy (NCEP Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult TreatmentPanel III), 2002; Hegglin and Keiser, 1955).

1.5. Genetic Susceptibility of Coronary Artery Disease

Differential susceptibility to CAD in randomly selected individuals and familial aggregation of the disease support for a genetic impact for the trait. Familial background of CAD may be directly related to the genetic determinants of the disease, or indirectly through sharing of lifestyle behaviors (eg, smoking, unhealthy diet, excess alcohol intake) or risk factors that may also have genetic determinants (eg., dyslipidemia, hypertension, DM). Two of the first studies that underlined FH as an independent risk factor for CAD was the Nurse's Health Study and the FHS (Colditz et al., 1986; Schildkraut et al.,1989). These studies were conducted on a population of women.

In later years, a twin study in the Swedish population has investigated premature CAD death in twins, after 26 years of follow-up. Among the monozygotic male twins, the relative hazard (RH) when a pair died before the age of 55 years was 8.1 (95% CI, 2.7-24.5). The RH was 3.8 (95% CI, 1.4-10.5) for dizygotic male twins. Among the monozygotic female twins, the RH when a pair died before the age of 65 years was 15.0 (95% CI, 7.1-31.9). Among dizygotic female twins, the RH was 2.6 (95% CI, 1.0-7.1). The results also demonstrated that the impact of genetic variants decreases with older age in both monozygotic and dizygotic twins and in both sexes, with adjustment for risk factors. Therefore, premature CAD death in one pair of twins is a strong predictor of risk of death from the same disease trait (Marenberg et al., 1994). The same study further investigated the proportional contribution of genetic variants to CAD mortality, after 36 years of follow-up. Specifically, the study showed that the heritability of CAD deaths was 0.57 (95% CI, 045-069) among male twins and 0.38 (95%CI, 0.26-0.50) among female twins (Zdravkovic et al., 2002). In younger individuals, it has been

demonstrated that genetic variants may contribute to a 20% to 60% increase in CAD risk (Chaer, Billeh and Massad, 2004).

The identification of genetic markers for CAD is a challenging task for scientists and therefore, research to identify these markers is extended and chequered during the last 25 years. CAD is a common and complex disease and does not exhibit Mendelian inheritance, therefore genetic variants may increase CAD risk without necessarily cause the disease. The effect size of a genetic contributor to a complex disease trait may be small and dispersed within a population, or large affecting a small population or may have a larger effect when interacting with an environmental factor (Mozaffarian et al, 2014).

The first genetic association studies were focused on the candidate gene approach, that is based on an a priori hypothesis of the plausible involvement of the selected gene in the pathogenesis and process of the disease under investigation. Many studies examined single nucleotide polymorphisms (SNPs) in single candidate genes, while others focused on more SNPs in the same gene. Historically, the success of the first candidate gene study in identifying a genetic variant in the susceptibility of CVD was published in Nature in 1992. Specifically, the study explored a variant found in the gene encoding angiotensin-converting enzyme (ACE) and showed that homozygotes for a deletion in the ACE gene were at higher MI risk (Cambien et al., 1992). Another example reflecting the success of this approach is the identification of variants in the apolipoprotein E (APOE) gene, which has a major role in cholesterol metabolism (Song, Stampfer and Liu, 2004). Candidate gene studies have been rather effective in identifying variants associated with rare and monogenic forms of CVD (Mendelian disorders). However, this methodological approach has appeared rather slow and unsuccessful in identifying novel genetic variants of polygenic CAD, mainly because of small genotyped samples sizes and limited tested SNPs.

In order to overcome the limitations from candidate gene studies, another methodological approach was developed, the genome-wide association study (GWAS), which was confronted with great excitement and optimism from scientists in the genetic epidemiology. The two methodologies derived from the candidate gene study and the GWAS are summarized in Table 1.5.1. (Frazier-Wood, 2015).

GWAS enabled the scanning of the entire genome in order to seek out associations between hundreds of thousands of SNPs and diseases. Historically, the first GWAS published results were in 2002 and evaluated the association of 92,788 SNPs with MI. The large-scale study underlined association between a marker on chromosome 6p21 and MI (Ozaki et al., 2002).

In the years after, GWAS on CVD identified some SNPs on 9p21.3 loci associated with CAD (Helgadottir et al., 2007; Mc Pherson et al, 2007; Wellcome Trust Case Control Consortium, 2007; Samani et al, 2007). The C4D Consortium identified five novel loci, including the lipase (LIPA) gene (Myocardial Infarction Genetics Consortium et al., 2009). A meta-analysis performed by the Coronary Artery Disease Genome-wide Replication and Meta-analysis (CARDIoGRAM) Consortium found that rs1333049 on the 9p21 region confers a 29% increase in MI risk per allele (Preuss et al., 2010). Another large-scale association analysis identified 13 new loci associated with CAD (Schunkert et al. ,2011). Specifically, the risk alleles for the new loci were associated with an increase in CAD risk ranging from 6 % to 17 % per allele. Up to 2011, 35 variants associated with CAD were reported (Peden JF, Farrall, 2011).

The GWAS are an asset in the toolbox of genetic epidemiology and so far many SNPs have been associated with complex diseases. However, the "GWA era" has been less rapid than expected to unravel the majority of the genetic variance associated with a polygenic disease. Up to date for instance, approximately 90% of the genetic variants for CAD, a commonly studied disease, remain unaccounted for. In addition, many initial GWA findings failed to replicate.

Table 1.5.1. Methodologies of the candidate gene study and genome-wide association study in comparison

Candidate gene study	Genome-wide association study
Small number of SNPs[a] are tested	Provides information on ≈80% of the human genome
An a priori hypothesis approach	A hypothesis-free approach
A statistically powerful approach.	Correction for multiple testing is needed. This leads to less power of the study
Identifies variants with small, modest and large effect	Identifies variants with modest or large effect dependant on the study size
Includes rare variants and non-SNP variants	Includes common variants (<1% in the population). However, some micro-arrays contain rare variants

[a]SNPs = single nucleotide polymorphisms

1.5.1. Genetic risk scores

In GWAS, an issue needed to be resolved was that significant associations were found for small effect sizes of variants but the significance did not survive after multiple testing. The

genetic risk score (GRS) was a methodological solution to overcome the limitation of small effect sizes. A GRS is a sum of alleles from various genes, known to be associated with the same phenotype. Summing up the alleles can show a larger effect size than the SNP alone. The following studies are illustrative examples of GRSs evaluation.

Two case-controls studies, the Ottawa Heart Genomics Study (OHGS) and the Wellcome Trust Consortium (WTCCC) compared the ability of prediction of 9p21.3 alone and a GRS constructed from a panel of 12 CAD SNPs. According to the findings of the study, the GRS demonstrated a greater predictive capability than risk factors or risk factors plus 9p21.3. (Davies et al., 2010).

Three GRSs were applied on 2887 participants of the FHS and were tested regarding CVD and Coronary Artery Calcium (CAC) incidence. A GRS based on 13 SNPs (13-GRS) robustly associated with CAD, a GRS based on 102 CAD SNPs and traditional risk factors (102 GRS) and a GRS based on 29 SNPs (13 SNPs from the 13 GRS plus 16 additional SNPs) (28 GRS). The 102 GRS was not associated with CAD or CVD events. The 13 GRS was associated with CAD incidence (HR 1.07; 95% CI 1.00-1.15; p=,0.04), CVD incidence (HR 1.05; 95% CI 1.01-1.09; p=0.03) and increased CAC (OR 1.18; 95% CI 1.11-1.26; $p=3.4\times10^{-7}$). The additional SNPs of the 29GRS did not result to an improved CAD predictive capability (Thanassoulis et al., 2012).

Relevant to the goals of genetic research is the identification of variants explaining a larger proportion of the disease. Therefore, a GRS implementation for the prediction of future events in clinical practice remains one of the perspectives in the near future.

1.6. Dietary Patterns and Coronary Artery Disease

Nutrition research has systematically tried to identify the health efficiency of a single dietary factor. However, nutrients and bioactive chemicals of food items are inter-correlated and research for association between a single nutrient and a chronic disease may underestimate the impact of the overall diet on health outcome (Tucker, 2010).

Intervention studies with nutrient supplementation failed to demonstrate a protective association between the nutrient intake and the investigated disease. For example, vitamin E or β-carotene supplementation did not yield the expected protective effect on cardiovascular health (Devaraj and Jialal, 2005; Omenn et al., 1996). In addition, interactions among dietary components of foods and within the food matrix are complicated and not yet completely understood by the scientific community. Given the above and as a consequence,

investigation on DPs rather than a single nutrient has gained ground and appreciation in nutritional research.

As aforementioned in section 1.4.2, DPs measure the nutritional intake, the quality and variety of food combinations in individuals or in groups within a population. The aim of DPs analysis is to test the hypothesis and investigate the association, if any, of DPs with health or disease outcomes. Therefore, it is important to include nutritional intake measurement in the design of the study. Nutritional intake assessment may be done through 24-h recall, diet records or food-frequency questionnaires (FFQ) (Newby and Tucker, 2004). DPs have been approached by using either an a priori index or data driven analysis.

The a priori approach starts by designing a DP based on recommendations from scientific consensus or proposed by researchers using solid scientific evidence. Subsequently, a score or index is derived from the points given to individuals according to their adherence to a set of foods or nutrients that consist the DP and after summing up the points for all constituents of the DP. A popular example of the a priori approach is the Mediterranean Diet Score (MedDietScore). The MedDietScore ranges from 0-55 and higher values of the score represents greater adherence to the MD (Panagiotakos, Pitsavos and Stefanadis, 2006). In the Greek population, it was demonstrated that a 10-unit increase of the MedDietScore is associated with 4% lower CAD risk for the subsequent ten years of life. Other examples of diet scores are the Dietary Approaches to Stop Hypertension (DASH), the Healthy Eating Index (HEI)-2005 and HEI-2010 (Guenther et al., 2013; Fung et al., 2008).

A second method for DPs is the posteriori analysis, which is a data-driven approach. Statistical methods that can be used for this purpose are the factor analysis (FA), or the very similar principal component analysis (PCA) and cluster analysis. Recently, reduced rank regression (RRR) has been proposed and used for DPs analysis. In posteriori analysis, DPs are defined in the existing dietary habits within a population and subsequently, investigators explore the association between the derive DPs and health outcomes or biomarkers. DPs reflect the association between the overall diet and health outcomes of interest (Tucker et al., 2010).

The majority of studies reported an inverse association between a healthy DP and CAD risk or mortality. Two meta-analysis have demonstrated that for each 2-unit increase in the adherence to the MD was associated a 10% reduction in CVD events (fatal or non-fatal) (Martinez-Gonzalez and Bes-Rastrollo, 2014; Sofi et al., 2013).

The results of the PREDIMED study demonstrated that individuals who followed a MD (plus extra virgin oil or nuts) had a decreased risk of CVD events (MI, stroke and deaths) compared to controls who followed dietary advice on fat reduction (Estruch, Ros and Martínez-González, 2013).

There is strong scientific evidence that a healthy DP has a beneficial impact on CVD. A healthy DP is characterized by higher consumption of whole grains, low fat dairies, fruits, vegetables and seafood; regular consumption of legumes and nuts; moderate consumption of alcohol; lower consumption of red and processed meat, refined grains and sugar-sweetened beverages (SSB).

1.7. Aims and Objectives

The objectives of the current PhD thesis were a) to record a wide range of markers between individuals with established CAD and individuals free of the disease, b) to identify novel loci that are associated with CAD risk through data sharing (Consortium participation), c) to evaluate whether GRS is associated with CAD risk and d) to identify DPs associated with CAD risk.

The aims of each report supporting the PhD thesis were:

Paper 1

To present the demographic characteristics, clinical characteristics/biochemical indices and lifestyle habits of the Greek sample; and to explore the potential association of exclusive olive oil consumption, in relation to lifestyle factors, with CAD risk.

Paper 2

To evaluate the association of a healthy lifestyle pattern in a Greek sample, by means of a preventive score, with glycaemic and adiposity traits; and to evaluate whether this lifestyle pattern modifies the association of known glucose-raising genetic variants on glycaemic traits.

Paper 3

To evaluate whether a GRS, constructed by 53 previously reported SNPs, is associated with CAD risk in Greek adults.

Paper 4

To identify DPs and to evaluate their association with CAD risk. To investigate the effect of genetic predisposition in the context of an environmental component, that is diet.

Paper 5

To identify additional susceptibility loci for CAD, using sample collections from European or south Asian descent. Additionally, to identify a broader set of SNPs; and to use this set to undertake network analysis to find key biological pathways underlying the pathogenesis of CAD.

Paper 6

To evaluate associations of a lifestyle score with fasting glucose (FG) and fasting insulin (FI) levels; to evaluate whether genotypes at known loci associated with FG and FI; and to modify the associations of diet with FG and FI, using data from multiple US and European cohort studies.

Chapter 2

SUBJECTS, MATERIALS & METHODS

2 | SUBJECTS, MATERIALS & METHODS

The study design, the materials and methods used are given below with subsequent details.

2.1. The THISEAS study

The **H**ellenic study of **I**nteractions between **S**ingle nucleotide polymorphisms and **E**ating in **A**therosclerosis **S**usceptibility (THISEAS) is a medical center-based case-control cohort. The pilot phase of the cohort was initiated in 2006. The sample under study comprised of CAD patients and subjects free of any CVD. Maintenance of confidentiality of the subjects' personal identity, medical and genetic information was ensured at all stages of the study. The study protocol was approved by the Ethics Committee of Harokopio University.

2.2. Study sample

Study participants were consecutively recruited from: (a) major hospitals, (b) Centers of Open Protection for the Elderly and (c) Municipalities, all placed in the region of Athens.

Controls were either outpatients for routine examinations or inpatients in departments other than the cardiology clinics, who visited the same hospitals and at the same period with the coronary patients. Healthy volunteers from Centers of Open Protection for the Elderly and Municipalities of the area of Athens were also enrolled in the control group. In order to reduce the effect of possible, unknown confounders and to eliminate the problem of misclassification, we retrieved precise information from medical history through the hospital or insurance records. Therefore, controls were subjects with negative coronary angiography findings, or negative stress test, or subjects without clinical symptoms of the disease, of any CVD, cancer, or inflammatory diseases.

Cases were enrolled from cardiology clinics of Athens' hospitals. Cases were subjects presenting with either acute coronary syndrome (ACS) or stable CAD defined as >50% stenosis in at least one of the three main coronary vessels assessed by coronary angiography. ACS was defined as acute MI ranging from ST-segment (STEMI) to non-ST-segment elevation (NSTEMI) and unstable angina (Kalra et al. 2008). ACS patients had also undergone coronary angiography that verified the presence of significant stenosis. Subjects with renal or hepatic disease were excluded from both study groups. In the group of cases,

60.4% were diagnosed with first-time CAD at the time of recruitment while, 49.6% were diagnosed with CAD six months or more before the time of recruitment.

All participants were informed about the goals of the study and gave their written consent. A total of 2886 people were invited to participate in the study and 2565 were finally recruited from 2006 to 2010.

2.3. Demographic assessment

Demographic and lifestyle characteristics, personal and family medical history of CVD risk factors, dietary habits and PA habits were assessed during recruitment of the study sample. Age (in years) was recorded: i) at the time of recruitment for controls and ii) at the time of first-time CAD diagnosis and at the time of recruitment for cases.

All participants were interviewed regarding their origins to ensure their Greek ancestry. Their educational status was measured by the years of schooling and was classified into four groups: primary school (0-6 years), high school (7-12 years), senior high or technical institutes (13-15 years), university (>=16 years). Senior high or technical institutes and university were combined into one group. The marital status of the participants was also recorded and defined as single, married, divorced or widowed. The annual financial status of the participants was classified into four groups: low (<8000 euro/ year), medium (8000-15000 euro/ year), high (15000-2000 euro/ year) and very high (>20000 euro/ year). Low and medium financial status were combined into one group (low/medium). Similarly, high and very high financial status was combined into one group (high/very high). Furthermore, data regarding the professional status was collected and the participants were accordingly characterized as public servants, private employees, freelancers, partially employed, unemployed, retired and householders.

2.4. Anthropometrical measurements

Volunteers underwent anthropometric measurements by trained dietitians of the research team. The equipment and the protocol used was the same for all subjects that participated in the study. Basic anthropometric indices, such as weight, height, WC and hip circumference (HC), were measured.

2.4.1. Weight and height measurements

Body weight and height were measured in all participants, wearing light clothing, without shoes. Weight was measured to the nearest 0.5kg using a leveled platform scale. Height was measured to the nearest 0.5cm using a wall-mounted stadiometer. Participants were asked to stand erect, with the feet, knees, buttock and shoulder blades in contact with the vertical surface of the stadiometer. The examiner then placed the participant in the Frankfort horizontal plane. BMI was computed as weight (kg) / height^2 (m) (Quetelet's equation): Body Mass Index= Weight (kg)/ Height^2 (m).

2.4.2. Waist and hip circumference measurements

WC and HC were measured using WHO protocol (WHO, 2008). A measurement tape was used for both measurements. WC was measured to the nearest 0.1cm by placing the measurement tape at the approximate midpoint between the lower margin of the last rib and the iliac crest. HC was measured around the widest portion of the buttocks to the nearest 0.1cm

The WHR was computed as the ratio of the circumference of the waist to the circumference to the hips: Waist-to- hip ratio = waist circumference (cm)/ hip circumference (cm).

2.5. Hematological and biochemical measurements

Blood samples were collected after 12-hour fasting between 6 and 10a.m. A physician performed venipuncture to obtain 10ml blood from each participant. The blood was distributed to two types of test tubes, one of which contained EDTA. In total three tubes were used for i) total blood count, ii) biochemical indices and iii) DNA isolation analyses for each participant. Furthermore, blood samples were centrifuged and both plasma and serum were stored at -80^{o}C for future measurements.

2.5.1. Total blood count

EDTA-blood was transferred on the same day in a local laboratory for analysis in a hematological auto-analyzer. Hematological indices obtained were leukocytes, neutrophiles, lymphocytes, monocytes, eosinophils, red blood cells (RBC), hemoglobin, hematocrit, mean corpuscular volume (MCV), mean corpuscluar hemoglobin (MCH), coefficient of variation of red cell distribution width (RDW-CV), platelets and mean platelet volume (MPV).

2.5.2. Biochemical measurements

Measurement of blood lipids, such as TC, HDL-C, TGs were determined using the enzymatic colorimetric assay (ACE analyzer). LDL-C was calculated from the Friedewald equation (Friedewald et al., 1972): LDL-C (mg/dl)= TC (mg/dl) – HDL-C (mg/dl) – (TGs(mg/dl)/5).

Glucose levels were also measured using the enzymatic colorimetric assay (ACE analyzer). Serum insulin was measured via immunofluorescence on an automatic analyzer with direct chemiluminescence immunometric assay, sandwich type, of two points, utilizing constant quantities of two antibodies according to the manufacturer's instructions (ST AIA pack IRI, Tosoh AIA System Analyzers, San Francisco, CA). Insulin resistance was estimated using the homeostasis model assessment (HOMA-IR) with the following equation (Matthews et al., 1985): HOMA-IR = fasting insulin (μIU/ml) × fasting glucose (mmol/L)/22.5.

2.5.3. DNA isolation

EDTA-blood was also used for DNA isolation. Genomic DNA (gDNA) was extracted from whole blood using the salting-out method (Miller et al., 1988). For each participant, DNA isolation was performed twice and each DNA sample was stored in two aliquots with TE buffer at -20°C. DNA isolation was performed in the Laboratory of Biology, Biochemistry, Physiology and Microbiology at Harokopio University.

2.6. Clinical assessment

A physician performed clinical assessment of the volunteers through a questionnaire during an interview. In order to eliminate recall bias, we also tried to reclaim medical information from the hospital or insurance records for both study groups, where available. Hypercholesterolemia was defined as TC levels greater than 200 mg/dl or use of hypolipidemic medication. Diabetic subjects were those with blood glucose levels greater than 125 mg/dl or subjects that were under special diet or treatment.

BP level was measured in right arm, having the volunteer seated and rested, by using a mercury sphygmomanometer. Hypertension was defined as BP levels greater than 140/90 mmHg or use of antihypertensive medication.

FH of premature CAD was considered when the age at which first-degree relatives develop symptomatic CAD is under 55 years for male and under 65 years for female.

2.7. Lifestyle assessment

Information regarding dietary intake, PA habits and smoking status was also collected.

2.7.1. Dietary assessment

Nutritional information was obtained by 60.4% (N= 737) of patients that were first diagnosed with CAD within six months by the time of their recruitment in the study. The volunteers were asked about their nutritional habits before CAD diagnosis. Nutritional information from volunteers with CAD diagnosis for more than six months at the time of recruitment was not collected, in order to eliminate a) bias regarding possible modifications of nutritional habits after CAD diagnosis and b) recall bias regarding nutritional habits before CAD diagnosis.

We tried to retrieve dietary information from double number of controls, in order to ensure 2:1 controls to cases ratio. After excluding subjects with incomplete or missing dietary data and extreme values of energy intake, the main sample with nutritional information included 499 cases and 832 controls.

Dietary data was collected through a 172-item picture-sort FFQ. Participants were asked to indicate how often they consumed various foods and beverages, as well as the portion size by comparison with photos. Daily consumption was calculated from the FFQ by multiplying the standard serving size of each food (as described by the Ministry of Health and Welfare, Supreme Scientific Health Council) by the value corresponding to each consumption frequency: never; 1–3 times/month; 1–2 times/week; 3–4 times/week; 5-6 times/week; 1 time/day (Supreme Scientific Health Council & Ministry of Health and Welfare of Greece, 1999). The FFQ was either self-administered or self-interviewed. In both cases, well-trained nutritional professionals were available to assist the participants during questionnaire completion.

Nutritionist Pro, version 2.2 software (Axxya Systems-Nutritionist Pro, Stafford, TX, USA) was used to analyze nutritional information regarding energy, macronutrient and micronutrient intakes.

2.7.2. Mediterranean Diet Score evaluation

The adherence to the MD was evaluated, by calculating a MedDietScore that is an a priori defined score. The range of that score is between 0–55; higher values suggest greater adherence to the traditional Mediterranean DP that is characterized by high consumption of

plant foods (non-refined cereals, vegetables, and fruits) and olive oil, by moderate consumption of dairy products, poultry and fish and by low consumption of meat and meat products (Panagiotakos et al., 2007). The MedDietScore was split by median into two categories to indicate the volunteers with low adherence (below median) and high adherence (above median) to the MD.

2.7.3. Evaluation of olive oil consumption

The FFQ included additional questions to assess the type of fat/oil used in cooking. Participants were asked to report the type of fat/oil used during the preparation of a) salads, b) cooked meal and c) fried meal. Each participant could indicate more than one choice of fat/oil namely, olive oil, seed oil, corn oil, sunflower oil, butter, and margarine, as depicted in **Table 2.1.**

Regarding olive oil consumption, participants were classified into two groups: those who exclusively used olive oil during meals and salad preparation and those who consumed other types of fat/oil as well.

Table 2.1. Questions used to assess the type of oils/ fats added in food preparation.

What type of fats do you use in cooking? (Please list all used)	What type of fats do you use in vegetables? (Please list all used)	What type of fats do you use in fried meals? (Please list all used)
☐ Olive oil	☐ Olive oil	☐ Olive oil
☐ Seed Oil	☐ Seed Oil	☐ Seed Oil
☐ Corn oil	☐ Corn oil	☐ Corn oil
☐ Sunflower oil	☐ Sunflower oil	☐ Sunflower oil
☐ Butter	☐ No oil/ Do not add fat	☐ Butter
☐ Margarine		☐ Margarine

2.7.4. Food groups

Food and nutrient intake by the FFQ were examined at the level of food groups and DPs. Dietary data were manually entered into an Excel spreadsheet database that translated the queried foods and beverages into food group equivalents. Regarding combinations of individual foods into one food item, the researcher consulted the ingredients, nutrient information and recipe from reference lists. The dietary guidelines for adults in Greece were used for portion size calculation (Supreme Scientific Health Council & Ministry of Health and

Welfare of Greece, 1999). In total, 26 food groups were estimated and are depicted in **Table 2.2**. For analysis, the number of food groups was further narrowed down to 21 by including two food groups into one (e.g. dairy, full fat and cheese, full fat).

Table 2.2. Assessment of food groups

Refined starch	Refined breakfast cereals, chocolate cereals, white toast bread, white bread, sesame bagel, rice, spaghetti,
Non-refined starch	All bran cereals, muesli, whole wheat toast bread, whole wheat bread
Fruits	Apple, orange, banana, pear, kiwi, tangerine, apricots, peach, watermelon, melon, strawberries, figs, grapes, cherries, dried figs
Fresh fruit juice	Fresh fruit juice, fruit juice (100%)
Fruit juice	Fruit juice (<100%)
Vegetables	Tomatoes, cucumber, onions, turnip, garlic, carrot, lettuce, cabbage, spinach, broccoli, greens, eggplants, pumpkins, artichokes
Potatoes	Baked potatoes, boiled potatoes, smashed potatoes
Potatoes	Fried potatoes
Legumes	Lentils, beans
Fish (including fried fish)	small fish, big fish
Poultry	Chicken breast, chicken leg
Red meat	Pork, beef, lamb/goat, liver
Processed meat	Salami, sausage, turkey, bacon,
Dairy, full-fat	Milk full fat, yogurt full fat
Dairy, semi-fat	Milk, semi-fat, yogurt semi-fat
Dairy, non-fat	Milk non-fat, yogurt semi-fat
Cheese, full-fat	Feta, parmesan, gruyere
Cheese, semi-fat	composite cheese
Eggs	Egg boiled, egg fried/scrambled
Fast foods	Pizza, cheese pie, ham-cheese pie, spinach pie
Sweets	Chocolate almond, dark chocolate, mini chocolate, chocolate croissant, cookies, chocolate cake, ice cream, ravani, baklava, pecan pie, donut, nougat, halvah, sugared bun, melomakarono
Nuts and seeds	Almonds, peanuts, hazelnuts, cashew, walnuts, pine nut
Beverages with sugar	Soft drinks
Beverages, non-sugar	Soft drinks light
Beverages with alcohol	Beer, red wine, white wine, whiskey, ouzo, liquor,
Coffee	Coffee

2.7.5. Physical activity assessment

Physical activity level (PAL) was assessed through Harokopio Physical Activity Questionnaire (HAPAQ) (Maraki et al., 2010) that evaluates the frequency, duration and intensity of occupational, household and leisure-time activities. PAL was calculated as the ratio of the work metabolic rate to the resting metabolic rate of 1 metabolic equivalent (MET). The work metabolic rate was defined as the person's total energy expenditure in a 24-hour period, based on performed physical activities. The caloric cost of each PA was estimated with the equation, kilocalories = MET × weight in kilograms × duration in hours. The 1 MET reference value of 1 $kcal \cdot kg^{-1} \cdot h^{-1}$, is used by convention and refers to a metabolic rate at rest of an average individual[18]. PA data was imported into an Excel spreadsheet that was properly designed to automatically calculate PAL value for each participant.

PA adoption was assessed as a categorical variable categorizing the participants into two groups, whether they reported leisure-time activities in a regular basis or not.

2.7.8. Smoking status assessment

Participants were asked to report whether they were current, never or former smokers. The average daily number of cigarettes, the duration of smoking and time of cessation were also estimated. Passive smoking exposure at home or work was also recorded by answering a close-ended type of question. Current smokers were defined as those who smoked at least one cigarette per day, non-smokers those who have never smoked in their life and former smokers those who have stopped smoking for at least six months. Former smokers were either combined together into one group with current smokers or with never-smokers.

Pack-years were also calculated by multiplying the number of cigarette packs smoked per day by the years of smoking.

Number of pack years = (number of cigarettes per day/20) × years of smoking

2.8. Genotyping

The genotyping procedure consisted of sample genotyping on the Metabochip and OmniExpress arrays and subsequent application of data quality control (QC).

2.8.1. The Metabochip micro-array

After the completion of DNA isolation, gDNA samples were transported in dry ice to Wellcome Trust Institute, Hinxton, UK, where genotyping took place. For this cause, the Metabochip was used, which is a custom genotyping array that comprises of 196,725 SNPs (iSELECT chip, Illumina, San Diego, CA, USA) associated with T2DM, CAD, MI and related quantitative traits (BMI, BP, lipid, FG and FI levels). It represents a cost-effective and valuable tool in genetic studies that endeavor to understand the architecture of common and complex human traits. The genotype calling algorithm was GenoSNP. A total number of 1887 samples were successfully genotyped.

2.8.2. The OmniExpress array

A subset of 1075 samples were also genotyped at Wellcome Trust Institute, Hinxton, UK, using the OmniExpress array (Human OmniExpress 12v1, Illumina, San Diego, CA, USA), which comprises of 733.202 SNPs. The genotype calling algorithm was Illuminus.

2.8.3. Data Quality Control

The next step after genotyping was the data QC. Data QC is an important procedure prior to association testing. A careful assessment of data quality could elude or diminish the rate of false-negative or false-positive disease associations. QC filters were applied in two levels, including sample quality metrics and SNP quality metrics.

Sample quality metrics were applied as described in the following lines.

- Samples with low call rates were filtered out. The standard threshold for excluding samples with low call rate was 95%.
- Samples with sex-discordance were identified, by comparing the sex ascertained during sample recruitment with the heterozygosity rate across all chromosome X SNPs. For female samples a high heterozygosity rate was expected, while the reverse was applied for male samples. The comparison between recorded and genotype sex was automatically performed by PLINK v.1.07 (Purcell et al., 2007). Sex-discordance could result from a sample-mix up or plating error.
- The mean and the standard deviation (SD) of heterozygosity in samples with over- or under-abundance of heterozygous SNPs were computed. Individuals with heterozygosity falling outside the mean ± SD were excluded.
- Duplicated and related samples were identified and removed from the cohort.

▪ Samples with divergent ancestry (non-Europeans) were also identified as ethnic outliers and removed from the cohort. The method used was multidimensional scaling implemented in PLINK.

SNP quality control set for the study is described in the following lines.

▪ Samples that diverge from Hardy-Weinberg equilibrium (HWE) are likely to be subject to genotyping or genotype errors. HWE filtering was based on controls since deviations could be also indicative of selection. In this study, the significance threshold for HWE was 10^{-4}. This practically means that SNPs with *p*-value < 10^{-4} were removed.

▪ Markers with a call rate less than 98% were removed from further study.

▪ The linkage disequilibrium (LD) cut-off for two SNPs in the same locus was $r^2 < 0.5$.

▪ The last step was to remove the SNPs with low minor allele frequency (MAF). A MAF threshold of 1% was applied.

SNPs that did not meet the quality control criteria were imputed. Imputation was performed based on the 1000 Genomes panel and IMPUTE2 (Howie, Donnelly and Marchini, 2009). The haplotyping software was SHAPEIT (Delaneau, Marchini and Zagury, 2012).

Genetic **r**isk **s**core **m**odeling

The weighted GRSs were calculated by counting the number of risk alleles (0, 1 or 2) carried by each individual for all SNPs selected, after multiplying the number of risk alleles per SNP by its β-estimate and then summing up the number of risk alleles across all SNPs tested (Hivert et al., 2011). Finally, the score per individual was divided by the average of the β-estimates of the SNPs (Rasmussen-Torvik et al., 2011). The unweighted GRSs were calculated by counting the number of risk alleles (0, 1 or 2) carried by each individual for all SNPs and then summing up the number of risk alleles across all SNPs tested.

2.9. Statistical analysis

Variables and **A**ssociations

Continuous variables are presented as mean values and SD, while categorical variables are presented as relative frequencies. Differences between categorical variables and groups of the study were assessed using the χ2 test. P–P plots were applied to assess the normality of the distribution of the continuous variables. Student's t test or the Mann–Whitney test was applied to evaluate differences in continuous variables between the two study groups.

Regarding genotypic data information, raw data manipulation and recoding for GRSs construction was performed in PLINK v.1.07 tool set (Purcell et al., 2007).

Logistic regression analysis was performed in order to estimate the relative risks of developing CAD by the calculation of ORs and their corresponding 95% CIs. In addition, unadjusted or adjusted logistic regression analysis was performed in order to test the association between GRS (based on a panel of 53 CAD variants, GRS-53) and CAD risk, by calculating the OR and their corresponding 95% of CIs. The calculation of ORs for CAD risk was also performed on the basis of covariates namely, age, sex, BMI, hypertension (yes vs. no), T2DM (yes vs no), smoking status I (current/former vs. never smokers), smoking status II (current vs. never/former smokers), PA adoption or physical inactivity (no vs. yes), energy intake (above or equal to vs below median) and WHR (above or equal to vs below median). Energy intake and WHR were split into two quantiles with regard to the median, separately for men and women. Logistic regression was also used for associations between each dietary component with CAD risk, without adjustments or after controlling for age, sex, BMI. For the interaction associations, we included a Glucose Preventive Score (GPS) × GRS cross product along with the covariates (GPS, GRS, age, sex, BMI and energy intake) and the glucose levels as an outcome. In order to further characterize the direction of the interaction, we conducted a stratified analysis. For the lifestyle score and the GRS, the coefficient of determination (R) was used as a measure to express the proportion of total variation explained by the model. Genotype by environment (G × E) variance contribution to the total variance of glucose levels was estimated using genome-wide complex trait analysis (GCTA).

The impact of each variant on CAD risk was examined through logistic regression, unadjusted or after confounding and assuming an additive genetic model.

In order to evaluate the combined effect of genetic predisposition and adoption of the MD we calculated the ORs for each of the following categories: a) low GRS + high western diet adherence, b) high GRS + low western diet adherence and c) high GRS + high western diet adherence, using as reference subjects with low GRS and low western diet adherence. Low and high GRS, as well as low and high western diet adherence were set on the basis of the median value for both variables.

Factor **A**nalysis

In addition, FA technique was conducted to identify DPs. Exploratory FA was carried out to evaluate validity, disclose underlying structures and reduce the number of variables. FA was chosen as extraction method using orthogonal rotation (Varimax rotation) in order to generated non-correlated components (namely, non-correlated DPs). The food variables that were highly correlated showed factor loadings (correlation coefficients) greater that |0.4|. The cut-off point for Eigen values was greater than 1.0.

Statistical **A**nalyses **S**oftwares

Analyses were based on 2-sided tests, while statistical significance was set at $p \leq 0.05$. The statistical software packages IBM SPSS Statistics 13.0 and 21.0 (SPSS Inc., Texas, USA) were used for all statistical calculations, where appropriate. Genetic analyses were performed in SNPTEST v.2.5.2 (https://mathgen.stats.ox.ac.uk/genetics_software/snptest/snptest.html). The statistical threshold for these analyses was based on Bonferroni correction for multiple testing. We also used Quanto v1.2.4 for power calculations (http://hydra.usc.edu/gxe/).

Chapter 3

RESULTS & DISCUSSION

3 | RESULTS & DISCUSSION

This section presents the manuscripts derived during the implementation of the current dissertation. The manuscripts (published or unpublished reports) are grouped in the following subsections as depicted in the following outlook.

Outlook of the papers included in the doctoral dissertation		
Published reports (subsection 3.1)	**Unpublished reports** (subsection 3.2)	**Consortia participation** (subsection 3.3)
Paper 1 Design of the THISEAS study and exclusive olive consumption on CAD risk	**Paper 3** The additive influence of genetic variants on CAD risk, by means of a GRS	**Paper 3** Novel loci associated with CAD. (CARDoGRAMplus C4D Consortium).
Paper 2 Lifestyle pattern x gene interaction in glycaemic traits	**Paper 4** DPs and CAD risk via FA Combined effect of gene susceptibility and a dietary component on CAD risk	**Paper 4** Diet x gene interaction in glycaemic traits (CHARGE Consortium)

Results & Discussion

Published reports: Papers **1-2**

3.1. Published reports Paper 1 and Paper 2

Paper 1 | THE THISEAS STUDY

This paper presents an overview of the design and methodology of the THISEAS study. The clinical characteristics/biochemical indices, along with the lifestyle habits of the participants are presented. The association of a dietary factor, that is exclusive olive oil consumption, with CAD risk was also examined.

Key points

- A total of 2565 adults were enrolled in the case-control study, specifically 1221 patients with CAD diagnosis and 1334 healthy individuals, from a Greek sample.
- Blood collection and analysis, demographic, anthropometric, clinical, PA and dietary assessments were performed.
- Higher education and financial status were recorded to the control group compared to the case group
- The conventional CAD risk factor, namely hypercholesterolaemia, hypertension, DMT2 and FH were more prevalent in cases than controls.
- With regard to lifestyle habits, controls reported higher PAL when compared to cases.
- Also, higher prevalence of smoking was recorded in CAD patients, who reported a higher number of pack-years by double than controls.
- Exclusive olive oil consumption was associated with lower CAD risk, after adjusting for confounding factors.

Paper 2 | LIFESTYLE - GENE INTERACTION IN GLYCAEMIC TRAITS

In this work, data from individuals without diabetes, drawn from the THISEAS study, were utilized for the analyses.

Key points

- A healthy lifestyle pattern, through a preventive score, was tested for association with glycaemic traits. In addition, it was examined whether this lifestyle pattern modifies the association of known glucose-raising SNPs on glycaemic traits.

- A GPS was constructed based on dietary and PA data. A GRS was also constructed, based on a panel of 20 SNPs. Both the weighted and the unweighted score was used for the analyses.
- The GPS was inversely associated with glucose levels.
- The GRS was associated with higher glucose levels.
- The association of the GRS with glucose levels was attenuated after interaction with the GPS.

Public Health Nutrition: page 1 of 7 doi:10.1017/S1368980015002244

Exclusive olive oil consumption has a protective effect on coronary artery disease; overview of the THISEAS study

Maria Dimitriou[1,*], Loukianos S Rallidis[2], Eirini V Theodoraki[1], Ioanna Panagiota Kalafati[1], Genovefa Kolovou[3] and George V Dedoussis[1,*]

[1]Department of Nutrition and Dietetics, School of Health Science and Education, Harokopio University, El. Venizelou 70, 17671 Athens, Greece: [2]Second Department of Cardiology, Attikon Hospital, School of Medicine, University of Athens, Athens, Greece: [3]First Cardiology Department, Onassis Cardiac Surgery Center, Athens, Greece

Submitted 25 March 2015: Final revision received 8 June 2015: Accepted 24 June 2015

Abstract

Objective: The aims of the current report are to present the demographic characteristics, clinical characteristics/biochemical indices and lifestyle habits of the population and to explore the potential association of exclusive olive oil consumption, in relation to lifestyle factors, with coronary artery disease risk.

Design: Demographic, lifestyle, dietary and biochemical variables were recorded. Logistic regression analysis was performed in order to estimate the relative risks of developing coronary artery disease.

Setting: The Hellenic study of Interactions between Single nucleotide polymorphisms and Eating in Atherosclerosis Susceptibility (THISEAS), a medical centre-based case–control study conducted in Greek adults.

Subjects: We consecutively enrolled 1221 adult patients with coronary artery disease and 1344 adult controls.

Results: A higher prevalence of the conventional established risk factors was observed in cases than in controls. Physical activity level was higher in controls (1·4 (SD 0·2) than in cases (1·3 (SD 0·3); $P<0·001$). Regarding current and ex-smokers, the case group reported almost double the pack-years of the control group (54·6 (SD 42·8) *v.* 28·3 (SD 26·3), respectively; $P<0·001$). Exclusive olive oil consumption was associated with 37 % lower likelihood of developing coronary artery disease, even after taking into account adherence to the Mediterranean diet (OR = 0·63; 95 % CI 0·42, 0·93; $P=0·02$).

Conclusions: Exclusive olive oil consumption was associated with lower risk of coronary artery disease, even after adjusting for adoption of an overall healthy dietary pattern such as the Mediterranean diet.

Keywords
CVD
Coronary artery disease
Case–control
Olive oil
Mediterranean diet

Coronary artery disease (CAD) is the most common disorder of CVD. It causes narrowing of the lumen of one or more of the coronary arteries, resulting in angina pectoris, myocardial infarction or congestive heart failure. It is estimated that 17·3 million people died from CVD in 2008 and this number contributes 30 % of all deaths worldwide. Of these deaths, 7·3 million were due to CAD[1].

In Greece, the prevalence of CVD risk factors was originally examined by the Seven Countries Study[2]. That study paved the way for other observational studies to follow, which measured the prevalence of CVD risk factors at population level[3–5]. The results of these studies have mainly highlighted that dislipidaemia, hypertension, obesity and environmental factors such as unhealthy dietary habits, smoking, physical inactivity and stress contribute to the development of CAD. Among the established environmental factors, dietary habits and their relationship to CAD have gained the attention of many scientists for more than half a century. The Seven Countries study was again the first to highlight a dietary pattern, namely the Mediterranean diet (MedDiet), as cardioprotective[6–8]. Along the same lines, other studies have consistently reported that adherence to this type of dietary pattern is beneficial against cardiovascular risk[9,10]. It has been proposed that the MedDiet exerts a beneficial influence against CAD risk factors and CAD occurrence mainly due to the abundant consumption of olive oil rich in MUFA[11]. Many studies have demonstrated an inverse relationship between olive oil consumption and the risk of CVD[12,13]. Moreover, as obesity rates are rising and lipid profiles are deteriorating within the Greek population, total fat intake, SFA intake and levels of physical activity should also be considered before advising for a Mediterranean food pattern adoption

*Corresponding authors: Email dimitrioumelina@gmail.com and dedousi@hua.gr

independently of the fat content[14]. Therefore, more evidence is needed to elucidate the association between olive oil consumption, in the context of MedDiet adoption, and CAD risk. Furthermore, even more scarce data exist regarding the effect of exclusive olive oil consumption on CAD risk.

The Hellenic study of Interactions between Single nucleotide polymorphisms and Eating in Atherosclerosis Susceptibility (THISEAS) is a medical centre-based case–control study. The aims of the current report are to: (i) present the demographic characteristics, clinical characteristics/biochemical indices and lifestyle habits of the Greek sample; and (ii) explore the potential association of exclusive olive oil consumption, in relation to lifestyle factors, with CAD risk.

Materials and methods

Study population

Study participants were consecutively recruited from (i) major hospitals, (ii) Centers of Open Protection for the Elderly and (iii) municipalities, all located in the region of Athens.

Controls were either out-patients for routine examinations or in-patients in departments other than cardiology clinics, who visited the same hospitals and during the same period as the coronary patients. Healthy volunteers from Centers of Open Protection for the Elderly and municipalities of the area of Athens were also enrolled in the control group. In order to reduce the effect of possible unknown confounders and to eliminate the problem of misclassification, we retrieved precise information from medical history through the hospital or insurance records. Therefore, controls were individuals with negative coronary angiography findings or a negative stress test, or individuals without clinical symptoms of CAD, any CVD, cancer or inflammatory diseases.

Cases were enrolled from cardiology clinics of Athens' hospitals. Cases were individuals presenting with either acute coronary syndrome or stable CAD defined as >50 % stenosis in at least one of the three main coronary vessels assessed by coronary angiography. Acute coronary syndrome was defined as acute myocardial infarction ranging from ST-segment (STEMI) to non-ST-segment elevation (NSTEMI) and unstable angina[15]. Acute coronary syndrome patients had also undergone coronary angiography that verified the presence of significant stenosis. Individuals with renal or hepatic disease were excluded from both study groups. In the group of cases, 60·4 % were diagnosed with first-time CAD at the time of recruitment, while 39·6 % were diagnosed with CAD six months or more before the time of recruitment.

All participants were informed about the goals of the study and gave their written consent. A total of 2886 people were invited to participate in the study and 2565 were finally recruited from 2006 to 2010.

The study protocol was approved by the Ethics Committee of Harokopio University.

Blood sampling and analysis

Blood samples were collected between 06.00 and 10.00 hours after a 12 h fast. Blood analysis included haematological and biochemical indices: fasting glucose, total cholesterol (TC), HDL cholesterol (HDL-C) and TAG. LDL cholesterol (LDL-C) was calculated from the Friedewald equation[16]. Furthermore, plasma and serum were isolated and stored at −80°C for future measurements.

Demographic, anthropometric and lifestyle assessments

Demographic and lifestyle characteristics, personal and family medical history of CVD risk factors, dietary habits and physical activity habits were assessed during recruitment of the study sample.

All participants were interviewed regarding their origins to ensure their Greek ancestry. Their educational status was measured by the years of schooling and was classified into four groups: (i) primary school (0–6 years); (ii) high school (7–12 years); (iii) senior high or technical institutes (13–15 years); and (iv) university (≥16 years). Senior high or technical institutes and university were combined into one group. The marital status of the participants was also recorded and defined as single, married, divorced or widowed. The annual financial status of the participants was classified into four groups: (i) low (<8000 €); (ii) medium (8000–15 000 €); (iii) high (15 000–20 000 €); and (iv) very high (>20 000 €). Low and medium financial status were combined into one group (low/medium). Similarly, high and very high financial status were combined into one group (high/very high). Furthermore, data regarding professional status was collected and the participants were accordingly characterized as public servants, private employees, freelancers, partially employed, unemployed, retired and householders.

Current smokers were defined as those who smoked at least one cigarette daily, non-smokers as those who have never smoked in their life and former smokers as those who have stopped smoking for at least six months. Pack-years (cigarette packs per day × years of smoking) were also calculated.

Anthropometric and clinical assessments

A physician performed a clinical assessment of the participants through a questionnaire during an interview. In order to eliminate recall bias, we also tried to reclaim medical information from the hospital or insurance records for both study groups, where available. Hypercholesterolaemia was defined as TC level greater than 200 mg/dl or use of lipid-lowering medication. Diabetics were those with blood glucose level greater than 125 mg/dl or individuals under a special diet or treatment.

Blood pressure level was measured in the right arm, with the participant seated and rested, using a mercury sphygmomanometer. Hypertension was defined as blood pressure greater than 140/90 mmHg or use of anti-hypertensive medication.

Family history of premature CAD was considered when the age at which first-degree relatives developed symptomatic CAD was <55 years for males and <65 years for females.

Finally, body weight and height were measured for all participants (wearing light clothing, without shoes) using a levelled platform scale and a wall-mounted stadiometer, to the nearest 0·5 kg and 0·5 cm, respectively. BMI was computed as [weight (kg)]/[height (m)]^2.

Physical activity assessment

Physical activity level (PAL) was assessed through the Harokopio Physical Activity Questionnaire (HAPAQ)[17] that evaluates the frequency, duration and intensity of occupational, household and leisure-time activities. PAL was calculated as the ratio of the work metabolic rate to the resting metabolic rate of 1 metabolic equivalent (MET). The work metabolic rate was defined as the person's total energy expenditure in a 24 h period, based on performed physical activities. The caloric cost of each physical activity (PA) was estimated with the equation: kilocalories = MET × weight in kilograms × duration in hours. The 1 MET reference value of 1 kcal/kg per h is used by convention and refers to a metabolic rate at rest of an average individual[18]. PA adoption was assessed as a categorical variable categorizing the participants into two groups, whether they reported leisure-time activities on a regular basis or not.

Dietary assessment

Nutritional information was obtained from 60·4 % (*n* 737) of patients who were first diagnosed with CAD within six months by the time of their recruitment into the study. The volunteers were asked about their nutritional habits before CAD diagnosis. Nutritional information from participants with CAD diagnosis for more than six months at the time of recruitment was not collected, in order to eliminate: (i) bias regarding possible modifications of nutritional habits after CAD diagnosis; and (ii) recall bias regarding nutritional habits before CAD diagnosis.

We tried to retrieve dietary information from double the number of controls, in order to ensure a controls-to-cases ratio of 2:1. After excluding individuals with incomplete or missing dietary data and extreme values of energy intake, the main sample with nutritional information included 499 cases and 832 controls.

Dietary data were collected through a 172-item picture-sort FFQ. Participants were asked to indicate how often they consumed various foods and beverages, as well as the portion size by comparison with photos. Daily consumption was calculated from the FFQ by multiplying the standard serving size of each food (as described by the Ministry of Health and Welfare, Supreme Scientific Health Council) by the value corresponding to each consumption frequency: never; 1–3 times/month; 1–2 times/week; 3–4 times/week; 5–6 times/week; and 1 time/d[19]. The FFQ was either self-administered or self-interviewed. In both cases, well-trained nutritional professionals were available to assist the participants during questionnaire completion.

Nutritionist Pro version 2·2 software (Axxya Systems-Nutritionist Pro, Stafford, TX, USA) was used to analyse nutritional information regarding energy, macronutrient and micronutrient intakes.

Mediterranean diet score evaluation

Adherence to the MedDiet was evaluated by calculating a MedDietScore (MDS) that is an *a priori*-defined score. The range of the MDS is between 0 and 55; higher values suggest greater adherence to the traditional Mediterranean dietary pattern that is characterized by high consumption of plant foods (non-refined cereals, vegetables and fruits) and olive oil, by moderate consumption of dairy products, poultry and fish, and by low consumption of meat and meat products[20]. The MDS was split at the median into two categories to indicate the participants with low adherence (below the median) and high adherence (above the median) to the MedDiet.

Evaluation of olive oil consumption

The FFQ included additional questions to assess the type of fat/oil used in cooking. Participants were asked to report the type of fat/oil used during the preparation of: (i) salads; (ii) cooked meals; and (iii) fried meals. Each participant could indicate more than one choice of fat/oil, namely olive oil, seed oil, corn oil, sunflower oil, butter and margarine, as depicted in the online supplementary material, Supplemental Table 1. Regarding olive oil consumption, participants were classified into two groups: (i) those who exclusively used olive oil during meal and salad preparation; and (ii) those who consumed other types of fat/oil as well.

Statistical analysis

Continuous variables are presented as mean values and standard deviations, while categorical variables are presented as frequencies. Differences between categorical variables and groups of the study were assessed using the χ^2 test. P–P plots were applied to assess the normality of the distribution of the continuous variables. Student's *t* test or the Mann–Whitney test was applied to evaluate differences in continuous variables between the two study groups. Logistic regression analysis was performed in order to estimate the relative risks of developing CAD by the calculation of odds ratios and their corresponding 95 % confidence intervals. All reported *P* values are based on two-sided tests and compared with a significance level of 5 %. The statistical software package IBM SPSS Statistics 21·0 was used for all statistical calculations.

Table 1 Sociodemographic characteristics of the study participants; THISEAS study, 2006–2010.

	CAD patients (*n* 1221)		Controls (*n* 1344)		
	Mean or frequency	SD	Mean or frequency	SD	*P* value*
Age (years)	62·8	10·4	57·7	14·1	<0·001
Male sex (%)	81·8	–	46·3	–	<0·001
Educational status (%)	10·5	4·8	11·5	4·7	<0·001
Primary school	36·7	–	28·1	–	<0·001
High school	34·2	–	34·1	–	
Tertiary education/ university	29·1	–	37·8	–	
Profession (%)					
Public servants	7·9	–	17·8	–	<0·001
Private employees	10·1	–	18·2	–	
Freelancers	19·2	–	14·6	–	
Partially employed	0·2	–	1·4	–	
Unemployed	0·3	–	4·9	–	
Retired	54·3	–	33·1	–	
Householders	8·0	–	10·0	–	
Annual income (%)					
Low/medium	44·3	–	38·1	–	<0·001
High/very high	55·7	–	61·9	–	
Marital status (%)					
Single	21·8	–	12·3	–	
Married	69·7	–	66·4	–	<0·05
Divorced	5·3	–	15·6	–	
Widow/er	3·2	–	5·8	–	

THISEAS, The Hellenic study of Interactions between Single nucleotide polymorphisms and Eating in Atherosclerosis Susceptibility; CAD, coronary artery disease.
Data are expressed as mean and standard deviation or frequency (%).
**P* values derived from Student's *t* test or the χ^2 test.

Table 2 Clinical characteristics and biochemical indices of the study participants; THISEAS study, 2006–2010

	CAD patients (*n* 1221)		Controls (*n* 1344)		
	Mean or frequency	SD	Mean or frequency	SD	*P* value*
BMI (kg/m^2)	28·3	4·2	28·8	4·9	<0·05
Hypercholesterolaemia (%)	81·8	–	70·8	–	<0·001
Lipid-lowering medication (%)	80·0	–	27·3	–	<0·001
Hypertension (%)	76·5	–	51·0	–	<0·001
Diabetes mellitus (%)	31·7	–	14·7	–	<0·001
Family history of CAD (%)	39·4	–	21·4	–	<0·001
TC (mg/dl)	178·4	41·3	214·8	41·4	<0·001
HDL-C (mg/dl)	43·8	12·4	53·6	14·4	<0·001
LDL-C (mg/dl)	109·9	39·4	136·3	36·7	<0·001
TAG (mg/dl)	145·0	91·0	124·1	71·2	<0·001
SBP (mmHg)	132·3	20·2	136·0	20·2	<0·001
DBP (mmHg)	78·9	12·0	79·5	11·2	0·2
Glucose (mg/dl)	117·7	41·7	101·4	25·8	<0·001

THISEAS, The Hellenic study of Interactions between Single nucleotide polymorphisms and Eating in Atherosclerosis Susceptibility; CAD, coronary artery disease; TC, total cholesterol; HDL-C, HDL cholesterol; LDL-C, LD cholesterol; SBP, systolic blood pressure; DBP, diastolic blood pressure.
Data are expressed as mean and standard deviation or frequency (%).
**P* values derived from Student's *t* test or the χ^2 test.

Results

A total of 999 (81·8 %) of the CAD patients were male and 222 (18·2 %) were female. On the other hand, 622 participants of the control group were male (46·3 %) and 722 (53·7 %) were female. The observed age difference between cases and controls was significant ($P = 0·001$). The mean BMI value was also different between the two study groups ($P < 0·05$).

Table 1 summarizes the sociodemographic characteristics of the study's population. The control group reported more years of education than the group of cases ($P < 0·001$). A higher proportion of controls reported high levels of educational attainment and income in the upper economic classes than CAD patients ($P < 0·001$).

The clinical characteristics and biochemical indices of the study's population are summarized in Table 2. A higher prevalence of the conventional established risk factors was observed in cases than in controls. Interestingly, a high proportion of controls had TC over 200 mg/dl with only 27·3 % being treated with lipid-lowering agents. On the other hand, 80 % of CAD patients were receiving hypolipidaemic medication.

Lifestyle characteristics are presented in Table 3. PAL was higher in controls when compared with cases ($P < 0·001$). A higher prevalence of smoking was recorded in CAD patients than in controls, with patients reporting almost double the pack-years of the control group ($P < 0·001$). Concerning diet, the higher proportion of fat intake observed in the control group compared with cases was derived from the higher MUFA intake in the former study group ($P = 0·000$). Exclusive use of olive oil in cooking was reported by a higher percentage of controls (11 %) compared with cases ($P < 0·001$).

After adjusting for conventional CAD risk factors, exclusive olive oil consumption was associated with 48 % lower likelihood of developing CAD (OR = 0·52; 95 % CI 0·35, 0·77; $P = 0·001$), as depicted in Table 4. The previous finding was confirmed even after taking into account PA adoption and adherence to the MedDiet (OR = 0·63; 95 % CI 0·42, 0·93; $P = 0·02$).

Discussion

The current work presents the design and descriptive characteristics of the THISEAS study. According to the study results, a higher prevalence of the established risk factors was observed in the group of cases. Higher values and above the normal range for TC and HDL-C were reported within the control group compared with cases. These results can be explained by the fact that most CAD patients were receiving lipid-lowering agents, mainly statins, while the majority of controls with TC values above 200 mg/dl were not receiving hypolipidaemic medication.

Moreover, almost 71 % and 51 % of the control group had hypercholesterolaemia and hypertension, respectively. The prevalence of hypercholesterolaemia within the control group is higher than previously reported in the Greek

Table 3 Lifestyle characteristics of the study participants; THISEAS study, 2006–2010

	CAD patients (*n* 499)		Controls (*n* 832)		
	Mean or frequency	SD	Mean or frequency	SD	*P* value*
Current and ex-smokers (%)	79·3	–	50·6	–	<0·001
Pack-years	54·6	42·8	28·3	26·3	<0·001
PAL	1·3	0·3	1·4	0·2	<0·001
Energy intake (kJ)	8820	3588	7665	3286	<0·001
Energy intake (kcal)	2108·0	857·6	1831·9	785·3	<0·001
Protein intake (%E)	17·5	3·1	19·7	4·3	<0·001
Carbohydrate intake (%E)	45·8	8·3	44·3	9·6	<0·005
Fat intake (%E)	35·0	6·3	36·0	6·3	<0·005
SFA (%E)	11·6	3·5	11·7	3·1	0·7
MUFA (%E)	15·4	3·2	16·4	3·8	<0·001
PUFA (%E)	5·1	1·3	4·9	1·5	0·1
Exclusive olive oil consumption (%)	65·1	–	76·0	–	0·001

THISEAS, The Hellenic study of Interactions between Single nucleotide polymorphisms and Eating in Atherosclerosis Susceptibility; CAD, coronary artery disease; PAL, physical activity level; %E, percentage of energy intake. Data are expressed as mean and standard deviation or frequency (%).
**P* values derived from Student's *t* test or the χ^2 test.

Table 4 Results from the logistic regression models for the evaluation of exclusive olive oil consumption on the risk of developing CAD in 259 CAD patients and 679 controls; THISEAS study, 2006–2010

	OR	95 % CI	*P* value
Model 1*	0·52	0·35, 0·77	0·001
Model 2†	0·62	0·42, 0·92	0·02
Model 3‡	0·63	0·42, 0·93	0·02

CAD, coronary artery disease; THISEAS, The Hellenic study of Interactions between Single nucleotide polymorphisms and Eating in Atherosclerosis Susceptibility.
*Model 1 included age, male sex, BMI, hypercholesterolaemia, hypertension, diabetes mellitus, family history of CAD and pack-years.
†Model 2 included model 1 and physical activity adoption.
‡Model 3 included model 2 and adherence to the Mediterranean diet.

population (ATTICA study), while the results regarding the prevalence of hypertension are similar to those reported previously[5]. However, the study sample in the THISEAS study is a decade older than the study aforementioned; therefore the results are not totally comparable.

In industrialized countries, socio-economic status is inversely related with CAD risk factors and CAD incidence, thus CAD mortality and morbidity is observed more among the less educated and less affluent[21,22]. Previous studies such as the Minnesota Heart Survey and the ATTICA study have demonstrated that low socio-economic status can increase CAD risk factors or can predict unhealthy risk patterns[23,24]. Along the same lines, our results show that the control group reported higher education levels and higher financial status when compared with CAD patients.

PA is associated with lower incidence of CAD[25]. In our study, CAD patients recorded lower PAL than controls. According to PAL values, CAD patients could be described as sedentary while controls could be described as physically active at low levels. Lower levels of PA in CAD patients than controls have also been reported in previous studies[4,26]. These studies have also demonstrated a higher prevalence of current smoking and a higher number of pack-years in cases than controls, in line with our results.

Regarding dietary fat intake in both groups, the mean daily intakes of SFA and PUFA, as a percentage of total energy intake, are similar to those demonstrated for Greece by the EPIC (European Prospective Investigation into Cancer and Nutrition) study[27]. The proportion of total fat intake in the present study is lower than previously reported in the EPIC study for the Greek population but closer to the proposed guidelines[27,28].

Olive oil is considered an important component of the MedDiet pattern and many studies have highlighted its inverse association with CAD occurrence[29,30]. The Greek population traditionally uses olive oil during meal and salad preparation. The beneficial effects of olive oil on CAD have been attributed to oleic acid and also to minor components, such as flavonoids, squalene and phenolic compounds[31]. The proposed mechanisms through which olive oil exerts its beneficial effects on CAD involve a decrease of TC and LDL-C, a reduction of LDL oxidation, an improvement of endothelial function and a decrease in thrombosis[11].

In the present study, in addition to the well-established CAD risk factors, it is observed that the effect of exclusive olive oil consumption seems to play a significant role in CAD risk. Specifically, participants who reported exclusive use of olive oil during the preparation of meals and salads had 37 % lower risk of developing CAD, regardless of their adherence to the MedDiet. This finding is similar to the results of prior studies conducted in the Greek population, namely the CARDIO2000 and the ATTICA studies[5,30].

More recent prospective studies have shown that high olive oil consumption is inversely associated with CAD risk. In the EPICOR study, a reduction in CAD risk was observed among Italian women in the highest quartile of olive oil consumption during an approximately 8-year follow-up (hazard ratio = 0·56, 95 % CI 0·31, 0·99; $P = 0·04$)[32]. In the EPIC-Spain study, consumers in the highest quartiles of olive oil consumption had a 44 % lower risk in CVD mortality compared with non-consumers[33]. The beneficial impact of olive oil consumption against CVD occurrence was further supported by the PREDIMED study, where a MedDiet supplemented with extra-virgin olive oil was associated with a reduction in CVD events (hazard ratio = 0·70; CI 0·54, 0·92)[12]. Another study that examined olive oil consumption in the context of a MedDiet is the Three-City Study. That study's results revealed a lower stroke risk in high olive oil users, in elderly subjects. In agreement with our results, this outcome suggests that olive oil can be protective against disease regardless of other dietary elements[34].

Although this finding regarding exclusive olive oil consumption may suggest an important approach to preventive

nutrition, it does not suggest evidence for causality. Furthermore, it is worth mentioning that the quantity of olive oil intake was not assessed in the present work and there was no differentiation among different varieties of olive oil (common, virgin or extra-virgin olive oil).

Limitations of our case–control study are the selection and recall bias. A small percentage of future CAD patients may be wrongly assigned to the control group and recall bias may still exist regarding dietary information, smoking and PA habits. Nevertheless, we tried to eliminate both systematic errors as mentioned above (see 'Materials and methods' section).

Another important limitation is that there were differences in age, sex ratio and BMI between the two study groups resulting in an unmatched case–control study. However, we tried to control for these confounding factors during the logistic regression analysis of the study.

In addition, waist circumference data are not shown in our results due to the high number of missing values from CAD patients who were hospitalized and resting in bed during recruitment. Waist circumference is an independent predictor of CVD risk in overweight individuals[35,36] and is in clinical utility in an attempt to track changes during lifestyle modifications even if BMI is not affected.

The main outcome of the present report is the finding that exclusive olive oil consumption in food and salad preparation can reduce the risk of CVD, even after controlling for potential confounders. Although many studies have demonstrated the beneficial effect of olive oil with regard to the variety or quantity consumed, not many studies have elucidated the effect of exclusive olive oil consumption. If this outcome is further confirmed, it suggests a simple, feasible and easily understood dietary recommendation for primary prevention of CVD occurrence.

Conclusions

The presented data provided the demographic, clinical and lifestyle characteristics of the THISEAS study. A higher prevalence of the established risk factors was reported among CAD patients. Exclusive olive oil consumption in cooking seems to lower the odds of CAD risk. This association remained significant even after taking into account CAD risk factors and lifestyle choices (PA and diet). However, further investigation is needed in larger sample sizes in order to elucidate the impact of exclusive olive oil consumption on CAD risk.

Acknowledgements

Acknowledgements: The authors thank all the dietitians and clinicians for their contribution to the project. *Financial support:* This work was partially supported by a research grant (PENED 2003) from the Greek General Secretary of Research and Technology. The Greek General Secretary of Research and Technology had no role in the design, analysis or writing of this article. *Conflict of interest:* None. *Authorship:* M.D. participated in the data collection, carried out the data manipulation, analysis and interpretation, performed the statistical analysis and drafted the paper. L.S.R. contributed to the data collection and critically reviewed the paper. E.V.T. participated in the design of the study, participated in the data collection and critically reviewed the paper. I.P.K. participated in the data collection, data manipulation and critically reviewed the paper. G.K. contributed to the data collection and critically reviewed the paper. G.V.D. carried out the study design, supervised and coordinated the study and drafted the paper. All authors have approved the final article. *Ethics of human subject participation:* The study protocol was approved by the Ethics Committee of Harokopio University.

Supplementary material

To view supplementary material for this article, please visit http://dx.doi.org/10.1017/S1368980015002244

References

1. World Health Organization (2011) *Global Atlas on Cardiovascular Disease Prevention and Control*. Geneva: WHO.
2. Aravanis C, Concordilas A, Dontas AS *et al.* (1970) Coronary heart disease in seven countries. IX. The Greek islands of Crete and Corfu. *Circulation* **4**, 88–100.
3. Moulopoulos SD, Adamopoulos PN, Diamantopoulos EI *et al.* (1987) Coronary heart disease risk factors in a random sample of Athenian adults. The Athens Study. *Am J Epidemiol* **126**, 882–892.
4. Panagiotakos DB, Pitsavos C, Chrysohoou C *et al.* (2002) Risk stratification of coronary heart disease in Greece: final results from the CARDIO2000 Epidemiological Study. *Prev Med* **35**, 548–556.
5. Panagiotakos DB, Pitsavos C, Chrysohoou C *et al.* (2008) Five-year incidence of cardiovascular disease and its predictors in Greece: the ATTICA study. *Vasc Med* **13**, 113–121.
6. Keys A (editor) (1970) Coronary heart disease in seven countries. *Circulation* 41, 1–211.
7. Keys A, Menotti A, Aravanis C *et al.* (1984) The seven countries study: 2,289 deaths in 15 years. *Prev Med* **13**, 141–154.
8. Menotti A, Keys A, Aravanis C *et al.* (1989) Seven Countries Study. First 20-year mortality data in 12 cohorts of six countries. *Ann Med* **21**, 175–179.
9. de Lorgeril M, Salen P, Martin JL *et al.* (1999) Mediterranean diet, traditional risk factors, and the rate of cardiovascular complications after myocardial infarction: final report of the Lyon Diet Heart Study. *Circulation* **16**, 779–785.
10. Sofi F, Abbate R, Gensini GF *et al.* (2010) Accruing evidence on benefits of adherence to the Mediterranean diet on health: an updated systematic review and meta-analysis. *Am J Clin Nutr* **92**, 1189–1196.
11. López-Miranda, Perez-Jimenez F, Ros E *et al.* (2010) Olive oil and health: summary of the II international conference on olive oil and health consensus report, Jaén and Córdoba (Spain) 2008. *Nutr Metab Cardiovasc Dis* **20**, 284–294.

12. Estruch R, Ros E, Salas-Salvadó J *et al.* (2013) Primary prevention of cardiovascular disease with a Mediterranean diet. *N Engl J Med* **4**, 1279–1290.
13. Huang CL & Sumpio BE (2008) Olive oil, the Mediterranean diet, and cardiovascular health. *J Am Coll Surg* **207**, 407–416.
14. Ferro-Luzzi A, James WP & Kafatos A (2002) The high-fat Greek diet: a recipe for all? *Eur J Clin Nutr* **56**, 796–809.
15. Kalra S, Duggal S, Valdez G *et al.* (2008) Review of acute coronary syndrome diagnosis and management. *Postgrad Med* **120**, 18–27.
16. Friedewald WT, Levi RI & Fredrickson DS (1972) Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clin Chem* **18**, 499–502.
17. Maraki M, Kollia A, Gioxari A *et al.* (2010) Development, reliability and validity of the Athens Physical Activity Questionnaire. *Clin Nutr Suppl* **5**, 80.
18. Ainsworth BE, Haskell WL, Herrmann SD *et al.* (2011) 2011 Compendium of physical activities: a second update of codes and MET values. *Med Sci Sports Exerc* **43**, 1575–1581.
19. Ministry of Health and Welfare (1999) Dietary guidelines for adults in Greece. *Arch Hellenic Med* **16**, 516–524.
20. Panagiotakos DB, Pitsavos C, Arvaniti F *et al.* (2007) Adherence to the Mediterranean food pattern predicts the prevalence of hypertension, hypercholesterolemia, diabetes and obesity, among healthy adults; the accuracy of the MedDietScore. *Prev Med* **44**, 335–340.
21. Marmot M (1989) Socioeconomic determinants of CHD mortality. *Int J Epidemiol* **18**, 196–202.
22. Mackenbach JP, Kunst A, Cavelaars AE *et al.* (1997) Socioeconomic inequalities in morbidity and mortality in Western Europe. The EU Working Group on Socioeconomic Inequalities in Health. *Lancet* **349**, 1655–1659.
23. Luepker RV, Rosamond W, Murphy R *et al.* (1993) Socioeconomic status and coronary heart disease risk factor trends. The Minnesota Heart Survey. *Circulation* **88**, 2172–2179.
24. Panagiotakos DB, Pitsavos C, Chrysohoou C *et al.* (2008) The effect of clinical characteristics and dietary habits on the relationship between education status and 5-year incidence of cardiovascular disease: the ATTICA study. *Eur J Nutr* **47**, 258–265.
25. Thompson PD, Buchner D, Piña IL *et al.* (2003) AHA Scientific Statement. Exercise and physical activity in the prevention and treatment of atherosclerotic cardiovascular disease. *Circulation* **107**, 3109–3116.
26. Gustavsson J, Mehling K, Leander K *et al.* (2015) FTO gene variation, macronutrient intake and coronary heart disease risk: a gene–diet interaction analysis. *Eur J Nutr* (Epublication ahead of print version).
27. Linseisen J, Welch AA, Ocké M *et al.* (2009) Dietary fat intake in the European Prospective Investigation into Cancer and Nutrition: results from the 24-h dietary recalls. *Eur J Clin Nutr* **63**, Suppl. 4, S61–S80.
28. International Task Force for Prevention of Coronary Heart Disease (2003) *Pocket Guide to Prevention of Coronary Heart Disease*. Grünwald: Börm Bruckmeier.
29. Fernández-Jarne E, Martinez-Losa E, Prado-Santamaría M *et al.* (2002) Risk of first non-fatal myocardial infarction negatively associated with olive oil consumption: a case–control study in Spain. *Int J Epidemiol* **31**, 474–480.
30. Kontogianni MD, Panagiotakos DB, Chrysohoou C *et al.* (2007) The impact of olive oil consumption pattern on the risk of acute coronary syndromes: the CARDIO2000 case–control study. *Clin Cardiol* **30**, 125–129.
31. Waterman E & Lockwood B (2007) Active components and clinical applications of olive oil. *Altern Med Rev* **12**, 331–342.
32. Bendinelli B, Masala G, Saieva C *et al.* (2011) Fruit, vegetables, and olive oil and risk of coronary heart disease in Italian women: the EPICOR Study. *Am J Clin Nutr* **93**, 275–283.
33. Buckland G, Mayén AL, Agudo A *et al.* (2012) Olive oil intake and mortality within the Spanish population (EPIC-Spain). *Am J Clin Nutr* **96**, 142–149.
34. Samieri C, Féart C, Proust-Lima C *et al.* (2011) Olive oil consumption, plasma oleic acid, and stroke incidence: the Three-City Study. *Neurology* **77**, 418–425.
35. Lofgren I, Herron K, Zern T *et al.* (2004) Waist circumference is a better predictor than body mass index of coronary heart disease risk in overweight premenopausal women. *J Nutr* **134**, 1071–1076.
36. Freiberg MS, Pencina MJ, D'Agostino RB *et al.* (2008) BMI vs. waist circumference for identifying vascular risk. *Obesity (Silver Spring)* **16**, 463–469.

Nutrition, Metabolism & Cardiovascular Diseases (2016) **26**, 201–206

Available online at www.sciencedirect.com

Nutrition, Metabolism & Cardiovascular Diseases

journal homepage: www.elsevier.com/locate/nmcd

Lifestyle may modify the glucose-raising effect of genetic loci. A study in the Greek population

E. Marouli [a,b], S. Kanoni [c], M. Dimitriou [a], G. Kolovou [d], P. Deloukas [c,e], G. Dedoussis [a,*]

[a] Department of Nutrition and Dietetics, School of Health Science and Education, Harokopio University, Athens, Greece
[b] National and Kapodistrian University of Athens, Faculty of Biology, Athens, Greece
[c] William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK
[d] Department of Cardiology, Onassis Cardiac Surgery Center, Athens, Greece
[e] Princess Al-Jawhara Al-Brahim Centre of Excellence in Research of Hereditary Disorders (PACER-HD), King Abdulaziz University, Jeddah 21589, Saudi Arabia

Received 16 July 2015; received in revised form 2 October 2015; accepted 13 October 2015
Available online 3 November 2015

KEYWORDS
Glucose levels;
Genetic risk score;
Lifestyle;
Interaction

Abstract *Background and aims:* Lifestyle habits including dietary intake and physical activity are closely associated with multiple body processes including glucose metabolism and are known to affect human health. Recent genome-wide association studies have identified several single nucleotide polymorphisms (SNPs) associated with glucose levels. The hypothesis tested here is whether a healthy lifestyle assessed via a score is associated with glycaemic traits and whether there is an interaction between the lifestyle and known glucose-raising genetic variants in association with glycaemic traits.
Methods and results: Participants of Greek descent from the THISEAS study were included in this analysis. We developed a glucose preventive score (GPS) including dietary and physical activity characteristics. We also modelled a weighted genetic risk score (wGRS), based on 20 known glucose-raising loci, in order to investigate the impact of lifestyle–gene interaction on glucose levels. The GPS was observed to be significantly associated with lower glucose concentrations ($\beta \pm SE$: -0.083 ± 0.021 mmol/L, $P = 1.6 \times 10^{-04}$) and the wGRS, as expected, with increased glucose levels ($\beta \pm SE$: 0.020 ± 0.007 mmol/L, $P = 8.4 \times 10^{-3}$). The association of the wGRS with glucose levels was attenuated after interaction with the GPS. A higher GPS indicated decreasing glucose levels in the presence of an increasing wGRS ($\beta_{interaction} \pm SE$: -0.019 ± 0.007 mmol/L, $P = 0.014$).
Conclusion: Our results indicate that lower glucose levels underlie a healthier lifestyle and also support an interaction between the wGRS for known glycaemic loci and GPS associated with lower glucose levels. These scores could be useful tools for monitoring glucose metabolism.

Introduction

Glucose homeostasis deficiency may lead to chronic increase in blood glucose levels and could affect a large number of tissues and organs. The shift from early metabolic abnormalities that forego diabetes, such as impaired fasting glucose and impaired glucose tolerance, to diabetes is not direct. However, current evidence indicates that most individuals with pre-diabetic states finally develop diabetes [1–3]. Cardiovascular disease risk moderately increases during the pre-diabetic state [4], but it further increases with the development of diabetes, along with long-term complications affecting

* Corresponding author. Tel.: +30 210 9549 304; fax: +30 210 9577 050.
E-mail address: dedousi@hua.gr (G. Dedoussis).

http://dx.doi.org/10.1016/j.numecd.2015.10.003
0939-4753/

eyes, kidneys and nervous system. As the medical socioeconomic strain of type 2 diabetes mellitus is increased by its complications and highlights the burden on healthcare systems [5]. The number of single nucleotide polymorphisms (SNPs) associated with fasting glucose concentration levels has now increased to 36 [6]. A combination of genetic and environmental factors contributes to impaired glucose homeostasis [7]. A large number of genome-wide significant genetic loci associated with glycaemic traits [6] and type 2 diabetes development [8] have been recently identified. In addition to genetic factors involved in impaired glucose homeostasis, lifestyle patterns including dietary intake play a significant role in the pathogenic process [7,9–11]. Glycaemic control strategies include weight management as the primary nutritional strategy, accompanied by moderate physical activity [12,13]. Evaluation of a dietary pattern (rather than single nutrients) and exercise status have been used in nutritional epidemiology and could provide a direct approach to the quantification of disease prevention [14,15]. The optimal prevention of diabetes requires identification of its modifiable risk factors to be targeted for intervention. Investigation of the interactions between genetic variants and environment has helped elucidate the biological basis of diabetes mellitus and promote individualised health-promoting lifestyle recommendations [16]. Information on personal genetic profile and lifestyle components are touted for potential contribution to personalised medicine [17]. It is of great interest for clinicians and other health-related professionals to consider the impact of diet and physical activity on modification of glucose levels in individuals with increased genetic predisposition.

In the present study, we sought to 1) evaluate the association of a healthy lifestyle pattern in the Greek population, by means of a preventive score, with glycaemic and adiposity traits and 2) evaluate whether this lifestyle pattern modifies the association of known glucose-raising genetic variants on glycaemic traits.

Methods

Study population

Our sample consisted of unrelated individuals of Greek origin, aged 57.7 ± 14.1 years, drawn from the THISEAS (The Hellenic Study of Interactions between SNPs and Eating in Atherosclerosis Susceptibility) [18]. Individuals with type 2 diabetes (medical history or fasting glucose levels >7 mmol/L) and outliers with respect to energy intake were excluded (Supplementary Methods). Information about genotyping, the description of adiposity and biochemical measurements, as well as the assessment of dietary patterns and physical activity are provided in the Supplementary Material (Supplementary Methods).

Modelling of glucose preventive score (GPS) and weighted genetic risk score (wGRS)

Based on selected dietary and physical activity data, a 'Glucose Preventive Score' for each volunteer was calculated. The components used for the score showed a positive or negative association with glucose levels and are supported in literature (Supplementary Methods). The lifestyle parameters used for the score included three with glucose-lowering association (hours in movement during work per day, vegetable consumption (servings per day) and fruits and fresh juice; servings/day) and one with glucose-raising association (consumption of soft drinks and beverages with sugar; servings per day). Daily servings of the food groups were estimated in our sample as the sum of daily servings of each item included on the food frequency questionnaire [19]. The lifestyle variables were categorised in tertiles and each tertile received a point. The first three components were assigned increasing points per tertile and the last one decreasing points. The score was the sum of points for all components per individual. The score ranged from 0- to 10 points, with an increasing glucose-lowering effect (Supplemental Table 1).

In order to evaluate the cumulative association of known glucose-raising genetic variants with glucose metabolism, we constructed a weighted genetic risk score (wGRS) and an unweighted genetic risk score (GRS). We included the published sentinel SNPs from 20 glycaemic-related loci, which were identified from the MAGIC (the Meta-Analyses of Glucose and Insulin-related traits Consortium) MetaboChip Meta-analysis effort [6]. For each of the 20 SNPs, individuals carrying 0, 1 or 2 glucose-raising alleles received 0, 1 or 2 points, respectively, for the GRS estimation. The wGRS was calculated as the sum of points across the 20 SNPs weighted by their published effect sizes [6].

We then divided the score by the average effect size of all SNPs so that it is rescaled to represent the range of the possible number of weighted glucose-increasing alleles for each individual [20].

The wGRS score was split into quantiles. Individuals in the last quantile (wGRS > 25 points) were classified as high risk (12.7% of the total sample) (Supplementary Methods).

Statistical methods and analysis

Statistical methods used for the analyses are described in detail in the Supplementary Methods. Association analysis was performed using PLINK [21] and R version 3.1.1. Continuous variables are presented as mean ± standard deviation (SD) or median ± interquartile range and categorical as relative frequencies. Natural log-transformed values of insulin levels were used. The reported *p*-values were based on two-sided tests. Linear regression models assuming an additive genetic model were used to test the association of the 20 SNPs on glucose levels. Linear regression models were applied to test for the associations between each lifestyle variable and glucose levels as well

as the association of the GPS on glycaemic and adiposity traits. For our interaction associations, we included a GPS × wGRS cross product along with the covariates (GPS, wGRS, age, sex, body mass index (BMI) and energy intake) and the glucose levels as an outcome. In order to further characterise the direction of the interaction, we conducted a stratified analysis. For the lifestyle score and the wGRS, the coefficient of determination (R [2]) was used as a measure to express the proportion of total variation explained by the model. Genotype by environment (G × E) variance contribution to the total variance of glucose levels was estimated using genome-wide complex trait analysis (GCTA) [22]. We used Quanto v1.2.4 for power calculations (http://hydra.usc.edu/gxe/) (Supplementary Methods).

Results

Descriptive characteristics of the study cohort are given in Supplemental Table S2. The association of each component of the GPS with glucose levels was investigated. Most of the selected lifestyle variables were significantly associated with glucose levels after controlling for age, sex, BMI and total energy (Supplementary Methods; Supplemental Table S3). Dietary food intake variables including fruit/fresh juice consumption (servings/day) and vegetable consumption (servings/day) indicated a significant negative association with glucose levels (Supplemental Table S3). As expected, consumption of soft drinks and beverages with sugar had a significant positive association with glucose levels (Supplemental Table S3).

The mean GPS and wGRS in THISEAS was 3.729 ± 1.351 points and 22.285 ± 2.351 points, respectively. There was a significant difference between men and women for the moving hours/day (higher mean value in men, $P = 3.2 \times 10^{-6}$). Women reported higher consumption of fresh fruits and juices ($P = 5.5 \times 10^{-7}$) compared to men (Supplemental Table S2).

GPS was associated with overall lower glucose levels, showing reduction by 0.083 units (mmol/L) per increasing point of the preventive score; β ± SE: −0.083 ± 0.021, $P = 1.6 \times 10^{-04}$ (Table 1, Supplemental Fig. S1). GPS explained 1.17% of the glucose level variation.

Most of the previously published 20 genetic variants for glucose levels (Supplemental Table S4) [6] also showed evidence for association in our study. The strongest association was observed at the PRKAR2A locus (β ± SE: 0.074 ± 0.027 mmol/L, $P = 0.007$) (Supplemental Table S4). Variants in PRKAR2A, PDX1, CDKAL1 and KL locus were nominally associated with glucose levels, after adjusting for age and sex. The majority of the investigated variants (15 out of 20) had consistent direction of effect between the THISEAS and the MAGIC [6] meta-analysis (binomial sign test $P = 0.041$) (Supplemental Table S4). Significant associations were observed after adjusting for age, sex and BMI for PRKAR2A and CDKAL1 loci (Supplemental Table S5).

The wGRS was significantly associated with higher glucose levels as expected. For each increment point (weighted risk allele) in an individual's wGRS, fasting glucose concentrations increased by 0.020 mmol/L: β ± SE: 0.020 ± 0.007, $P = 8.4 \times 10^{-3}$ (Table 2A). The unweighted GRS was also significantly associated with glucose levels: β ± SE: 0.018 ± 0.007, $P = 0.011$ (Table 2B). Fig. 1 illustrates the increase in the mean fasting glucose levels across individuals carrying an increasing number of glucose-raising alleles, weighted by their published effect size, in our cohort. The positive correlation between the wGRS and glucose levels (mmol/L) in diabetes-free individuals after controlling for age, sex and BMI is also presented as a scatterplot (Supplemental Fig. S2). The difference in the mean glucose levels between subjects with higher wGRS (>25 weighted alleles) and those with the lower wGRS (<18 weighted alleles) was 0.200 mmol/L ($P = 0.01$). The association between wGRS and GRS with glucose levels was also significant in the subsample used for the interaction analysis (Supplementary Table S6).

We then investigated the impact of interaction of GPS with wGRS on blood glucose levels. Our results indicate an attenuating impact of the wGRS in interaction with the GPS on glucose concentration ($\beta_{interaction}$ ± SE:

Table 1 Associations of glucose preventive score with glycaemic traits.

	Beta[a]	SE[a]	P	N[b]
Glycaemic traits				
Glucose (mmol/L)	−0.083	0.021	1.6×10^{-04}	552
Insulin (ln-pmol/L)	0.031	0.029	0.298	243
HOMA.IR	−0.006	0.084	0.938	240

Regression models for the association of the GPS with glycaemic traits, adjusted for age, sex, BMI and total energy intake.
[a] Beta coefficient and standard error for the estimated difference in glycaemic traits per 1-unit increase in GPS.
[b] N indicates the sample size.

Table 2 Associations[a] of the genetic risk scores with glycaemic traits.

A. Associations[a] of the weighted genetic risk score with glycaemic traits.

	Beta[b]	SE[b]	P	N[c]
Glycaemic traits				
Glucose (mmol/L)	0.020	0.007	8.4×10^{-03}	1132
Insulin (ln-pmol/L)	−0.003	0.009	0.7	586
HOMA.IR	−0.003	0.024	0.9	582

B. Associations[a] of the unweighted genetic risk score with glycaemic traits

	Beta[b]	SE[b]	P	N[c]
Glycaemic Traits				
Glucose (mmol/L)	0.018	0.007	0.011	1132
Insulin (ln-pmol/L)	−0.006	0.008	0.439	586
HOMA.IR	−0.011	0.022	0.588	582

[a] Adjusted for age, sex and BMI.
[b] Beta coefficient and standard error for the estimated difference in glucose (mmol/L), fasting insulin (ln-pmol/L) concentration or HOMA.IR values per 1-unit increase in the weighted genetic risk score (wGRS).
[c] N indicates the sample size.

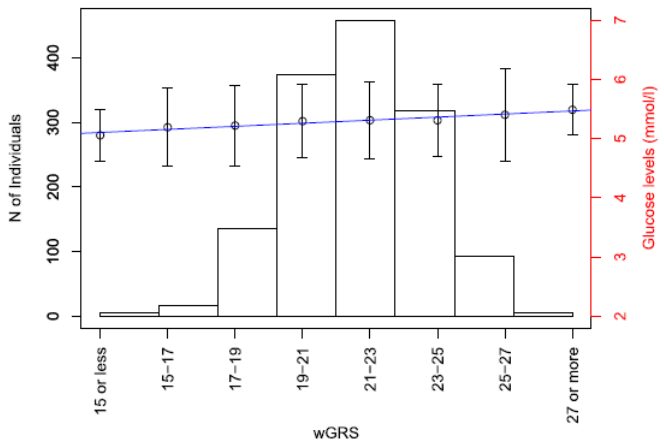


Figure 1 Glucose level increment by increasing the number of known glucose-increasing loci, weighted by the published effect size. The bar plots show the average and standard error of glucose in mmol/L for each genotype score group (right Y-axis). The histogram denotes the distribution of individuals in every genotype score group (left Y-axis).

−0.019 ± 0.007 mmol/L, $P = 0.014$). The estimated interaction regression coefficient indicated that a higher preventive score was associated with a decrease in glucose levels in the presence of an increasing genetic risk score in individuals. Our study reported a 0.019 mmol/L estimated decrease in fasting glucose concentration per 1 point increase of the wGRS combined with 1 point increase in the GPS. (Table 3A). An attenuating impact of the unweighted GRS in interaction with the GPS on glucose levels was also observed ($\beta_{interaction}$ ± SE: −0.015 ± 0.007 mmol/L, $P = 0.036$) (Table 3B).

In order to further investigate the GPS and wGRS interactions, stratified analyses were performed. Participants were stratified in tertiles based on their GPSs, and association tests of the wGRS with glucose levels were carried out within each GPS group. The association between the wGRS and glucose levels was stronger in the lower GPS group (β ± SE: 0.035 ± 0.019, $P = 0.068$) compared to the higher GPS group (β ± SE: −0.043 ± 0.021, $P = 0.043$) (Table 4).

GCTA [22] was used to assess the genetic variance contribution of the 20 known glucose SNPs (included in the GPS) to the glucose level variance. When the G × E was not included in the model, the additive genetic heritability for glucose levels was 1.1%. The heritability explained by G × E interaction was 1.8% and the total heritability explained was 2.9%. G × E heritability was calculated as the G × E variance divided by the total phenotypic variance after adjusting for age, sex and BMI.

Table 3 Interaction analysis of the genetic risk scores and GPS on glucose levels.

A. Interaction analysis of the weighted genetic risk score and GPS on glucose levels

	Beta[a]	SE	*P*	*N*[b]
Glycaemic trait Glucose (mmol/L)	−0.019	0.007	0.014	533

B. Interaction analysis of the unweighted genetic risk score and GPS on glucose levels

	Beta[a]	SE	*P*	*N*[b]
Glycaemic trait Glucose (mmol/L)	−0.015	0.007	0.036	533

Abbreviations: wGRS: weighted genetic risk score, GPS: glucose preventive score.
Regression model for interaction analysis of the GPS and wGRS on the glucose levels. Beta coefficient and standard error for the estimated difference in glucose (mmol/L), per 1-unit increase in the wGRS, assuming the additive genetic model, interacting with a 1-point increase in the lifestyle score.
[a] Results adjusted for age, sex, BMI and total energy intake.
[b] *N* indicates the sample size.

Discussion

Using data from diabetes-free participants of the THISEAS, we observed favourable associations between adherence to a healthy lifestyle and lower glucose levels and adiposity indices. We found that fruit and fresh juice consumption and vegetable consumption were significantly associated with lower glucose levels. Physical activity parameters showed a decreasing trend in glucose levels, while soft drinks and beverages were associated with increased glucose levels. Our results support the lowering impact of work and home-related physical activity on blood glucose levels.

Healthy lifestyle was represented by the GPS, including both physical activity and dietary parameters. Our study demonstrated a negative association between the GPS and glucose levels in our study. These findings could be useful for promoting healthy eating patterns and physical activity and also for glucose homeostasis control.

As expected, the derived wGRS was found to be associated with glucose concentration in THISEAS. The effect size of wGRS was approximately 0.020 mmol/L per point increment of wGRS, and calculation of a weighted genetic risk score based on the 20 genome-wide significant SNPs could be useful for the identification of at-risk individuals.

We tested the hypothesis that a healthy lifestyle could modify the cumulative impact of glucose-raising variants by performing an interaction association analysis. Our results indicate the impact of a significant gene–lifestyle interaction on glucose levels in diabetes-free individuals. A glucose risk-allele carrier who adheres to a healthier lifestyle is likely to have lower glucose levels compared to a risk-allele carrier with a less healthy lifestyle ($\beta_{interaction}$ ± SE: −0.019 ± 0.007 mmol/L, $P = 0.014$).

Stratified analysis was performed to further characterise the direction of the interaction between the GPS and the wGRS. When participants were stratified according to their GPS, a change in the direction of the wGRS association with glucose levels was observed. The association of the wGRS with glucose levels was stronger in the lower GPS group compared to the higher GPS group, where the opposite direction for the association with glucose levels was observed. Adoption of a healthier lifestyle, as reflected by the different strata of GPS used, attenuates the genetic predisposition for increasing glucose levels.

Table 4 Stratified associations of the weighted genetic risk score with glycaemic traits per GPS tertile.

Glycaemic trait	GPS tertile											
	<3 points				3–5 points				>5 points			
	Beta[a]	SE[a]	*P*	*N*[b]	Beta[a]	SE[a]	*P*	*N*[b]	Beta[a]	SE[a]	P	N[b]
Glucose (mmol/L)	0.035	0.019	0.068	104	0.007	0.014	0.602	287	−0.043	0.021	0.043	142

Regression models for the association of the wGRS with glucose levels, adjusted for age and sex.
[a] Beta coefficient and standard error for the estimated difference in glucose (mmol/L) per 1-unit increase in the weighted genetic risk score (wGRS) for each GPS tertile.
[b] *N* indicates the sample size.

GCTA was used to explore the G × E variance contribution of the environmental preventive score to the variation in glucose levels. Our results indicate that the GPS affects the genetic susceptibility of increasing glucose levels. The combination of the genetic additive variance and G × E variance of the GPS explained more variation for glucose than genetic effects alone. These data indicate the significance of the G × E variance of the preventive score for glucose levels and also suggest the magnitude to which G × E interactions contribute to the variation in the trait. These findings strongly support the idea of lifestyle intervention in individuals with a genetic predisposition for increasing glucose levels.

The study of interactions is of great importance for improving the accuracy of the assessment of both environmental and genetic impact; it facilitates prediction and provides customised recommendations for disease prevention [23]. Although many publications detected gene–environment interactions, very few have been replicated [24,25]. There is an immense need to develop prevention strategies for diabetes, particularly focussing on intensive lifestyle changes which could be cost-effective [26]. Lifestyle modification via physical activity and dietary intake could improve glycaemic control and it is the most preferable first-line measure for the management of glucose levels.

Some caveats could appear in the interpretation of our data. Limitations of the study include bias in the assessment of lifestyle and dietary measurements, which is common in other similar studies. In addition, the magnitude of the interaction between the lifestyle preventive score and the genetic risk score was moderately significant; this could be attributed to the insufficient sample size and power. Furthermore, the selected loci included for the wGRS calculation explain a fraction of glucose level variation. Regardless of the fact that the fundamental impact on mechanisms and novel treatment and prevention approaches to glucose intolerance and diabetes would be small, further investigation of these scores in increased sample size could offer compelling results. Moreover, such data and analyses do not support information about the prospective effect of a change in the lifestyle pattern. For this purpose, longitudinal studies are needed in order to conduct observations of the same subjects for a period of time after adopting a healthier lifestyle. Observational studies, especially cross-sectional study designs, such as ours, are characterised by limited potential to offer causality or of predictive value conclusions.

In summary, screening of predisposing genetic variants, reliably associated with glycaemic traits that demonstrate an attenuated impact under the influence of protective lifestyle behaviour, could contribute to better recommendations for glucose homeostasis control. Further research in this direction could contribute to a better comprehension of how dietary intake and physical activity recommendations could be customised to the individual's genetic background. A combination of these tools would be useful to evaluate the gene–environment interaction, elucidate our understanding in terms of the biological pathways involved and help in prognosis, prevention and monitoring of glucose homeostasis.

Funding

Recruitment for THISEAS was partially funded by a research grant (PENED 2003) from the Greek General Secretary of Research and Technology. PD's work forms part of the research themes contributing to the translational research portfolio of Barts Cardiovascular Biomedical Research Unit which is supported and funded by the National Institute for Health Research.

Duality of interest

The authors declare that there is no duality of interest associated with this manuscript.

Contribution statement

EM participated in analysis design, statistical analysis and data interpretation and drafted the manuscript; SK participated in analysis design and genotyping; MD participated in sample recruitment, carried out data manipulation and performed dietary and physical activity data analyses; GK participated in sample recruitment; PD participated in coordination and genotyping of the study; and GD participated in coordination of the study.

All authors read and approved the final manuscript.

Acknowledgements

We thank all the dieticians and clinicians for their contribution to the project.

Authors would like to thank all participants and volunteers for their contribution in THISEAS.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.numecd.2015.10.003.

References

[1] Tuomilehto J, Lindstrom J, Eriksson JG, Valle TT, Hamalainen H, Ilanne-Parikka P, et al. Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. N Engl J Med 2001;344:1343–50.

[2] Santaguida PL, Balion C, Hunt D, Morrison K, Gerstein H, Raina P, et al. Diagnosis, prognosis, and treatment of impaired glucose tolerance and impaired fasting glucose. Evid Rep Technol Assess Summ 2005:1–11.

[3] de Vegt F, Dekker JM, Jager A, Hienkens E, Kostense PJ, Stehouwer CD, et al. Relation of impaired fasting and postload glucose with incident type 2 diabetes in a Dutch population: the Hoorn Study. JAMA 2001;285:2109–13.

[4] Balkau B, Hu G, Qiao Q, Tuomilehto J, Borch-Johnsen K, Pyorala K. Prediction of the risk of cardiovascular mortality using a score that includes glucose as a risk factor. The DECODE Study. Diabetologia 2004;47:2118–28.

[5] Snell-Bergeon JK, Wadwa RP. Hypoglycemia, diabetes, and cardiovascular disease. Diabetes Technol Ther 2012;14(Suppl. 1): S51–8.

[6] Scott RA, Lagou V, Welch RP, Wheeler E, Montasser ME, Luan J, et al. Large-scale association analyses identify new loci influencing glycaemic traits and provide insight into the underlying biological pathways. Nat Genet 2012;44:991–1005.

[7] Nettleton JA, Hivert MF, Lemaitre RN, McKeown NM, Mozaffarian D, Tanaka T, et al. Meta-analysis investigating associations between healthy diet and fasting glucose and insulin levels and modification by loci associated with glucose homeostasis in data from 15 cohorts. Am J Epidemiol.177:103–115.

[8] Frayling TM. Genome-wide association studies provide new insights into type 2 diabetes aetiology. Nat Rev Genet 2007;8: 657–62.

[9] Nettleton JA, McKeown NM, Kanoni S, Lemaitre RN, Hivert MF, Ngwa J, et al. Interactions of dietary whole-grain intake with fasting glucose- and insulin-related genetic loci in individuals of European descent: a meta-analysis of 14 cohort studies. Diabetes Care.33:2684–2691.

[10] Mann JI, De Leeuw I, Hermansen K, Karamanos B, Karlstrom B, Katsilambros N, et al. Evidence-based nutritional approaches to the treatment and prevention of diabetes mellitus. Nutr Metab Cardiovasc Dis 2004;14:373–94.

[11] Turner KM, Keogh JB, Clifton PM. Dairy consumption and insulin sensitivity: a systematic review of short- and long-term intervention studies. Nutr Metab Cardiovasc Dis.25:3–8.

[12] Yu LL, Zhang JL, Yan JY, Zhao JS, Wang LS, Zhang BY, et al. Physical activity level is associated with fasting plasma glucose among women aged 35-55 without diabetes in Shandong, China. Nutr Metab Cardiovasc Dis 2015;25:889–90.

[13] Kavouras SA, Panagiotakos DB, Pitsavos C, Chrysohoou C, Anastasiou CA, Lentzas Y, et al. Physical activity, obesity status, and glycemic control: the ATTICA study. Med Sci Sports Exerc 2007;39: 606–11.

[14] Matsuo T, So R, Shimojo N, Tanaka K. Effect of aerobic exercise training followed by a low-calorie diet on metabolic syndrome risk factors in men. Nutr Metab Cardiovasc Dis 2015;25:832–8.

[15] Wu Y, Zhang D, Jiang X, Jiang W. Fruit and vegetable consumption and risk of type 2 diabetes mellitus: a dose–response meta-analysis of prospective cohort studies. Nutr Metab Cardiovasc Dis 2015;25:140–7.

[16] Franks PW. Gene x environment interactions in type 2 diabetes. Curr Diab Rep 2011;11:552–61.

[17] Pearson H. Genetic testing for everyone. Nature 2008;453:570–1.

[18] Theodoraki EV, Nikopensius T, Suhorutsenko J, Peppes V, Fili P, Kolovou G, et al. Fibrinogen beta variants confer protection against coronary artery disease in a Greek case-control study. BMC Med Genet 2010;11:28.

[19] Dietary guidelines for adults in Greece. Arch Hellenic Med 1999; 16(5):516–24.

[20] Rasmussen-Torvik LJ, Li M, Kao WH, Couper D, Boerwinkle E, Bielinski SJ, et al. Association of a fasting glucose genetic risk score with subclinical atherosclerosis: the Atherosclerosis Risk in Communities (ARIC) study. Diabetes 2011;60:331–5.

[21] Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81: 559–75.

[22] Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet, 88:76–82.

[23] Ottman R. Gene-environment interaction: definitions and study designs. Prev Med 1996;25:764–70.

[24] Ioannidis JP, Boffetta P, Little J, O'Brien TR, Uitterlinden AG, Vineis P, et al. Assessment of cumulative evidence on genetic associations: interim guidelines. Int J Epidemiol 2008;37:120–32.

[25] Franks PW, Nettleton JA. Invited commentary: gene X lifestyle interactions and complex disease traits–inferring cause and effect from observational data, sine qua non. Am J Epidemiol 2010;172: 992–7. discussion 8–9.

[26] Saha S, Gerdtham UG, Johansson P. Economic evaluation of lifestyle interventions for preventing diabetes and cardiovascular diseases. Int J Environ Res Public Health 2010;7:3150–95.

Unpublished reports: Papers 3-4

3.2. Unpublished reports: Paper 3 and Paper 4

Paper 3 is the collaborative effort of CARDIoGRAMplusC4D Consortium to identity new risk loci for CAD. Paper 4 examines the cumulative effect of the variants published in two recent reports from CARDIoGRAMplusC4D Consortium within the THISEAS study.

Paper 3 | THE EFFECT OF A MULTI-LOCUS GRS ON CAD RISK

A GRS was tested on a total of 998 cases and controls from the THISEAS study.

Key points

- A GRS was constructed based on a panel of 53 previously reported variants on 422 CAD patients and 576 controls. Both the weighted and the unweighted score were used for the analyses.
- The majority of the 53 previously reported SNPs revealed consistent directional effects with those previously reported.
- The weighted GRS was associated with increased CAD risk, after adjusting for age and sex.
- Similar results were demonstrated for the unweighted GRS.

Paper 4 | DIETARY PATTERNS - COMBINED EFFECT OF A DIETARY COMPONENT AND GENE SUCEPTIBILITY ON CAD RISK

Participants consisted of 356 cases and 661 controls drawn from the THISEAS study. Dietary components were generated through FA. The combined effect of a dietary component and gene susceptibility on CAD risk was also examined.

Key points

- FA analysis generated eight dietary components.
- The Western type dietary component showed modest association with CAD risk.

The findings possibly suggest that the effect of high adherence to an unhealthy pattern is not similar among individuals with different genetic backgrounds. To further support this finding, it needs to be replicated in larger samples sizes so as robust associations to be found.

Additive influence of genetic variants increases the risk of coronary artery disease in a Greek sample: The THISEAS study

Abstract

Objective: The main aim of the current study was to evaluate whether a genetic risk score (GRS), constructed by 53 previously reported single nucleotide polymorphisms (SNPs), is associated with coronary artery disease (CAD) risk in Greek adults.

Methods and Results: The study population consisted of 422 CAD patients and 576 controls, drawn from the THISEAS database. We developed a weighted GRS (wGRS-53), based on 53 variants associated with CAD and previously reported in two of the largest genetic studies. One variant, rs8042271 at MFGE8-ABHD2 locus, reached significant levels for association with CAD risk (p=0.0005). Two variants, rs602633 at SORT1 locus and rs6725887 at WDR12 locus demonstrated nominal evidence for association (*p*=0.032 and *p*=0.015, respectively) with CAD risk. In total, 35 out of 53 SNPs revealed consistent directional effects with those previously reported. The wGRS-53 was associated with increased CAD risk (OR=1.03, 95% CI=1.01-1.05, *p*=0.011), after adjusting for age and sex. In addition, the ORs of main risk factors on CAD risk remained almost the same after adding as a covariate the additive influence of genetic variants in regression models.

Discussion: This is the first study in a Greek sample that examined the association of a GRS, consisting of a set of variants from the largest CAD genetic studies. Future prospective studies could evaluate the utility of this GRS in the Greek population. GRSs could be useful for identifying individuals at high CAD risk and/or predicting CAD incidence.

Keywords
Genetic risk score
Coronary artery disease
Variants
Case- control

Coronary artery disease (CAD) remains one of the primary causes of death among non-communicable diseases worldwide (Lozano et al., 2010). Although death rates attributable to CAD have declined at an aggregate level, the burden of the disease on premature mortality still remains high (Lozano et al., 2010; Sidney et al., 2013). The existing scientific literature offers ample evidence regarding the environmental factors that contribute to CAD development, such as unhealthy dietary habits, physical inactivity and cigarette smoking (Hu, 2009; Jhamnani et al., 2015). In addition, genetic mechanisms are a strong component

to the disease, since genetic variants may contribute to a significant proportion (38-57%) to CAD mortality (Zdravkovic et al., 2002).

GWAs have successfully identified common genetic variants associated to CAD (McPherson et al., 2007; Samani et al., 2007; Wellcome Trust Case Control Consortium, 2007), although these variants confer a modest explanation of CAD heritability. The results from studies in larger sample sizes along with the collaborative effort of the CARDIoGRAM and C4D consortia, revealed 35 common variants associated with CAD (Peden & Farrall, 2011). Within a year, the CARDIoGRAMplusC4D Consortium increased the number of susceptibility chromosomal loci to 46 (CARDIoGRAMplusC4D Consortium et al., 2013). The same Consortium has most recently updated the list of loci that have been reliably associated with CAD risk with ten more loci with small effect sizes (an 1000 Genome-based GWA meta-analysis) (Nikpay et al., 2015).

The reported variants by the CARDIoGRAMplusC4D Consortium provide us with the challenge to evaluate their additive effects on CAD risk through a GRS. The combined effect of these variants into a score has not been previously reported and the utility of GRSs has not been adequately tested in samples from the Greek population.

The aims of the current report were (i) to evaluate whether a GRS, constructed by recently reported single nucleotide polymorphisms (SNPs), confers to a substantial increment of CAD risk in Greek adults and (ii) to estimate the odds ratios of risk factors on CAD, after taking into account genetic susceptibility.

Materials and methods

Study sample

The study population comprised up to 998 subjects of Greek origin, drawn from The Hellenic Study of Interactions between Single nucleotide polymorphisms and Eating in Atherosclerosis Susceptibility (THISEAS) database. The THISEAS study is a medical center-based case-control study conducted in the region of Athens. The analysis of the present report was restricted to 422 CAD patients and 576 controls, depending on the genetic data availability of the SNPs selected for the GRS construction. All participants were informed about the goals of the study and gave their written consent. The study protocol was approved by the Ethics Committee of Harokopio University.

Details of the study design are described elsewhere (Dimitriou et al., 2015).

SNP selection, Genotyping and Genetic Risk Score (GRS) calculation

Previously reported variants from the CARDIoGRAMplusC4D Consortium were selected in order to investigate their joint effects (CARDIoGRAMplusC4D Consortium et al., 2013; Nikpay et al., 2015). These large meta-analyses studies have included a total of 61 CAD associated variants. In the current report, we narrowed down the number of previously published variants to 53. Specifically, we included those variants with published results under the additive model. In addition, for two variants in the same locus, we selected the one with the most significant published association.

Genomic DNA was extracted from whole blood using the salting-out method (Miller et al. 1988). Genotyping was performed at Wellcome Sanger Institute, Hinxton, UK using the Illumina Omni Express array. Sample exclusion criteria included (i) sample call rate <95%, (ii) samples with sex discrepancies, (iii) ethnic outliers, (iv) samples with genome-wide heterozygosity higher than ± 3SD and (v) duplicated samples. SNP exclusion criteria was (i) Hardy–Weinberg Equilibrium (HWE) $p<10^{-4}$ and (ii) call rate ≥98% The variants that failed to meet quality control criteria were imputed. Imputation was performed based on the 1000 Genomes panel and IMPUTE2. Specifically, imputed data were used for 31 variants (from a total of 53). Specific information regarding the directly genotyped and imputed variants are provided in **Supplementary Table 1**.

Subsequently, we created a multilocus weighted GRS (wGRS-53) on an *α priori* basis. The wGRS-53 was calculated by counting the number of risk alleles (0, 1 or 2) carried by each individual for 53 SNPs, after multiplying the number of risk alleles per SNP by its β-estimate and then summing up the number of risk alleles across all SNPs tested (Hivert et al., 2011). Finally, the score per individual was divided by the average of the β-estimates of the 53 SNPs (Rasmussen-Torvik et al., 2011). We also provide results regarding the unweighted GRS (unwGRS-53) based on the same panel of SNPs.

Measurements of covariates

Information regarding several risk factors for CAD was collected at the time of the enrolment. For the purpose of this study, the examined risk factors were: hypertension, Type 2 diabetes mellitus (T2DM), smoking status, waist-to-hip ratio, energy intake and physical activity (PA) adoption.

Waist and hip circumference were measured using WHO protocol (WHO, 2012). The waist-to-hip ratio (WHR) was computed as the ratio of the circumference of the waist to the circumference of the hips.

Regarding lifestyle characteristics, smoking status was assessed as a categorical variable in two ways. Specifically, former smokers were either combined into one group with current smokers (smoking status I; current/former vs. never smokers) or were combined into the same group with never smokers (smoking status II; current vs. never/former smokers.). With respect to PA, the Harokopio Physical Activity Questionnaire (HAPAQ) was used (Maraki et al., 2010). PA adoption was defined as any type of leisure-time exercise in a regular basis and was assessed as a categorical variable whether subjects report leisure-time PA or not. Dietary information was collected through a 172-item picture sort food frequency questionnaire (FFQ). Mean energy intake was estimated using the Nutritionist Pro, version 2.2 software (Axxya Systems-Nutritionist Pro, Stafford, TX, USA) for participants with complete dietary information. Participants with extreme values in energy intake were excluded from the analysis.

Further details regarding the methods of measurements of interest, including body mass index (BMI), clinical, dietary and PA assessments can be found elsewhere (Dimitriou et al., 2015).

Statistical analysis

Continuous variables are expressed as mean and standard deviations, while categorical variables are expressed as relative frequencies. Unadjusted or adjusted logistic regression analysis was performed in order to test the association between wGRS-53 and CAD risk, by calculating the odds ratio (OR) and their corresponding 95% of confidence intervals (CIs). The calculation of ORs for CAD risk was also performed on the basis of hypertension (yes vs. no), type 2 diabetes mellitus (T2DM) (yes vs no), smoking status I (current/former vs. never smokers), smoking status II (current vs. never/former smokers), PA adoption (no vs. yes), energy intake (above or equal to vs below median) and WHR (above or equal to vs below median). Energy intake and WHR were split into two quantiles with regard to the median, separately for men and women. All analyses were based on 2-sided tests, while statistical significance was set at $p \leq 0.05$. The statistical package IBM SPSS Statistics 21.0 was used for statistical calculations aforementioned.

The impact of each variant on CAD risk was examined through logistic regression, either unadjusted or adjusted for age and sex and assuming an additive genetic model in SNPTEST v2.5.2 (https://mathgen.stats.ox.ac.uk/genetics_software/snptest/snptest.html). The statistical threshold for this analysis was set at a level of *p*=0.0009, based on Bonferroni correction for 53 tests.

Results

The descriptive characteristics of the study are presented in **Table 1.** The mean value of the wGRS-53 was higher in CAD patients per 1-unit compared to the control group (*p*=0.002) **(Table 1)**.

The distributions of the wGRS-53 in the study sample and separately in the two study groups (controls and cases) are presented in **Figures 1a. and 1.b.** In the study sample, the minimum and maximum values of the wGRS-53 were 31.1 and 69.5, respectively. The case group presented higher minimum and maximum wGRS-53 values (35.1 and 69.5, respectively) compared to the control group (31.1 and 68.9), respectively (data not shown). As expected, the mean values of the unwGRS-53 differed in the same direction between the two study groups **(Supplementary Table 3)**.

The association summary statistics of the 53 SNPs are depicted in **Table 2**. The logistic regression revealed consistent effects on CAD risk in the THISEAS study, for 35 out of the 53 SNPs, One variant, rs8042271 at MFGE8-ABHD2 locus, reached significance to a Bonferroni threshold (OR=1.48m 95% CI=1.24-1.64, p=0.00005). Furthermore, CAD risk was nominally associated with two variants, rs602633 at SORT1 locus (OR=1.27, 95% CI=1.02-1.59, *p*=0.032) and rs6725887 at WDR12 locus (OR=1.35, 95% CI=1.06-1.73, *p*=0.0015). The evaluation of the wGRS-53 on CAD risk, demonstrated that 1-unit increase of the wGRS-53 was associated with 3% increase of CAD risk, after adjusting for age and sex (*p*=0.011). This finding was further confirmed after taking into account the BMI (*p*=0.004) **(Table 3)**.

We also examined the effect of conventional risk factors on CAD risk, after controlling for age, sex and BMI. We also run the logistic regression analysis controlling for age, sex, BMI and wGRS-53. Adding the wGRS-53 as a covariate did not change the point of estimator of the OR of risk factors on CAD risk. The estimates for the ORs are presented in **Table 4**. Hypertension, T2DM, increased energy intake, absence of PA and increased WHR confer an increase in CAD risk. Current smokers had an approximate 1-fold increase in CAD risk when

compared to former/never smokers (smoking status II) (OR=2.07, 95% CI= 1.47-2.93, *p*=0.000) (Model 2). The estimate of OR further increased when former smokers were excluded from the never smokers reference group and were combined together with current smokers (smoking status I) (OR=3.30, 95% CI= 2.25-4.85, *p*=0.000) (Model 2).

Logistic regression models on the unwGRS-53 and estimates for the ORs with the unwGRS-53 as a confounder are demonstrated on **Supplementary Tables 4** and **5**.

Discussion

In this work, we tested the individual associations and evaluated the cumulative effects of the SNPs reported by the CARDIoGRAMplusC4D meta-analyses, on CAD risk. Although only one variant reached significance at MFGE8-ABHD2 and two variants at SORT1 and WDR12 loci met nominally significance threshold, $p<0.05$, the THISEAS findings revealed consistent directional effects and similar effect sizes for the majority of the tested variants, with those reported in the two meta-analysis from the CARDIoGRAMplusC4D Consortium.

Logistic regression analyses, either unadjusted or adjusted for main confounding risk factors, demonstrated that the multi-locus genetic score was associated with CAD risk. Specifically, CAD risk was increased by 3% for every 1-unit increase of the wGRS-53. Our finding regarding the association of GRS with CAD risk reflects a previous report on Swedish population. In line with our results, a GRS based on 46 SNPs pooled by the CARDIoGRAMplusC4D consortium, was applied in more than 10.000 Swedish and was associated with increased CAD risk (Ganna et al., 2013). In a current study, a GRS27 was applied on 23595 participants drawn from the Malmö Diet and Cancer study and was associated with CAD incidence. A GRS50 based on a panel of additional 23 SNPs further improved the predictive capability of CAD (Tada et al., 2015). However, the results from published studies that have used GRSs to estimate CAD risk are not directly comparable, since they have calculated GRSs based on different panels of SNPs.

As expected, logistic regression analysis showed that exposure to conventional risk factors, such as hypertension, T2DM, increased energy intake and increased WHR, increased the estimates of CAD risk, after taking into account genetic susceptibility. In addition, exposure to smoking increased CAD risk, after controlling for genetic susceptibility. These results are compatible with a recent population-based study conducted in Greece (Yiannakouris et al.,

2014). Logistic regression by CAD risk factors, adjusted for the unwGRS-53 instead of the wGRS-53, did not change the associations .

At this point, we should mention some limitations of the present study. The case and control group differed in the main variables (age, sex ratio and BMI). This unmatched case-control study design could possibly result to a less gain in efficiency than in matched case-control studies. However, we tried to eliminate for confounding by adjusted logistic regression models. Another limitation of the present study is the small sample size and therefore, is underpowered to detect consistent directional associations for all tested loci with those of the meta-analyses or to confirm significant associations at a Bonferroni threshold. In addition, we cannot claim evidence for causality or that the GRS provides evidence for CAD prediction.

Previously important prospective studies have used GRSs in an attempt to investigate the aggregate effects of genetic variants in order to predict cardiovascular disease risk (Thanassoulis et al., 2012; Tikkanen et al., 2013; Paynter et al. 2010). However, the use of a GRS in clinical practice still remains unattainable. It is likely that in the future, a GRS that aggregates rare variants with strong effects will be more successful in providing a larger increment on CAD risk.

Despite the limitations of the present study, the results are noteworthy, since up to date, genetic research in the Greek population is very limited. Furthermore, scientific evidence regarding the joint effects of genetic variants associated with CAD identified from GWAs studies is even more scarce. This is the first study in a Greek sample that examined the association of a GRS, consisting of a set of variants from the largest CAD genetic studies.

To conclude, we investigated the individual and cumulative effects of established associated loci with CAD risk in a Greek sample. The main outcome is that a GRS composed of 53 SNPs is significantly associated with CAD risk. We expect that this study will pave the way for prospective cohorts to further investigate the utility of this GRS in the Greek population. GRSs could be useful in personalised medicine for identifying individuals at high CAD risk and/or predicting CAD incidence.

References

The references for this report are provided in chapter 5, References.

Tables

Table 1. Descriptive characteristics of the study

	CAD patients (N= 422)		Controls (N= 576)		*p*-value*
	Mean or frequency	SD	Mean or frequency	SD	
Age (years)	61.7	10.3	54.9	13.2	0.006
Male sex (%)	82.5%		49.7%		0.000
BMI	27.9	4.0	28.7	4.7	0.000
wGRS-53	49.3	6.5	48.1	6.1	0.002

CAD=coronary artery disease, BMI=body mass index, wGRS-53= weighted genetic risk score

Data are expressed as mean and standard deviation (SD) or frequency (%)

**p*-values derived from student's t test or χ^2 test

Table 2. Association summary statistics for known susceptibility CAD loci

			CARDIoGRAMplusC4D association summary statistics				THISEAS association summary statistics			
SNP	**Candidate Gene**	**Chr**	**Effect allele/ Non-effect allele**	**EAF***	**OR ***	***p*-value***	**EAF**	**OR**	**95% CI**	***p*-value**
rs602633	**SORT1**	**1**	**C/A**	0.77	1.13	2.19×10^{-18}	**0.80**	**1.27**	**1.02-1.59**	**0.032**
rs11206510	PCSK9	1	T/C	0.84	1.04	5.09×10^{-3}	0.80	1.13	0,90-1.41	0.288
rs6725887	**WDR12**	**2**	**C/T**	0.11	1.10	5.29×10^{-8}	**0.15**	**1.35**	**1.06-1.73**	**0.015**
rs9818870	MRAS	3	T/C	0.14	1.05	1.83×10^{-3}	0.14	1.20	0.93-1.55	0.157
rs12190287	TCF21	6	C/G	0.59	1.04	6.48×10^{-4}	0.64	0.95	0.74-1.13	0.613
rs2048327	SLC22A3-LPAL2-LPA	6	G/A	0.35	1.05	1.09×10^{-5}	0.27	0.92	0.75-1.13	0.437
rs11556924	ZC3HC1	7	C/T	0.65	1.08	1.45×10^{-9}	0.67	0.95	0.73-1.13	0.580
rs1333049	CDKN2BAS1	9	C/G	0.47	1.21	1.08×10^{-34}	0.44	1.16	0.97-1.39	0.081
rs579459	ABO	9	C/T	0.21	1.04	2.13×10^{-2}	0.21	1.04	0.83-1.29	0.740
rs12413409	CYP17A1-CNNM2-NT5C2	10	G/A	0.89	1.08	4.12×10^{-3}	0.89	1.17	0.89-1.38	0.201
rs2505083	KIAA1462	10	C/T	0.42	1.06	2.82×10^{-7}	0.46	0.93	0.89-1.22	0.390
rs974819	PDGFD	11	A/G	0.29	1.08	2.03×10^{-9}	0.37	1.00	0.80-1.17	0.972
rs3184504	SH2B3	12	T/C	0.40	1.07	6.13×10^{-3}	0.48	0.99	0.80-1.16	0.954
rs9515203	COL4A1-COL4A2	13	T/C	0.74	1.08	1.13×10^{-8}	0.80	1.14	0.92-1.31	0.196
rs2895811	HHIPL1	14	C/T	0.43	1.04	1.18×10^{-4}	0.39	1.02	0.85-1.22	0.855
rs12936587	RAI1-PEMT-RASD1	17	G/A	0.59	1.04	2.06×10^{-4}	0.61	1.01	0.82-1.18	0.885
rs1122608	LDLR	19	G/T	0.76	1.06	3.72×10^{-6}	0.71	0.96	0.74-1.15	0.704
rs9982601	Gene desert (KCNE2)	21	T/C	0.13	1.10	8.69×10^{-9}	0.13	1.07	0.83-1.39	0.604
rs17114036	PPAP2B	1	A/G	0.91	1.09	2.68×10^{-5}	0.91	1.02	0.67-1.27	0.914
rs12205331	ANKS1A	6	C/T	0.81	1.01	4.36×10^{-1}	0.90	1.12	0.83-1.34	0.372
rs9369640	PHACTR1	6	A/C	0.65	1.09	1.11×10^{-12}	0.58	1.07	0.89-1.28	0.464
rs2047009	CXCL12	10	C/A	0.48	1.05	9.66×10^{-6}	0.47	1.11	0.93-1.33	0.228
rs11203042	LIPA	10	T/C	0.44	1.03	9.86×10^{-3}	0.43	0.99	0.80-1.16	0.977
rs15563	UBE2Z	17	C/T	0.52	1.01	2.44×10^{-1}	0.53	1.10	0.92-1.31	0.298
rs2281727	SMG6	17	C/T	0.36	1.04	8.46×10^{-4}	0.42	0.99	0.83-1.19	0.943
rs2075650	ApoE-ApoC1	19	G/A	0.14	1.11	5.86×10^{-11}	0.11	1.03	0.77-1.37	0.840
rs17464857	MIA3	1	T/G	0.87	1.02	1.55×10^{-1}	0.84	0.93	0.64-1.16	0.603

Table 2. Association summary statistics for known susceptibility CAD loci *(continued)*

			CARDIoGRAMplusC4D association summary statistics				THISEAS association summary statistics			
SNP	**Candidate Gene**	**Chr**	**Effect allele/ Non-effect allele**	**EAF***	**OR ***	***p*-value***	**EAF**	**OR**	**95% CI**	***p*-value**
rs12539895	7q22	7	A/C	0.19	1.02	4.00×10^{-2}	0.18	0.80	0.64-1.01	0.580
rs9326246	ZNF259-APOA5-APOA1	11	C/G	0.10	1.04	2.90×10^{-2}	0.10	1.02	0.68-1.27	0.865
rs7173743	ADAMTS7	15	T/C	0.58	1.06	2.46×10^{-7}	0.50	1.10	0.93-1.25	0.242
rs4845625	IL6R	1	T/C	0.47	1.04	3.46×10^{-8}	0.47	0.96	0.76-1.13	0.661
rs515135	APOB	2	G/A	0.83	1.08	2.17×10^{-8}	0.81	1.06	0.85-1.33	0.585
rs2252641	ZEB2-ACO74093.1	2	G/A	0.46	1.04	1.27×10^{-4}	0.46	0.99	0.83-1.19	0.958
rs1561198	VAMP5-VAMP8-GGCX	2	A/G	0.45	1.05	2.57×10^{-6}	0.45	1.05	0.88-1.25	0.616
rs7692387	GUCY1A3	4	G/A	0.81	1.06	1.85×10^{-5}	0.80	1.09	0.86-1.27	0.391
rs273909	SLC22A4-SLC22A5	5	C/T	0.14	1.09	2.00×10^{-7}	0.08	0.93	0.68-1.28	0.653
rs10947789	KCNK5	6	T/C	0.76	1.06	1.22×10^{-5}	0.79	0.94	0.69-1.15	0.623
rs4252120	PLG	6	T/C	0.73	1.06	1.82×10^{-5}	0.68	1.08	0.89-1.24	0.364
rs264	LPL	8	G/A	0.86	1.05	7.30×10^{-4}	0.84	1.15	0.93-1.33	0.169
rs9319428	FLT1	13	A/G	0.32	1.05	5.70×10^{-6}	0.29	1.05	0.86-1.28	0.627
rs17514846	FURIN-FES	15	A/C	0.44	1.05	7.35×10^{-7}	0.45	0.98	0.82-1.17	0.812
rs2954029	TRIB1	8	A/T	0.55	1.04	7.75×10^{-5}	0.59	1.07	0.88-1.22	0.441
rs6544713	ABCG5-ABCG8	2	T/C	0.30	1.06	1.57×10^{-7}	0.36	1.12	0.94-1.27	0.178
rs1878406	EDNRA	4	T/C	0.15	1.06	3.54×10^{-3}	0.17	1.01	0.87-1.39	0.429
rs2023938	HDAC9	7	G/A	0.10	1.07	5.25×10^{-5}	0.13	0.83	0.64-1.09	0.184
rs17087335	REST-NOA1	4	T/G	0.21	1.06	4.60×10^{-8}	0.18	0.99	0.79-1.25	0.940
rs3918226	NOS3	7	T/C	0.06	1.14	1.70×10^{-9}	0.08	1.32	0.96-1.80	0.086
rs10840293	SWAP70	11	A/G	0.55	1.06	1.30×10^{-8}	0.61	1.06	0.88-1.27	0.518
rs56062135	SMAD3	15	C/T	0.79	1.07	4.50×10^{-9}	0.79	1.08	0.85-1.26	0.471
rs8042271	**MFGE8-ABHD2**	**15**	G/A	0.90	1.10	3.70×10^{-8}	**0.93**	**1.48**	**1.24-1.64**	**0.0005**
rs7212798	BCAS3	17	C/T	0.15	1.08	1.90×10^{-8}	0.15	1.14	0.90-1.45	0.265
rs663129	PMAIP1-MC4R	18	A/G	0.26	1.06	3.20×10^{-8}	0.25	1.13	0.93-1.39	0.238
rs180803	POM121L9P-ADORA2A	22	G/T	0.97	1.20	1.60×10^{-10}	0.99	1.50	0.65-1.81	0.097

SNP=single nucleotide polymorphism, Chr=chromosome, EAF=effect allele frequency, OR=odds ratio, 95% CI=95% Confidence Interval, HWE=Hardy Weinberg Equilibrium Results were obtained using logistic regression models assuming an additive effect, adjusted for age and sex. Allelic test p-value, OR and 95% CI are shown for each single SNP. *EAF, OR and allelic test p-value as published by previous literature; Bold high-lighted locus indicated nominal evidence for association with CAD risk.

Table 3. Results from the logistic regression models for the evaluation of wGRS-53 on the risk of developing CAD

	OR	95% CI	*p*-value
Model 1*	1.03	1.01-1.05	0.002
Model 2†	1.03	1.01-1.05	0.011
Model 3‡	1.04	1.01-1.06	0.004

GRS-53= genetic risk score, CAD=coronary artery disease, OR=odds ratio, CI=confidence interval

* **Model 1** unadjusted

† **Model 2** adjusted for age, sex

‡ **Model 3** adjusted for age, sex and body mass index

Table 4. Logistic regression for CAD incidence by risk factors in the THISEAS study

	N	**ORs**	**95% CI**	***p*- value**
Hypertension (yes vs. no)	964			
*Model 1**		6.51	4.35-9.72	0.000
Model 2†		6.38	4.27-9.56	0.000
Type 2 diabetes mellitus (yes vs. no)	983			
Model 1		2.79	1.95-3.99	0.000
Model 2		2.71	1.90-3.89	0.000
Smoking status I (current/former vs. never smokers)	851			
Model 1		3.30	2.25-4.85	0.000
Model 2		3.35	2.28-4.94	0.000
Smoking status II (current vs. never/former smokers)	851			
Model 1		2.07	1.47-2.92	0.000
Model 2		2.07	1.47-2.93	0.000
Physical activity adoption (no vs. yes)	650			
Model 1		1.95	1.17-3.26	0.010
Model 2		1.95	1.16-3.26	0.011
Energy intake (above or equal to vs below median)	538			
Model 1		1.94	1.29-2.91	0.001
Model 2		1.94	1.29-2.91	0.001
Waist-to-hip ratio (above or equal to vs below median)	693			
Model 1		2.18	1.48-3.19	0.000
Model 2		2.20	1.50-3.24	0.000

CAD=coronary artery disease, THISEAS= The Hellenic study of Interactions between Single nucleotide polymorphisms and Eating in Atherosclerosis Susceptibility

*****adjusted for age, sex and body mass index

†adjusted for age, sex, body mass index and weighted genetic risk score

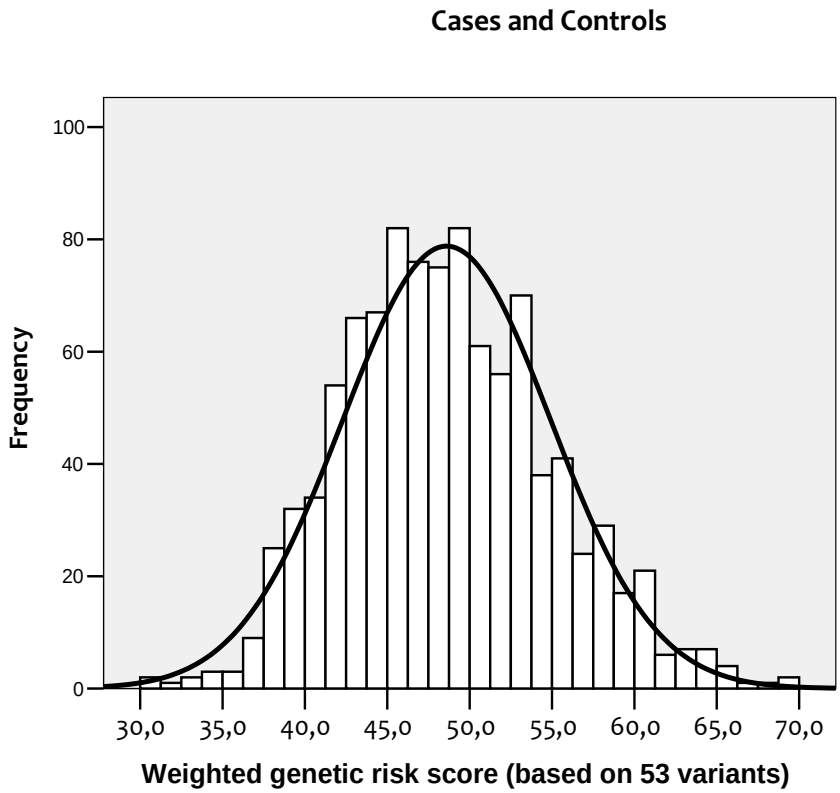


Figure 1. a. Distributions of the weighted genetic risk score (wGRS-53) in the study sample.

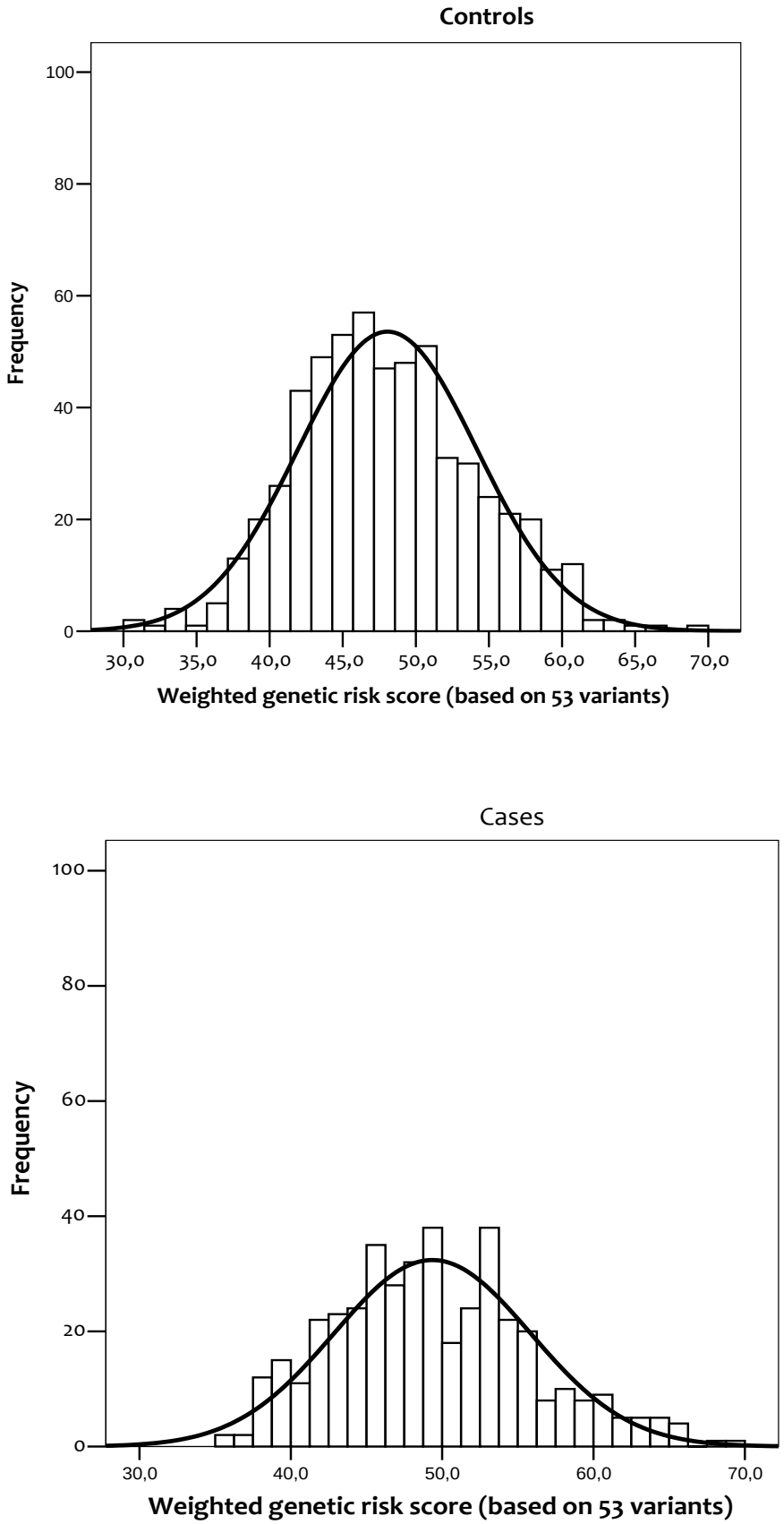


Figure 1.b. Distribution of the weighted genetic risk score (wGRS-53) separately in the control and case group.

The Association between dietary patterns and coronary artery disease using factor analysis.

In the present analysis the hypothesis tested was whether DPs of CAD patients and controls, as derived from FA, were associated with CAD risk.

Methods

Study sample

The study population comprised up to 1017 subjects of Greek origin, drawn from the THISEAS database. The analysis of the present report was restricted to 356 cases diagnosed with first-time CAD at the time of the recruitment and 661 controls, depending on the dietary data availability.

Statistical analysis

FA analysis was used to generate DPs (dietary components). Details regarding the study design, food groups and FA analysis are described in chapter 2. In total, 21 food groups were coded as servings per day and entered the FA analysis. The statistical packages IBM SPSS Statistics13.0 and 21.0 were used for statistical calculations.

Results

The descriptive characteristics (namely, the demographic, lifestyle and clinical) of the participants are presented in **Table 1.** As expected, the prevalence of physical inactivity, smoking, hypertension, hypercholesterolemia and T2DM was higher in cases compared to controls.

Table 2 summarizes the score coefficients (factor loadings) derived from FA. Absolute values greater than 0.4, indicate that the food variables are highly correlated and contribute more to the development of a dietary component. In this report, eight components were generated from the initial 21 food groups, that explained 53.5% of the total variation in intake. Specifically, the derived components were: a western type pattern (component 1), which included red meat, processed meat, fried potatoes and fast foods; a vegetarian type

pattern (component 2), which is mainly characterized by vegetables, legumes and potatoes (boiled, baked or smashed); a starch pattern (component 3) which was loaded with refined and unrefined starch (with the unrefined starch prevailing); a pattern that included the consumption of coffee and alcohol (component 4); a pattern that was characterized by the consumption of poultry, fish and seafood (component 5); a pattern that included dairy and eggs (component 6); a binge-eating type pattern (component 7) that included sweets and nuts intake; a pattern that included soft drinks and fruit drinks (component 8). Fruits and fresh fruit juice were not loaded.

In order to evaluate the associations between each extracted dietary component and CAD risk, logistic regression models were performed either without adjustments (Model 1) or after controlling for age, sex and BMI (Model 2). **Table 3** summarizes the results from logistic regression that evaluated the association between each dietary component and the likelihood of CAD. The unadjusted regression showed that component 1 (OR=1.10; 95% CI=1.01-1.10, *p*=0.034), component 4 (OR=1.03; 95% CI=1.02-1.04, *p*=0.000) and component 6 (OR=1.09; 95% CI=1.03-1.16, *p*=0.003) were positively associated with CAD risk. On the other hand, component 3, component 5 and component 8 were inversely associated with CAD risk. After adjusting for the aforementioned confounders, only the association of components 1 and 4 remained significant (OR=1.20; 95% CI=1.09-1.32, *p*=0.000 and OR=1.02; 95% CI=1.01-1.03, *p*=0.000). In particular component 1 revealed a modest effect on CAD risk, while component 4 revealed a small effect on CAD risk. The positively associated component 6, along with the inversely associated components 3, 5 and 8 lost their significance, in adjusted analysis. Component 7 was inversely correlated with CAD risk in both regression models. Component 2 was not associated with the likelihood of CAD, in both unadjusted and adjusted analysis.

Also, in an attempt to apply multiple regression analysis to model the associations between DPs and dependent variables to CAD (TC, LDL-C, TGs, SBP and blood glucose) in pooled sample or in controls did not reveal significant results (data not shown).

In a subset of individuals, with genetic and dietary data available, we sought the combined effect of genetic predisposition (using the wGRS-53 in the aforementioned report) and the western type diet adherence on CAD risk (**Table 4**).

The unadjusted analyses, showed that all three categories [low GRS+ high Western type diet adherence (p= 0.012), high wGRS-53 and low Western type diet adherence (p=0.015), high wGRS-53 and high Western type diet adherence (p=0.002) increased the odds of having CAD

compared to the reference group of low wGRS-53 + low Western type diet adherence. After controlling for age, sex and BMI, it was shown that high wGRS-53 + high Western type diet adherence increased the odds of having CAD by more that 1-fold, compared to low wGRS-53 + low Western type diet adherence (OR=2.30; 95% CI= 1.26-4.20, p=0.006). In the first two groups, the combined effect of genes and (unhealthy) diet on CAD risk, lost the significance. High Western type diet adherence did not increase the odds of CAD risk compared to low Western type diet adherence, among individuals with a genetic predisposition score in the highest quantile (OR=1.40; 95% CI= 0.79-2.49, p=0.255). Along similar lines, our data did not reveal association between high type Western diet adherence and CAD risk, among individuals with a genetic predisposition score in the highest quantile than among individuals with a genetic predisposition score in the lowest quantile (OR=1.30; 95% CI= 0.74-2.27, p=0.360).

Discussion

The present report assessed the likelihood of having CAD as a function of dietary components. DPs were identified from 21 food groups that entered the FA analysis, a *posteriori* technique. Specifically, eight components (patterns) were generated.

The first component could be described as a Western type pattern (unhealthy pattern); it was mainly characterized by the consumption of red meat, processed meat, fried potatoes and fast foods. This pattern was positively associated with CAD risk and the odds were further increased after controlling for age, sex and BMI. This finding is in concordance with the previously reported positive association between Western patterns and CAD in studies with high methodological quality (Mente et al., 2009). This component is characterized by the SFA intake and this type of fat has been implicated as a CAD risk factor. In another Greek study, a pattern characterized by meat intake and meat products has been associated with higher WC and lower HDL-C (Panagiotakos et. al 2007).

Most studies of alcohol and CAD risk have demonstrated a U-shaped relationship of CAD risk with total alcohol (Hill, 2005). Also, a U-shaped relationship was found between coffee consumption and all-cause mortality (Zhao et al., 2015). In the present report, the fourth component, which was characterized by coffee and alcohol consumption, was positively associated with the likelihood of having CAD. However, it is noteworthy to mention that the percentage of current smokers was almost double in cases than in controls and smokers tend to drink alcohol or coffee more than non-smokers. Therefore, we cannot rule out the

possibility that it is the lifestyle habit (smoking) and not the compliance with this pattern that affects CAD risk.

Component 2, a vegetarian type pattern, which is mainly characterized by the consumption of vegetables, legumes and potatoes (excluded fried potatoes) is a healthy dietary pattern. Vegetables and legumes are usually components of prudent patterns, which are inversely associated with CAD risk (Mente et al., 2009). In our findings, this component showed an expected directional effect on CAD risk, although not significant.

The inverse association of the seventh pattern (sweets and nuts) is in concordance with the results of a case-control Norwegian study. The latter demonstrated an inverse association of sweets with MI (Lockheart et al., 2007). The authors mention that although this was an unexpected outcome, this food group contained food items namely, nuts, almond paste and chocolate, which have been associated with reduced CAD risk. Similarly, in our report, this component also included nuts; in addition, among other sweets, almond chocolate and dark chocolate were also included. We cannot rule out the possibility that the latter food items have affected the direction of the association with CAD risk.

It has been shown that diets mainly characterized by foods with high glycaemic index scores have been associated with high TG and low HDL-C levels (Jeppesen et al., 1997) In addition, beverages with added sugar are associated with CAD risk (Fung et al., 2009). However, in this report, soft drinks and fruit drinks (component 8) did not reveal associations with CAD risk, after confounding. The third, fifth and sixth components also lacked significant associations, after confounding.

Our data could possibly suggest that the effect of high adherence to an unhealthy pattern (Western type diet) is not similar among individuals with different genetic backgrounds. However, to further support this finding, it needs to be replicated in larger samples sizes so as robust associations to be found.

Potential limitations of this report are the recall bias of food intake which may have resulted in underreported or overreported dietary intakes. Many cases had received dietary advice by the time of the interview, so we cannot rule out the possibility that they may have shaded their diet report towards favourable foods of the dietary advice. This is a case-control investigation that cannot support causality. FA as a technique also has some limitations, since the extracted components are based on subjective decisions.

Despite the limitations of this report, we assume that these results are noteworthy, since there is a gap in literature examining DPs on CAD risk or the combined effect of genes and

diet on CAD risk in the Greek population. We assume that some of the non-significant findings may well have been significant if the sample size was larger and had a larger discriminatory power.

References

The references for this report are provided in chapter 5, References.

Table 1. Descriptive characteristics of the participants

	Controls (N=661)		Cases (N=356)		
	Mean	**±SD***	**Mean**	**±SD**	**p-value**
Demographic & Lifestyle characteristics					
Age (years)	54.1	±14.1	62.5	±10.1	0.000
Years of education	12.3	±4.6	11.5	±4.9	0.007
	Relative Frequency (%)		**Relative Frequency (%)**		**p-value**
Physical inactivity	79.5%		90.9%		0.000
Current smokers	26.4%		46.7%		0.000
Clinical characteristics					
	Mean	**±SD**	**Mean**	**±SD**	**p-value**
Body mass index	28.4	±4.9	27.8	±3.8	0.040
Systolic blood pressure (mmHg)	133.8	±18.2	134.4	±20.2	0.944
Diastolic blood pressure (mmHg)	79.6	±11.2	80.0	±12.7	0.658
Total cholesterol (mg/dl)	210.3	±38.5	191.6	±47.8	0.000
Low-density lipoprotein cholesterol (mg/dl)	133.2	±34.7	122.6	±42.0	0.000
Triglyceride (mg/dl)	114.1	±64.2	148.4	±103.2	0.000
Blood glucose (mg/dl)	98.1	±22.8	113.2	±34.8	0.000
	Relative Frequency (%)		**Relative Frequency (%)**		**p-value**
Prevalence of hypertension	47.5%		90.3%		0.000
Use of antihypertensive medication	29.0%		85.1%		0.000
Prevalence of hypercholesterolemia	69.5%		88.4%		0.000
Use of lipid lowering medication	21.5%		79.4%		0.000
Prevalence of diabetes mellitus	10.6%		35.0%		0.000
Use of anti-diabetic medication	5.7%		21.8%		0.000

*SD=Standard deviation.

Table 2. Loadings from principal component analysis regarding food groups consumed by the participants from the THISEAS study

	Component[a]							
	1	2	3	4	5	6	7	8
Red meat	**0.584**	0.300	-0.117	0.281	0.043	0.087	0.033	-0.122
Processed meat	**0.671**	-0.076	0.097	0.082	-0.006	0.031	0.066	0.082
Potatoes, fried	**0.549**	0.135	-0.175	0.021	0.007	0.189	-0.189	0.039
Fast foods	**0.630**	-0.083	-0.165	-0.037	0.037	0.152	0.149	0.173
Vegetables	-0.105	**0.625**	0.164	0.081	0.268	0.089	-0.097	0.198
Legumes	-0.036	**0.666**	-0.147	-0.019	-0.010	-0.046	0.297	-0.145
Potatoes, boiled /baked/smashed	0.208	**0.615**	-0.054	-0.144	-0.034	0.082	-0.031	-0.010
Refined starch	0.139	0.256	**-0.645**	0.250	-0.057	0.056	0.056	0.124
Unrefined starch	-0.030	0.096	**0.813**	0.034	0.020	0.054	-0.041	-0.092
Coffee	0.059	-0.033	-0.104	**0.734**	-0.076	0.027	0.161	-0.006
Alcohol	0.070	-0.060	-0.13	**0.747**	0.168	0.070	-0.056	-0.001
Fish	-0.007	0.208	0.031	0.147	**0.606**	-0.048	-0.004	-0.168
Seafood	0.204	-0.108	-0.059	-0.003	**0.688**	0.065	-0.028	0.135
Poultry	0.537	0.082	0.098	-0.065	***0,431***	-0.331	-0.026	-0.099
Dairy, full fat	0.096	0.298	0.003	0.216	0.089	**0,565**	-0.193	0.041
Dairy, semi/non fat	-0.033	-0.020	0.443	-0.045	-0.201	**-0.440**	-0.214	0.292
Eggs	0.121	-0.026	0.019	-0.013	-0.084	**0.624**	0.193	-0.011
Sweets	0.214	-0.038	0.032	-0.126	0.132	0.336	**0.498**	0.292
Nuts	0.024	0.095	0.024	0.173	-0.029	-0.029	**0.739**	-0.070
Soft drinks	0.180	0.112	-0.023	0.066	-0.255	-0.058	-0.225	**0.701**
Fruit drinks	-0.02	0.081	-0.158	-0.047	0.224	-0.037	0.184	**0.590**
Fruits	-0.219	0.243	0.0201	-0.118	0.366	-0.020	0.279-	0.215

Numbers in bold indicate loadings with absolute value >0.4 (higher correlation of the food group with the component); Total %variance explained equals 53.5.

[a] Component description: Component 1= A Western type pattern; Component 2= A vegetarian type pattern; Component 3= A Starch type pattern; Component 4= A pattern that is characterized by the consumption of alcohol and coffee; Component 5=A pattern that is mainly characterized by the consumption of poultry, seafood and fish; Component 6= A pattern that is mainly characterized by the consumption of dairy and eggs; Component 7= A binge-eating type pattern that is mainly characterized by the consumption of sweets and nuts; Component 8= A pattern that is mainly characterized by the consumption of soft drinks and fruit drinks.

Table 3. Results from logistic regression that evaluated the association between dietary components and the likelihood of having coronary artery disease

	Odds ratio	95% CI	p-value
Component 1: [a]Western type diet pattern			
[*]*Model 1*	1.10	1.01-1.10	0.034
[**]*Model 2*	1.20	1.09-1.32	0.000
Component 2: [b]Vegetarian type diet pattern			
Model 1	0.97	0.90-1.05	0.484
Model 2	0.95	0.84-1.04	0.259
***Component 3:* [c]Starch type pattern**			
Model 1	0.95	0.91-0.99	0.007
Model 2	0.98	0.94-1.03	0.447
Component 4: Coffee and alcohol			
Model 1	1.03	1.02-1.04	0.000
Model 2	1.02	1.01-1.03	<0.001
Component 5: Poultry, fish and seafood			
Model 1	0.75	0.61-0.92	0.005
Model 2	0.85	0.67-1.07	0.156
Component 6: Dairy, cheese and eggs			
Model 1	1.09	1.03-1.16	0.003
Model 2	1.06	1.00-1.14	0.063
Component 7: [d]Binge-eating type pattern			
Model 1	0.79	0.71-0.89	0.000
Model 2	0.84	0.75-0.95	0.005
Component 8: [e]Soft/fruit drinks			
Model 1	0.66	0.52-0.83	0.001
Model 2	0.77	0.77-1.00	0.052

[*]No adjustments;

[**]Adjustments= age, sex and body mass index

[a]A pattern mainly characterized by the consumption of meat, processed meat, fast foods, fried potatoes and fast-food; [b] A pattern mainly characterized by the consumption of legumes, vegetables and potatoes; [c]A pattern mainly characterized by the consumption of unrefined starch; [d]A pattern mainly characterized by the consumption of sweets and nuts; [e]A pattern mainly characterized by the consumption soft drinks and fruit drinks

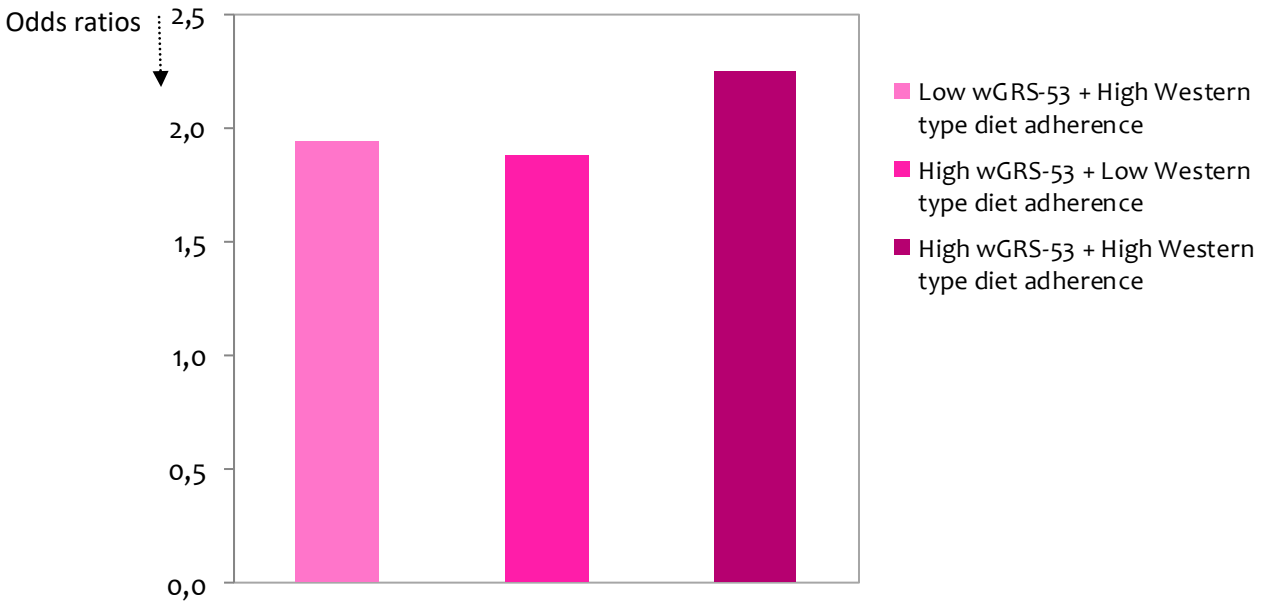
Table 4 Logistic regression results for the joint effect of genetic risk score and Western type diet adherence

wGRS-53 + Western type diet intake	N (controls/cases)	OR	95% CI	p-value	Adjusted OR	Adjusted 95%CI	Adjusted p-value
Low wGRS-53 + Low Western type diet adherence	143 (106/37)		*referent*			*referent*	
Low wGRS-53 + High Western type diet adherence	129 (77/52)	1.94	1.16-3.23	0.012	1.79	0.97-3.32	0.064
High wGRS-53 + Low Western type diet adherence	139 (84/55)	1.88	1.13-3.11	0.015	1.84	0.99-3.39	0.052
High wGRS-53 + High Western type diet adherence	134 (75/59)	2.25	1.36-3.74	0.002	2.30	1.26-4.20	0.006
	N (controls/cases)	**OR**	**95% CI**	**p-value**	**Adjusted OR**	**Adjusted 95%CI**	**Adjusted p-value**
High wGRS-53 + Low Western type diet adherence	139 (84/55)		*referent*			*referent*	
High wGRS-53 + High Western type diet adherence	134 (75/59)	1.42	0.84-2.41	0.189	1.40	0.79-2.49	0.255
Low wGRS-53 + High Western type diet adherence	129 (77/52)		*referent*			*referent*	
High wGRS-53 + High Western type diet adherence	134 (75/59)	1.16	0.71-1.90	0.542	1.30	0.74-2.27	0.360

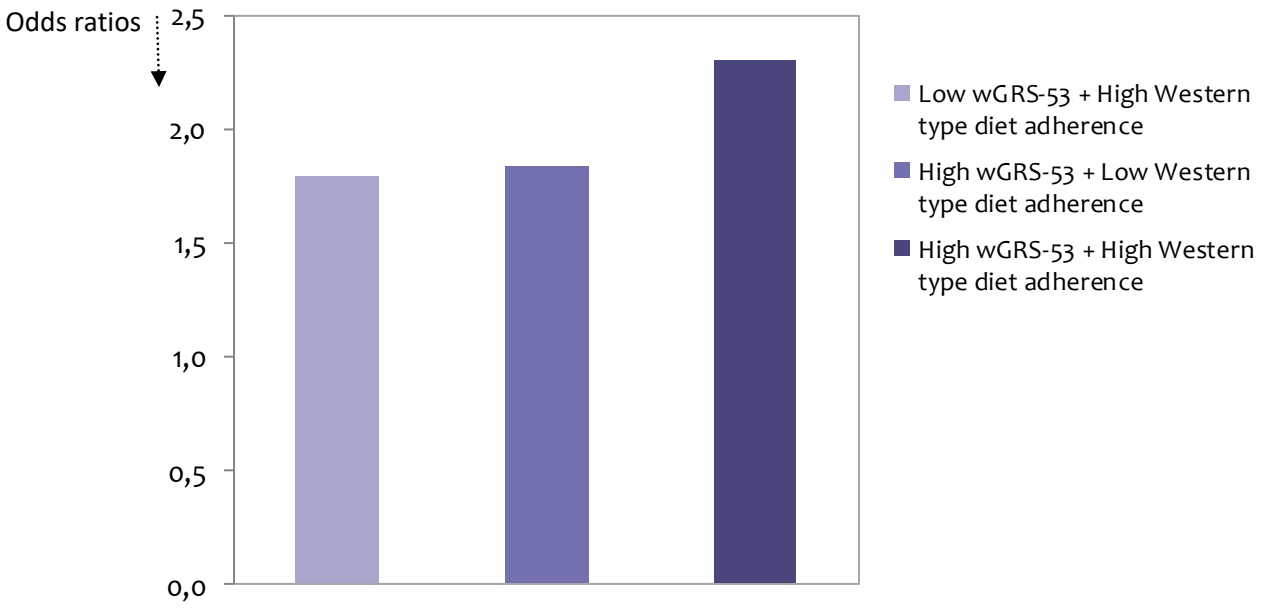
wGRS-53= weighted genetic risk score (based on a panel of 53 SNPs); OR= odds ratio; CI= confidence interval;
Adjustments= age, sex, body mass index

Figure 1. The combined effect of the genetic predisposition and Western type diet adherence on CAD risk in A) unadjusted logistic regression analyses and B) adjusted logistic regression analyses. (Reference group = Low wGRS-53 + Low Western type diet adherence)

A)



B)



Consortia participation: Papers **5-6**

3.3. Consortia participation: Paper 5 and Paper 6

Paper 5 | NEW RISK LOCI FOR CAD (CARDIoGRAMplusC4D CONSORTIUM)

This paper reflects the collaborative effort of investigators included in the CARDIoGRAMplusC4D Consortium.

Key points

- The paper presents an association analysis in 63,746 CAD cases and 130,681 controls.
- This is a 2-stage analysis using the Metabochip array and imputed GWAs data of European or south Asian descent.
- The results demonstrated 15 new loci at genome-wide significant to CAD. These novel loci, along with the 30 previously known CAD susceptibility loci, explained approximately 6% of the genetic susceptibility to CAD.
- A set of 104 independent SNPs (with a FDR 5%) were also identified.
- In total, the variants identified in this meta-analysis explained the additive genetic variance of CAD by 10.6%.

Paper 6 | DIETARY PATTERNS-GENE INTERACTION IN GLYCAEMIC TRAITS (CHARGE CONSORTIUM)

This paper reflects the collaborative effort of investigators included in the nutrition working group of the Cohorts for Heart and Ageing Research in Genomic Epidemiology (CHARGE) Consortium.

Key points

- Data from 15 US and European cohorts, comprising 51,289 individuals without diabetes, were included in the present meta-analysis. The purpose of the meta-analysis were to examine whether diet and genetic variation interact to influence FG and FI levels.
- 16 previously reported FG-associated loci and 2 FI-associated loci were selected for association analyses. A GRS was constructed based on the 16 FG-associated variants. A diet score was also constructed based on 9 food groups.
- Healthier diets were associated with lower FG and FI levels.
- The GRS was associated with FG levels.
- No interactions were demonstrated between the diet score and the GRS or the 16 FG-associated variants or the 2 FI-associated variants.

nature genetics

Large-scale association analysis identifies new risk loci for coronary artery disease

The CARDIoGRAMplusC4D Consortium[1]

Coronary artery disease (CAD) is the commonest cause of death. Here, we report an association analysis in 63,746 CAD cases and 130,681 controls identifying 15 loci reaching genome-wide significance, taking the number of susceptibility loci for CAD to 46, and a further 104 independent variants ($r^2 < 0.2$) strongly associated with CAD at a 5% false discovery rate (FDR). Together, these variants explain approximately 10.6% of CAD heritability. Of the 46 genome-wide significant lead SNPs, 12 show a significant association with a lipid trait, and 5 show a significant association with blood pressure, but none is significantly associated with diabetes. Network analysis with 233 candidate genes (loci at 10% FDR) generated 5 interaction networks comprising 85% of these putative genes involved in CAD. The four most significant pathways mapping to these networks are linked to lipid metabolism and inflammation, underscoring the causal role of these activities in the genetic etiology of CAD. Our study provides insights into the genetic basis of CAD and identifies key biological pathways.

Coronary artery disease and its main complication, myocardial infarction, is the leading cause of death worldwide. Although, epidemiological studies have identified many risk factors for CAD, including plasma lipid concentrations, blood pressure, smoking, diabetes and markers of inflammation, a causal role has been proven only for some (for example, low-density lipoprotein (LDL) cholesterol and blood pressure), primarily through randomized clinical trials of drug therapy directed at the risk factor[1]. Twin and family studies have documented that a significant proportion (40–50%) of susceptibility to CAD is heritable (for a review, see ref. 2). Because genotypes are not confounded by environmental exposures, genetic analysis has the potential to define which risk factors are indeed causal and to identify pathways and therapeutic targets[3,4]. To date, genome-wide association studies (GWAS) have collectively reported a total of 31 loci, associated with CAD risk at genome-wide significance ($P < 5 \times 10^{-8}$)[5–13]. However, variants at these loci explain less than 10% of the heritability of CAD. One likely reason for this is that, given the polygenic nature of complex traits and the relatively small observed effect sizes of the loci identified, many genuinely associated variants do not reach the stringent *P*-value threshold for genome-wide significance. Indeed, there is increasing evidence that the genetic architecture of common traits involves a large number of causative alleles with very small effects[14]. Addressing this will require the discovery of additional loci while leveraging large-scale genomic data to identify the molecular pathways underlying the pathogenesis of CAD. Such discovery is facilitated by building molecular networks, on the basis of DNA, RNA and protein interactions, which have nodes of known biological function that also show evidence of association with risk variants for CAD and related metabolic traits.

In the largest GWAS meta-analysis of CAD undertaken to date by the Coronary ARtery DIsease Genome-wide Replication and Meta-analysis (CARDIoGRAM) Consortium[5], which involved 22,233 cases and 64,762 controls, in addition to loci reported at genome-wide significance, a linkage disequilibrium (LD)-pruned set of 6,222 variants achieved a nominal association *P* value of less than 0.01. Here, we test these 6,222 SNPs in a meta-analysis of over 190,000 individuals, with the primary aim of identifying additional susceptibility loci for CAD. To this end, we used the Metabochip array[15], which is a custom iSELECT chip (Illumina) containing 196,725 SNPs, designed to (i) follow-up putative associations in several cardiometabolic traits, including CAD, and (ii) fine map confirmed loci for these traits. All SNPs on the array with data in the CARDIoGRAM study were considered for analysis (79,138 SNPs, of which 6,222 were the replication SNPs and 20,876 were fine-mapping SNPs in the 22 CAD susceptibility loci identified at the time at which the array was designed; the remaining SNPs were submitted by the other consortia contributing to the Metabochip array[15]). In addition, we assess whether the genome-wide significant CAD risk alleles act through traditional risk factors by considering the available large GWAS for these traits[16–20]. Finally, we identify a broader set of SNPs passing a conservative FDR threshold for association with CAD and use this set to undertake network analysis to find key biological pathways underlying the pathogenesis of CAD.

RESULTS

Study design

We expanded the CARDIoGRAM discovery data set (22,233 cases and 64,762 controls[5], stage 1) with 34 additional CAD sample collections (stage 2) of European or south Asian descent comprising 41,513 cases and 65,919 controls (study descriptions and sample characteristics are given in **Supplementary Tables 1a** and **2a**, respectively) and undertook a 2-stage meta-analysis to test SNPs on the Metabochip array

[1]A full list of authors and affiliations appears at the end of the paper.

Received 24 April; accepted 2 November; published online 2 December 2012; doi:10.1038/ng.2480

for disease association in a total of 63,746 cases and 130,681 controls. A further set of 3,630 cases and 11,983 controls from 4 independent studies was used for replication of SNPs that reached $5 \times 10^{-8} < P < 1 \times 10^{-6}$ in combined stage 1 and 2 analysis (stage 3; **Supplementary Tables 1b** and **2b**). An overview of the study design is provided in **Supplementary Figure 1**. Cases were selected for inclusion following the standard criteria for CAD and myocardial infarction used in the CARDIoGRAM study[5] (details for the stage 2 and 3 cohorts are given in **Supplementary Table 2**). Collections were typed with either the Metabochip array (60% of samples) or provided GWAS data imputed using HapMap (**Supplementary Table 3**). We applied standard quality control criteria to each study and corrected for population stratification if λ_{GC} was ≥1.05 (estimated for samples typed on the Metabochip using 4,310 SNPs associated with long QT syndrome and located at least 5 Mb away from established CAD risk loci; Online Methods). Case-control association analyses were adjusted for sex and age. For the 79,138 SNPs on the Metabochip with both stage 1 and 2 data, we combined (2-sided) *P* values from stage 1 with their respective (1-sided) *P* values for stage 2 using Fisher's method (Online Methods). In stage 3, we validated SNPs at $5 \times 10^{-8} < P < 1 \times 10^{-6}$ and combined evidence across all stages (1– 3) using a sample size–weighted meta-analysis.

Genome-wide significant loci

We first examined the 30 CAD risk loci previously reported in individuals of European ancestry at genome-wide significance (the *ADTRP* (*C6orf105*) locus has been reported only in Chinese)[12] in the stage 2 samples. For the 26 loci in which we could test the known lead SNP or a suitable proxy ($r^2 > 0.8$), we found highly significant associations in the stage 2 samples (**Table 1**). Notably, in four of these loci (*CDKN2B-AS1*, *COL4A2*, *CXCL12* and *APOE*), we detected additional SNPs not in LD ($r^2 < 0.5$) with the lead SNP, which also reached genome-wide significance and were conditionally independent when analyzed with GCTA software[21]. The additional SNP in the *APOE* locus, rs445925 ($P = 9.42 \times 10^{-11}$; $r^2 = 0.015$ with rs207560 in 1000 Genomes Project data), is located near *APOC1*, a gene previously suggested to confer risk for CAD[22]. The r^2 value between rs445925 ($P = 9.42 \times 10^{-11}$; $n = 31$ studies) and rs7412 ($P = 8.86 \times 10^{-4}$; $n = 21$ studies), which tags the *APOE* e2 allele,is 0.588. The *LIPA* locus also harbors a strong independent signal, which, however, did not reach genome-wide significance. Findings for the strongest associated variant available on the Metabochip for the other four loci (*MIA3*, 7q22, *ZNF259-APOA5-APOA1* and *ADAMTS7*) for which we did not have a good proxy for the previously reported lead SNP are also given (**Table 1**). Notably, for *ADAMTS7*, rs7173743 ($r^2 = 0.38$ with rs3825807, the published lead SNP) also achieved genome-wide significance.

We next examined the association of the 6,222 SNPs with $P < 0.01$ in CARDIoGRAM (we excluded SNPs in all loci listed in **Table 1**). Distribution of the absolute *z* scores for these SNPs in the stage 2 samples showed strong enrichment in positive scores corresponding to SNPs with directionally consistent signals between stages 1 and 2 under the null distribution, which is defined by mean = 0 and s.d. = 1 (4,260 SNPs observed versus 3,111 SNPs expected; binomial 2-sided $P = 7.5 \times 10^{-187}$) (**Supplementary Fig. 2**). In total, 19 loci showed association at $P < 1 \times 10^{-6}$ in the combined stage 1 and 2 analysis, with 13 of them reaching genome-wide significance, namely *IL6R*, *APOB*, *VAMP5-VAMP8-GGCX*, *SLC22A4-SLC22A5*, *ZEB2-AC074093.1*, *GUCY1A3*, *KCNK5*, *LPL*, *PLG*, *TRIB1*, *ABCG5-ABCG8*, *FURIN-FES* and *FLT1* (**Table 2**; Forest and regional association plots are given in **Supplementary Figs. 3** and **4**, respectively). The 6 loci with associations not reaching $P < 5 \times 10^{-8}$ were further validated (stage 3) in 4 independent studies (3,630 cases and 11,983 controls; **Supplementary Table 1b**). Two loci, *EDNRA* and *HDAC9* replicated at $P < 0.05$ and reached genome-wide significance in a combined analysis of stages 1–3 (**Table 2**); findings for those SNPs not meeting the above criteria are shown in **Supplementary Table 4**.

Of the newly associated loci reaching genome-wide significance, *TRIB1* and *ABCG5-ABCG8* were recently reported to reach study-wide significance ($P < 3 \times 10^{-6}$) in a large candidate gene (IBC array) study of CAD[13]. The same study reported rs2706399 in the *IL5* locus, which is located 200,349 bp away from the SNP we detected in the *SLC22A4-SLC22A5* locus (rs273909; **Table 2**). Although located in the same recombination interval, these SNPs are not in LD ($r^2 = 0.02$), and conditional analysis in a subset of 85,136 samples (up to19,200 cases) suggested that the 2 signals are conditionally independent; when conditioning on rs2706399 (*IL5* locus), the *P* value for rs273909 (*SLC22A4* locus) was 5.54×10^{-3} (1.33×10^{-3} initially), whereas the converse conditioning gave a *P* value of 3.34×10^{-2} for rs2706399 (*IL5*; 7.55×10^{-3} initially). We also detected a second signal in the *FES* locus (rs2521501; $P = 1.31 \times 10^{-9}$); conditional analysis with rs17514846 and rs2521501 ($r^2 = 0.43$ in 1000 Genomes Project data) showed the two signals not only to be independent but to also increase in strength upon conditioning (rs17514846 associated at $P = 1.07 \times 10^{-25}$ when conditioned on rs2521501; conversely, the *P* value for rs2521501 was 9.24×10^{-26}).

Subgroup analyses

Genetic risk of CAD could vary by age and gender and could also specifically influence the risk of its main adverse outcome, myocardial infarction[23]. We therefore undertook exploratory association analyses in subgroups partitioned by either gender, age at event (with individuals of <50 years of age being defined as young cases) or history of myocardial infarction (Online Methods). For the 46 genome-wide significant CAD risk loci, we observed no trend for higher odds ratios (ORs) in any of the subgroup analyses (**Supplementary Table 5**). However, one new locus reached genome-wide significance in males and in young CAD cases (rs16986953; $P = 1.89 \times 10^{-8}$ and 1.67×10^{-8}, respectively), which is located in a gene desert (with nearest transcript AK097927), 1.3 Mb away from the *APOB* gene. Interaction analysis conducted in a subset of studies ($n = 12$) where we had individual-level data provided suggestive evidence of an association with age ($P = 0.033$) but not with sex ($P = 0.708$); further studies are required to confirm this finding.

Wider Metabochip content

In addition to SNPs provided by the CARDIoGRAM Consortium, the Metabochip array contains a further 113,248 SNPs submitted for a range of cardiometabolic traits[15] other than CAD itself (associated at $P > 0.01$ with CAD in CARDIoGRAM samples or not tested). For these SNPs, we did not detect any new locus reaching genome-wide significance in our data set (including stage 1 and 3 data, when available). In total, therefore, we discovered 15 newly associated loci at genome-wide significance, increasing the total number of genome-wide significant loci to 45 in individuals of European and south Asian ancestry.

Localizing candidate CAD genes

To identify potential causal CAD-associated genes at the 15 new susceptibility loci identified in our study, we first analyzed genome-wide expression quantitative trait locus (eQTL) data in multiple tissues (circulating monocytes, liver, fat, skin, omentum, aortic media and adventitia, mammary artery and lymphoblastoid cell lines (LCLs)). We found that the lead SNP or a proxy in high LD ($r^2 \geq 0.8$) in three of the new loci was associated in *cis* with variable expression levels of the *GGCX-VAMP8*, *PLG* and *FES* genes (**Supplementary Table 6**).

Table 1 Association findings for known CAD susceptibility loci

Known loci[a]	Published lead SNP or proxy	New SNP (r^2 with lead SNP)	Chr.	Effect/non-effect allele (frequency)	Stage 2 OR	Stage 2 *P*	Combined *P*	Combined OR
SORT1[b]	rs602633 (tagging rs599839; $r^2 = 1.00$)		1	C/A (0.77)	1.13	2.19×10^{-18}	1.47×10^{-25}	1.12
PCSK9	rs11206510		1	T/C (0.84)	1.04	5.09×10^{-3}	1.79×10^{-5}	1.06
WDR12	rs6725887		2	C/T (0.11)	1.10	5.29×10^{-8}	1.16×10^{-15}	1.12
MRAS	rs9818870		3	T/C (0.14)	1.05	1.83×10^{-3}	2.62×10^{-9}	1.07
TCF21	rs12190287		6	C/G (0.59)	1.04	6.48×10^{-4}	4.94×10^{-13}	1.07
SLC22A3-LPAL2-LPA	rs3798220		6	C/T (0.01)	1.28	4.90×10^{-5}	N/A	N/A
		rs2048327 (0.03)	6	G/A (0.35)	1.05	1.09×10^{-5}	6.86×10^{-11}	1.06
ZC3HC1	rs11556924		7	C/T (0.65)	1.08	1.45×10^{-9}	6.74×10^{-17}	1.09
CDKN2BAS1	rs1333049		9	C/G (0.47)	1.21	1.08×10^{-34}	1.39×10^{-52}	1.23
		rs3217992 (0.50)	9	A/G (0.38)	1.14	7.27×10^{-32}	7.75×10^{-57}	1.16
ABO	rs579459		9	C/T (0.21)	1.04	2.13×10^{-2}	2.66×10^{-8}	1.07
CYP17A1-CNNM2-NT5C2	rs12413409		10	G/A (0.89)	1.08	4.12×10^{-3}	6.26×10^{-8}	1.10
KIAA1462	rs2505083		10	C/T (0.42)	1.06	2.82×10^{-7}	1.35×10^{-11}	1.06
PDGFD	rs974819		11	A/G (0.29)	1.08	2.03×10^{-9}	3.55×10^{-11}	1.07
SH2B3	rs3184504		12	T/C (0.40)	1.07	6.13×10^{-7}	5.44×10^{-11}	1.07
COL4A1-COL4A2	rs4773144		13	G/A (0.42)	1.06	2.34×10^{-6}	1.43×10^{-11}	1.07
		rs9515203 (0.01)	13	T/C (0.74)	1.08	1.13×10^{-8}	5.85×10^{-12}	1.08
HHIPL1	rs2895811		14	C/T (0.43)	1.04	1.18×10^{-4}	4.08×10^{-10}	1.06
RAI1-PEMT-RASD1	rs12936587		17	G/A (0.59)	1.04	2.06×10^{-4}	1.24×10^{-9}	1.06
LDLR	rs1122608		19	G/T (0.76)	1.06	3.72×10^{-6}	6.33×10^{-14}	1.10
Gene desert (*KCNE2*)	rs9982601		21	T/C (0.13)	1.10	8.69×10^{-9}	7.67×10^{-17}	1.13
PPAP2B	rs17114036		1	A/G (0.91)	1.09	2.68×10^{-5}	5.80×10^{-12}	1.11
ANKS1A	rs12205331 (tagging rs17609940; $r^2 = 0.85$)		6	C/T (0.81)	1.01	4.36×10^{-1}	4.18×10^{-5}	1.04
PHACTR1	rs9369640 (tagging rs12526453; $r^2 = 0.90$)		6	A/C (0.65)	1.09	1.11×10^{-12}	7.53×10^{-22}	1.09
CXCL12	rs501120		10	A/G (0.83)	1.06	7.13×10^{-5}	1.79×10^{-9}	1.07
		rs2047009 (0.05)	10	C/A (0.48)	1.05	9.66×10^{-6}	1.59×10^{-9}	1.05
LIPA	rs2246833 (tagging rs1412444; $r^2 = 0.98$)		10	T/C (0.38)	1.04	2.76×10^{-2}	9.49×10^{-6}	1.06
		rs11203042 (0.39)	10	T/C (0.44)	1.03	9.86×10^{-3}	6.08×10^{-6}	1.04
UBE2Z	rs15563 (tagging rs46522; $r^2 = 0.93$)		17	C/T (0.52)	1.01	2.44×10^{-1}	9.37×10^{-6}	1.04
SMG6	rs2281727 (tagging rs216172; $r^2 = 0.96$)		17	C/T (0.36)	1.04	8.46×10^{-4}	7.83×10^{-9}	1.05
ApoE-ApoC1	rs2075650		19	G/A (0.14)	1.11	5.86×10^{-11}	N/A	N/A
		rs445925 (0.03)	19	C/T (0.90)	1.13	8.76×10^{-9}	N/A	N/A
MIA3	N/A	rs17464857 (0.18)	1	T/G (0.87)	1.02	1.56×10^{-1}	6.06×10^{-5}	1.05
7q22	N/A	rs12539895 (0.64)	7	A/C (0.19)	1.02	4.00×10^{-2}	5.33×10^{-4}	1.08
ZNF259-APOA5-APOA1	N/A	rs9326246 (0.63)	11	C/G (0.10)	1.04	2.90×10^{-2}	1.51×10^{-7}	1.09
ADAMTS7	N/A	rs7173743 (0.38)	15	T/C (0.58)	1.06	2.46×10^{-7}	6.74×10^{-13}	1.07

Chr., chromosome.
[a]Locus *C6orf105*, which has been reported only in Chinese and has no good proxy SNP (Utah residents of Northern and Western European ancestry (CEU) or Han Chinese in Beijing, China (CHB)) on the Metabochip. The best available proxy is rs9348953 ($r^2 = 0.01$), with combined $P = 2.81 \times 10^{-3}$. [b]rs12740374, which was reported as a functional variant in this locus and has $r^2 = 0.895$ with rs599839, has combined $P = 8.25 \times 10^{-18}$ (OR = 1.135) based on the random-effects model used (*P* in stage 2 alone was 6.48×10^{-21} under the fixed-effect model).

We then assessed allele-specific expression data in monocytes, fibroblasts and LCLs and found three loci where the lead SNP was associated with an imbalance in expression of either *LPL*, *GGCX* or *FES*; *IL6R* showed some evidence of allele-specific expression in the fibroblast sample (**Supplementary Table 6**). Finally, we examined the new CAD risk loci for genes with relevant disease trait associations in mouse knockout models; six loci harbor a gene for which a mouse knockout model has a relevant cardiovascular phenotype, namely *ABCG8*, *APOB*, *GUCY1A3*, *PLG*, *LPL* and *FES* (**Supplementary Table 7**). *PLG* is adjacent to *LPA*, and, although the *PLG* risk variant rs4252120[T] was strongly associated with elevated Lp(a) lipoprotein levels ($P = 5 \times 10^{-24}$) in 3,698 PROCARDIS cases, it was associated with CAD independent of the LPA-linked variant at rs3798220. A detailed discussion of the genes in each locus is provided in the **Supplementary Note**. Of the 30 previously reported CAD susceptibility loci in individuals of European and south Asian ancestry, mouse knockout models for the candidate genes *PEMT*, *APOE*, *LDLR*, *COL4A1*, *LIPA*, *APOA1-APOA5*, *PPAP2B* and *PCSK9* also show phenotypic characteristics directly relevant to disease (**Supplementary Table 7**). In total, approximately a third of the 45 CAD loci contain a known functionally relevant candidate gene.

Overlap with traditional risk factors

We assessed both the known and new CAD susceptibility loci for overlap of associations with a number of relevant traits for which summary statistics have been made available: lipid levels (GLGC)[16], blood pressure (ICBPG)[17], diabetes (DIAGRAM)[18], glucometabolic traits (fasting insulin and fasting glucose concentrations, HOMA-B

Table 2 Additional loci showing genome-wide significant association with CAD

				Stage 1 (18,014 cases and 40,925 controls)[a]		Stage 2 (40,365 cases and 63,714 controls)		Combined (stages 1 and 2)	Stage 3 (5,055 cases and 5,617 controls)		Combined (stages 1–3)	
SNP	Chr.	Nearest gene(s)	Effect/non-effect allele (frequency)	OR	*P*	OR	*P*	*P*	OR	*P*	*P*	Biological relevance[b]
New												
rs4845625	1	*IL6R*	T/C (0.47)	1.06	4.84×10^{-5}	1.04	3.46×10^{-5}	3.55×10^{-8}	1.09	1.58×10^{-3}	3.64×10^{-10}	2
rs515135	2	*APOB*	G/A (0.83)	1.07	8.63×10^{-4}	1.08	2.17×10^{-8}	4.80×10^{-10}	1.03	4.02×10^{-1}	2.56×10^{-10}	1
rs2252641	2	*ZEB2-AC074093.1*	G/A (0.46)	1.06	1.37×10^{-5}	1.04	1.27×10^{-4}	3.66×10^{-8}	1.00	9.54×10^{-1}	5.30×10^{-8}	
rs1561198	2	*VAMP5-VAMP8-GGCX*	A/G (0.45)	1.06	7.47×10^{-5}	1.05	2.57×10^{-6}	4.48×10^{-9}	1.07	1.75×10^{-2}	1.22×10^{-10}	A,1
rs7692387	4	*GUCY1A3*	G/A (0.81)	1.08	1.04×10^{-5}	1.06	1.89×10^{-5}	4.57×10^{-9}	1.13	5.47×10^{-4}	2.65×10^{-11}	1
rs273909	5	*SLC22A4-SLC22A5*	C/T (0.14)	1.07	3.24×10^{-3}	1.09	2.00×10^{-7}	1.43×10^{-8}	1.11	2.43×10^{-2}	9.62×10^{-10}	A,1
rs10947789	6	*KCNK5*	T/C (0.76)	1.07	6.07×10^{-5}	1.06	1.22×10^{-5}	1.63×10^{-8}	1.01	7.03×10^{-1}	9.81×10^{-9}	3
rs4252120	6	*PLG*	T/C (0.73)	1.07	1.18×10^{-5}	1.06	1.82×10^{-5}	5.00×10^{-9}	1.07	9.58×10^{-2}	4.88×10^{-10}	1
rs264	8	*LPL*	G/A (0.86)	1.11	2.99×10^{-7}	1.05	7.30×10^{-4}	5.06×10^{-9}	1.06	1.60×10^{-1}	2.88×10^{-9}	1
rs9319428	13	*FLT1*	A/G (0.32)	1.06	7.88×10^{-5}	1.05	5.70×10^{-6}	1.01×10^{-8}	1.10	1.37×10^{-3}	7.32×10^{-11}	1
rs17514846	15	*FURIN-FES*	A/C (0.44)	1.07	2.37×10^{-5}	1.05	7.35×10^{-7}	4.49×10^{-10}	1.04	3.02×10^{-1}	9.33×10^{-11}	A,1
Previously reported at array-wide level of significance ($P < 3 \times 10^{-6}$)												
Rs2954029	8	*TRIB1*	A/T (0.55)	1.06	2.79×10^{-5}	1.04	7.75×10^{-5}	4.53×10^{-8}	1.05	8.56×10^{-2}	4.75×10^{-9}	4
Rs6544713	2	*ABCG5-ABCG8*	T/C (0.30)	1.06	2.22×10^{-4}	1.06	1.57×10^{-7}	8.72×10^{-10}	0.96	3.56×10^{-1}	2.12×10^{-9}	1
New (stage 3 replication)												
Rs1878406	4	*EDNRA*	T/C (0.15)	1.10	2.37×10^{-6}	1.06	3.54×10^{-3}	1.65×10^{-7}	1.09	2.01×10^{-2}	2.54×10^{-8}	1
Rs2023938	7	*HDAC9*	G/A (0.10)	1.08	6.81×10^{-4}	1.07	5.25×10^{-5}	6.49×10^{-7}	1.13	4.09×10^{-2}	4.94×10^{-8}	1

[a]Total sample sizes do not include the CHARGE sample sizes. [b]A, *cis* eQTL in LCLs; 1, mouse model available with cardiovascular phenotype; 2, mouse model has homeostatic and immune phenotypes; 3, mouse model has respiratory, nervous system, mortality, aging, growth and renal phenotypes; 4, mouse model has growth and immune phenotypes.

(homeostatic model assessment-β score) and HOMA-IR (insulin resistance); MAGIC)[19] and anthropometric traits (GIANT)[20,24]. After applying a Bonferroni correction for the 51 independent CAD-associated alleles tested (44 loci; no data available for rs16986953 and rs2521501), 12 loci showed evidence of association ($P < 1 \times 10^{-4}$) between the lead CAD risk SNP and 1 or more plasma lipid trait (total cholesterol, LDL cholesterol, high-density lipoprotein (HDL) cholesterol and triglyceride concentration) in the expected direction (the CAD risk allele was associated with higher total cholesterol, LDL cholesterol and triglyceride concentrations and lower HDL cholesterol concentration). These lead SNPs were most strongly associated with LDL cholesterol concentration at eight loci (*APOB*, *ABCG5-ABCG8*, *PCSK9*, *SORT1*, *ABO*, *LDLR*, *APOE* and *LPA*), with triglyceride concentration at two loci (*TRIB1* and the *APOA5* cluster) and with HDL cholesterol concentration at one locus (*ANKS1A*). There was near-equivalent association for triglyceride and HDL cholesterol concentrations at one locus (*LPL*). All loci except *LPA* and *ANKS1A* showed genome-wide significance for association with a lipid trait. These results underscore the importance of LDL cholesterol as a causal CAD risk factor (**Supplementary Table 8**). At the *SH2B3* locus, the CAD risk allele for rs3184504 was associated with both lower LDL cholesterol ($P = 1.73 \times 10^{-9}$) and HDL cholesterol ($P = 4.97 \times 10^{-6}$) concentration; one likely explanation is the presence of independent variants for CAD and LDL cholesterol. Two known CAD risk loci (*CYP17A1-NT5C2* and *SH2B3*) and two of the new CAD susceptibility loci (*GUCY1A3* and *FES*) have previously been associated with systolic (SBP) and diastolic (DBP) blood pressure[17]. Significant evidence for association with DBP was also observed for *ZC3HC1* (**Supplementary Table 8**). In contrast to the results for lipid concentration and blood pressure, there was no significant association of any of the loci tested with type 2 diabetes (T2D). Consistent with this observation, none of the assessed glucometabolic traits (fasting insulin and fasting glucose concentrations, HOMA-B and HOMA-IR) were related to these CAD variants (at the *ANKS1A* locus, it was not the CAD risk SNP that was associated with fasting insulin concentration and HOMA-IR). Suggestive associations ($P < 1 \times 10^{-4}$) with body mass index (BMI) and waist-hip ratio were observed in the *CYP17A1-CNNM2-NT5C2* and *RAI1-PEMT-RASD1* loci, respectively.

Additional suggestive associations

The genome-wide significance threshold, $P < 5 \times 10^{-8}$, we used is the accepted criterion for reporting individual association signals, as for each experiment it controls the error rate among common variants to less than 5%. However, SNPs showing suggestive association with a phenotype but not meeting this genome-wide threshold are likely to include additional true positive signals in well-powered studies (**Supplementary Fig. 1**). Such SNPs may also be informative in predicting CAD risk and in constructing CAD-associated biological networks. To identify such variants, we undertook an FDR analysis to assess the proportion of false positive signals in a set of (nominally) significant SNPs[25]. The Metabochip array contains both SNPs with priors in terms of association to CAD (CARDIoGRAM study $P < 0.01$) and blocks of highly correlated SNPs in fine-mapping regions. Therefore, to normalize the distribution of SNPs considered for FDR analysis, we (i) removed all SNPs in the CAD fine-mapping regions and LD-pruned ($r^2 < 0.2$) SNPs in the non CAD fine-mapping regions and (ii) adjusted the combined *P* values of all SNPs with priors in stage 1 ($P < 0.01$) using fixed-effect inverse variance–weighted meta-analysis *P* values for all other SNPs (Online Methods). In addition, we obtained 104 SNPs at an FDR threshold of 5% and LD threshold of $r^2 < 0.2$ (**Supplementary Table 9**). The median OR for CAD for these SNPs was 1.054 (interquartile range of 0.0199) per risk allele (**Supplementary Fig. 5**).

On the basis of a heritability estimate of 40% for CAD, the combination of the known and newly associated SNPs within the 45 susceptibility loci (**Tables 1** and **2**) explains approximately 6% of the additive genetic variance of CAD. The addition of the 104 SNPs from FDR analysis increased the fraction explained to 10.6% (Online Methods).

a

Canonical pathways

Biological function	*P* value
Atherosclerosis signaling	1.67×10^{-6}
LXR/RXR activation	2.14×10^{-6}
FXR/RXR activation	5.76×10^{-6}
Acute phase response signaling	4.36×10^{-5}

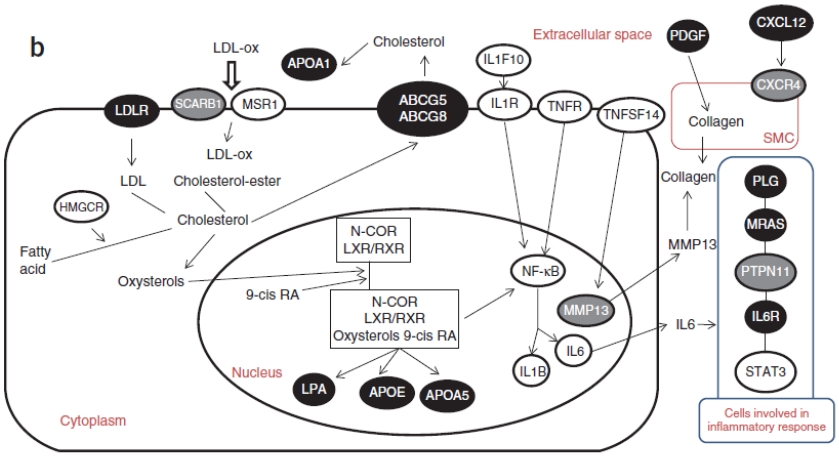


Figure 1 Canonical pathway analysis. (**a**) The four most significant canonical pathways represented in networks 3, 5 and 9, and overlapping networks ON1 (includes networks 1, 2, 6 and 8) and ON2 (includes networks 4 and 7); all molecules are listed by network in **Supplementary Table 10**. (**b**) Schematic showing parts of the atherosclerosis signaling, LXR/RXR activation and acute phase response signaling canonical pathways (Ingenuity) that are involved in both lipid metabolism and inflammation. Genes in confirmed CAD susceptibility loci (including both previously and newly reported) and in loci showing suggestive association with an FDR of <10% are depicted as black and gray ovals, respectively. Other key genes are depicted as white ovals; notably, some of them, such as *IL1F10-IL1B*, *STAT3* and *HMGCR*, have SNPs ranking in the top 1,000 in the FDR analysis. The process leading to myocardial infarction involves multiple cell types that are depicted in this schematic as a composite cell (large oval) and its nucleus (inner oval) in the extracellular space; the smooth muscle cell is shown separately (SMC; red oval), whereas the blue oval depicts cell types involved in the inflammatory response.

Network analysis

In contrast to estimating heritability where we want to keep the false positive rate as low as possible, in network analysis, we want to maximize the representation of potential network nodes in the gene set used. Thus, to perform network analysis, we selected the top 222 SNPs defined by the FDR analysis (10% FDR; final $P < 6.6 \times 10^{-4}$) at an LD threshold of $r^2 \leq 0.7$ and assigned 239 candidate genes on the basis of either eQTL data or physical proximity (**Supplementary Table 10**). We mapped 238 of the 239 genes in the Ingenuity Knowledge Base and considered 233 for network construction (Online Methods) on the basis of available data on interactions in humans, mice and/or rats (51 genes within the 46 genome-wide significant loci (set A) and 182 genes within the loci selected at FDR < 10% (set B)). Including neighboring genes, Ingenuity generated 9 networks comprising 553 nodes; these included 48 (94.1%) of the genes in set A and 156 (85.7%) of those in set B (**Supplementary Table 10**). We obtained 2 overlapping networks: ON1, which included networks 1, 2, 6 and 8, comprising the majority of genes in both sets (33 and 83 in sets A and B, respectively), and ON2, which included networks 4 and 7 (**Supplementary Table 10**). The nine networks were strongly enriched for genes (query set) known to be involved in lipid metabolism ($P = 1.48 \times 10^{-9}$), cellular movement (blood and endothelial cells; $P = 1.35 \times 10^{-7}$) and processes such as tissue morphology (size and area of atherosclerotic lesion, quantity of leukocytes, macrophages and smooth muscle cells; $P = 9.66 \times 10^{-10}$) and immune cell trafficking (migration and adhesion; $P = 1.12 \times 10^{-7}$). As a negative control in the network analysis, we used a set of 368 genes selected from the least significant SNPs in the FDR analysis; the resulting networks showed no significant enrichment in relevant molecular functions and process (results described in detail in the **Supplementary Note**).

We then assessed how genes in the networks overlap with canonical pathways in the Ingenuity database. The four most significant canonical pathways represented in these networks are shown in **Figure 1a**. The top three pathways, atherosclerosis signaling, liver X receptor (LXR)/retinoid X receptor (RXR) activation and farnesoid X receptor (FXR)/RXR activation, all harbor genes involved in lipid metabolism, including ten CAD risk loci (*ABCG5-ABCG8*, *APOA1*, *APOA5*, *APOB*, *APOE*, *CXCL12*, *LDLR*, *LPA*, *LPL* and *PDGFD*). This is in agreement with our finding that 12 CAD risk loci are associated with lipid levels at $P < 1 \times 10^{-4}$ (**Supplementary Table 8**). Notably, three of the top four pathways also contain genes involved in inflammation. In addition to the atherosclerosis signaling and LXR/RXR activation pathways, the acute phase response signaling (AAPRS) pathway, which includes four CAD risk loci (*APOA1*, *MRAS*, *IL6R* and *PLG*), is involved in inflammation and, more specifically, the rapid inflammatory response that is triggered, among other factors, by tissue injury. Genes from both the lipid metabolism and inflammation-related pathways map to all networks, except network 9, which harbors only two genes (**Supplementary Table 10**). As shown for overlapping network ON1 (**Supplementary Fig. 6**), genes in lipid metabolism and inflammation are interconnected and include both CAD-associated loci reaching genome-wide significance and candidate loci at FDR < 10%. Key interactions between CAD susceptibility genes (known, new and the FDR set) involved in lipid metabolism and inflammation are shown in **Figure 1b**; macrophages take up oxidized LDL (ox-LDL) through their cell surface scavenger receptors to form foam cells. Foam cells secrete proinflammatory cytokines, such as interleukin (IL)-1, IL-6 and matrix metalloproteinases, which can amplify the local inflammatory response and stimulate smooth muscle cell proliferation and initial migration toward the lesion[26]. Regulation of collagen secretion by smooth muscle cells in the extracellular matrix is regulated by matrix metalloproteinases. Reduction of collagen in the extracellular matrix will destabilize the plaque. Both *COL4A1* and *COL4A2* encode subunits of type IV collagen, which is the major structural component of basement membranes lining the inner surface of blood vessels. Metalloproteinases have a role in the maintenance of the extracellular matrix and remodeling, contributing to the transition of plaques from stable to vulnerable states (**Fig. 1b**).

DISCUSSION

Here, we report the largest genetic study to date assessing the impact of common variation on CAD risk. On the basis of analyses involving 63,746 CAD cases and 130,681 controls, we identified 15 new risk alleles at genome-wide significance, bringing the total number of confirmed CAD susceptibility loci in individuals of European and south Asian ancestry to 45. We also identified a further set of 104 likely independent ($r^2 < 0.2$) SNPs associated at an FDR of 5% with ORs between 1.031 and 1.126 per risk allele. In total, we estimate that these variants explain approximately 10.6% of the additive genetic variance of CAD (although we note that this may be an overestimate, given that it was not obtained in an independent sample). Our data also support the presence of additional true signals among the tested common SNPs that are likely to further contribute in explaining heritability; for example, the *P*-value adjustment we applied in the FDR analysis penalized the replication SNPs.

Among the 45 loci in individuals of European and south Asian ancestry that were confirmed to be associated with CAD, we found that 12 were significantly associated with the concentrations of blood lipids (mainly with LDL cholesterol), and 5 were associated with blood pressure. These data support the known etiological relationships of plasma lipids and blood pressure with CAD. People with T2D seem to have a 1.5- to 2-fold higher risk of CAD than those without diabetes[27], but none of the 45 risk loci were associated with diabetes status or with continuous levels of various glucometabolic traits. We note that, for the binary variable of T2D status, inability to show associations with CAD risk loci may reflect limited statistical power. The temporal relationship for comorbidity with both diabetes and CAD is complex: individuals with CAD without diabetes at diagnosis often subsequently develop T2D[28]. Furthermore, despite clear benefits in preventing microvascular disease (for example, retinopathy and nephropathy), intensive glucose control in diabetics reduces the risk of cardiovascular disease relatively modestly[29]. However, before a final conclusion can be reached, as many cohorts contributing to this meta-analysis focused by design on early disease manifestation or excluded diabetic individuals, a formal testing of the relationship of T2D and CAD in Mendelian randomization experiments will be necessary. To this end, the large genetic association data set on CAD assembled here will also facilitate testing of the causal relationship of other putative risk factors for CAD.

A desirable clinical goal is to integrate genetic information into a risk score for CAD in an attempt to provide improved predictive power over traditional risk factors in asymptomatic subjects, such that preventative measures, where available, can be more appropriately targeted. Our findings provide an appropriate framework of 153 CAD risk variants (at those established as susceptibility loci meeting the genome-wide significance threshold and additional suggestive loci with an FDR of <5%) for assessing a genetic risk score in well-powered prospective studies to determine whether they are sufficiently informative and independent predictors to have potential for use in day-to-day practice.

Allowing for inherent limitations in selecting likely candidate genes at each locus, our network analysis identified lipid metabolism and inflammation as key biological pathways involved in the genetic pathogenesis of CAD. Indeed, there was significant crosstalk between the lipid metabolism and inflammation pathways identified (**Fig. 1**). The emergence of lipid metabolism as a key pathway provides a positive control for the network and pathway analysis. On the other hand, this analysis provides strong new evidence at the molecular level in support of the causal involvement of inflammatory mechanisms in the pathogenesis of coronary atherosclerosis[30]. The role of inflammation in atherosclerosis is well documented in the literature[26]; for example, risk factors such as fat diet, smoking, hypertension, hyperglycemia, obesity and insulin resistance can trigger the expression of adhesion molecules (upregulated by atherogenic lipoproteins such as ox-LDL, very-low-density lipoprotein (VLDL) and Lp(a) lipoprotein) by endothelial cells, leading to the attachment of monocytes to the arterial wall. Although our analysis identified as significant the rapid inflammatory response pathway (mediated by NF-κB, MAPK and JAK-STAT signaling) that is primarily involved in innate immunity, many of the effector pathways in innate and adaptive immunity are heavily overlapping, and both are likely to have a role in CAD pathogenesis[26]. The five CAD-related networks constitute a useful framework for further functional and mechanistic studies to elucidate the biological processes underlying CAD pathogenesis and to investigate gene-environment interactions.

URLs. QVALUE software for FDR analysis, http://genomics.princeton.edu/storeylab/qvalue/; coronary heart disease statistics, http://www.bhf.org.uk/publications/view-publication.aspx?ps=1002097; top 10 causes of death fact sheet 310, http://www.who.int/mediacentre/factsheets/fs310/en/index.html; Uppsala Platform, http://molmed.medsci.uu.se/SNP+SEQ+Technology+Platform/Genotyping.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. Summary statistics for the 79,138 SNPs considered in this study for association with CAD (SNPs with stage 1 and stage 2 data) are available at ftp://ftp.sanger.ac.uk/pub/cardiogramplusc4d/.

Note: Supplementary information is available in the online version of the paper.

ACKNOWLEDGMENTS

We thank the personnel of the Wellcome Trust Sanger Institute (WTSI) Genotyping Facility, in particular S. Edkins, for supervising the genotyping of the AMC-PAS, Cardiogenics, GLACIER, MORGAM, PROMIS, THISEAS, and WTCCC cohorts.

AMC-PAS/SANQUIN.
We thank A.A. Soussan for technical assistance.

We thank personnel from the Estonian Genome Center of the University of Tartu (EGCUT) and the Estonian Biocentre, especially M. Hass and V. Soo, for data generation.

FINCAVAS.
We thank the staff of the Department of Clinical Physiology for collecting the exercise test data.

The GLACIER Study.
The GLACIER study is a nested study within the Northern Sweden Health and Disease Study; phenotyping was conducted as part of the Västerbotten Intervention Project. We thank the participants and the investigators from these studies for their valuable contributions, with specific thanks to L. Weinehall, Å. Agren, K. Enquist and T. Johansson.

GoDARTS Dundee.
We are grateful to all the participants who took part in this study, to the general practitioners, to the Scottish School of Primary Care for their help in recruiting the participants and to the whole team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. We acknowledge the support of the Health Informatics Centre at the University of Dundee in managing and supplying the anonymized data and National Health Service (NHS) Tayside, the original data owner.

Heart Protection Study.
The study was designed and conducted by the Clinical Trial Service Unit & Epidemiological Studies Unit (CTSU) at the University of Oxford. Genotyping was supported by a grant to Oxford University and Centre National de Genotypage (CNG) from Merck. The funders had no role in the design of the study or in the data collection or analysis. We especially acknowledge the participants in the study,

the Steering Committee and our collaborators. J.C.H. acknowledges support from the British Heart Foundation (BHF) Centre of Research Excellence.

LOLIPOP.
We thank the participants and research staff who made the study possible.

MORGAM study.
We thank the contributing sites and key personnel, as detailed below.

Finland: We thank FINRISK, National Institute for Health and Welfare, Helsinki: V.S. (principal investigator), A. Juolevi, E. Vartiainen and P. Jousilahti; Alpha-Tocopherol, Beta-Carotene Cancer Prevention (ATBC) study, National Institute for Health and Welfare, Helsinki: J. Virtamo (principal investigator) and H. Kilpeläinen; the MORGAM Data Centre, National Institute for Health and Welfare, Helsinki: K. Kuulasmaa (responsible person), Z. Cepaitis, A. Haukijärvi, B. Joseph, J. Karvanen, S. Kulathinal, M. Niemelä and O. Saarela; and the MORGAM Central Laboratory, National Institute for Health and Welfare, Helsinki: M.P. (responsible person), P. Laiho and M. Sauramo.

France: We thank the National Coordinating Centre, National Institute of Health and Medical Research (U258), Paris: P. Ducimetière (national coordinator) and A. Bingham; Prospective Epidemiological Study of Myocardial Infarction (PRIME)/Strasbourg, Department of Epidemiology and Public Health, EA 3430, University of Strasbourg, Faculty of Medicine, Strasbourg: D. Arveiler (principal investigator), B. Haas and A. Wagner; PRIME/Toulouse, Department of Epidemiology, Toulouse University School of Medicine, Toulouse: J.F. (principal investigator), J.-B. Ruidavets, V. Bongard, D. Deckers, C. Saulet and S. Barrere; PRIME/Lille, Department of Epidemiology and Public Health, INSERM U744–Université Lille Nord de France–Institut Pasteur de Lille, Lille: P. Amouyel (principal investigator), M. Montaye, B. Lemaire, S. Beauchant, D. Cottel, C. Graux, N. Marecaux, C. Steclebout and S. Szeremeta; and the MORGAM Laboratory, INSERM U937, Paris: F.C. (responsible person), L. Tiret and V. Nicaud.

Italy: We thank Centro Ricerche EPIMED–Epidemiologia e Medicina Preventiva, Dipartimento di Medicina Clinica e Sperimentale; Università dell' Insubria, Varese: M.M.F. (principal investigator) and G. Veronesi; and Research Centre on Public Health, University of Milano–Bicocca, Monza: G. Cesana.

UK: We thank PRIME/Belfast, Queen's University Belfast, Belfast: F.K. (principal investigator), A.E. (former principal investigator), J. Yarnell and E. Gardner; and the MORGAM Coordinating Centre, Queen's University Belfast, Belfast: A.E. (MORGAM coordinator), S. Cashman and F.K.
MORGAM management group: A.E. (chair), S.S.B., F.C., M.M.F., K. Kuulasmaa, A. Palotie, M.P., A.P., V.S., H. Tunstall-Pedoe and P.G. Wiklund. Previous members: K. Asplund, L. Peltonen, D. Shields and B. Stegmayr. The PRIME Study is organized under an agreement between INSERM and the Merck, Sharpe and Dohme-Chibret Laboratory, with the following participating laboratories: The Strasbourg MONICA Project, Laboratoire d'Epidémiologie et de Santé Publique, and the Université de Strasbourg, Strasbourg, France (D. Arveiler and B. Haas); The Toulouse MONICA Project, UMR INSERM 1027, and the Department of Epidemiology, Toulouse University School of Medicine, Université Paul Sabatier, Toulouse, France (J.F. and J.-B. Ruidavets); The Lille MONICA Project, INSERM U744, Institut Pasteur de Lille and Université Lille Nord de France, Lille, France (P. Amouyel and M. Montaye); The Department of Epidemiology and Public Health, Queen's University, Belfast, Belfast, UK (A.E., J. Yarnell and F.K.); The Department of Atherosclerosis, INSERM U545, Institut Pasteur de Lille, Faculté de Médecine and Université Lille Nord de France, Lille, France (G. Luc and J.-M. Bard); The Laboratory of Haematology, INSERM U626, and Hôpital La Timone, Marseille, France (I. Juhan-Vague and P. Morange); The Laboratory of Endocrinology, INSERM U563, Toulouse, France (B. Perret); The Vitamin Research Unit, The University of Bern, Bern, Switzerland (F. Gey); The Nutrition and Metabolism Group, Centre for Public Health, Queen's University Belfast, Belfast, UK (J. Woodside and I. Young); The DNA Bank, INSERM/Université Pierre et Marie Curie (UPMC), Paris Université Unite Mixte de Recherche (UMRS) 937, Paris (F.C.); The Coordinating Centre, Institut Fédératif de Recherche Santé Publique (IFR 69), Villejuif, France (P. Ducimetière); and INSERM U970, Villejuif, France, and University Paris V, Paris Cardiovascular Research Centre (PAARC), Paris (A. Bingham).

PIVUS/Swedish Twin Registry.
We thank the SNP&SEQ Technology Platform in Uppsala (see URLs) for genotyping, in particular T. Axelsson, A.-C. Wiman and C. Pöntinen for excellent assistance.

Ulm (EMIL).
We thank the Centre of Excellence Baden-Wuerttemberg Metabolic Disorders.

WTCCC.
We thank the BHF Family Heart Study Research Group for the collection of the cases.

AUTHOR CONTRIBUTIONS

Writing committee: P. Deloukas, S. Kanoni, C.W., M.F., T.L.A., J.R.T., E.I., D. Saleheen, J.E., M.P. Reilly, R. Collins, S. Kathiresan, A.H., U.T., J.S.K., J.D., C.N.A.P., R.R., H.W., H.S. and N.J.S. **Steering committee:** P. Deloukas, S. Kanoni, C.W., M.F., T.L.A., J.R.T., E.I., D. Saleheen, J.E., M.P. Reilly, R. Collins, S. Kathiresan, A.H., U.T., J.S.K., J.D., C.N.A.P., R.R., H.W., H.S., N.J.S., S.S.B., B.O.B., J.C.C., R. Clarke, G.D., P.W.F., C.H., G.K.H., Jong-Young Lee, T.L., W.M., A.M., M.S.N., C.O., M.P., S. Ripatti, M.S.S., S.S., A. Siegbahn, C.J.W. and P.A.Z. **Analysis committee:** B.A.G., K. Stirrups, I.R.K., J.-B.C., Å.J., T.E., L.F., A.G., A.S. Havulinna, W.K.H., J.C.H., N.E., M.E.K., K. Kristiansson, P.L., L.-P.L., S. Rafelt, D. Shungin, R.J.S., G. Thorleifsson, E.T., N.V.Z., B.F.V., L.L.W., W.Z. and A.Z. **Genotyping:** D. Absher, I.B., C.B., S.C.-B., DIAGRAM Consortium, N.E.M., K.F., P.F., B.G., L.G., S.G., J.H., B.-G.H., S.E.H., T.K., J.W.K., C. Langenberg, C. Langford, M.I.M., M.M.-N., K.N., J.F.P., S. Rosinger, D.R., M.P. Rumpf, A. Schäfer, A.F.R.S., P.J.W. and Wellcome Trust Case Control Consortium. **Array design:** H.M.K. and N.W.R. **Functional analyses:** E.G., P.E., A.F.-C., A.L., O.M., S.M., MuTHER Consortium, T.-P.Y., A.H.G., E.S., T.P. and A.-C.S. **Samples and phenotyping:** (ADVANCE) A.S.G., C.I. and T.Q.; (AMC-PAS/SANQUIN) C.E.v.d.S. and H.B.; (Angio-Lüb/KORA) P. Diemert; (CADomics) P.S.W.; (CARDIOGENICS) F.C. and W.H.O.; (CHARGE) E.B., A.L.C., A.D. and V.G.; (Corogene) M.-L.L. and J.S.; (deCODE) G. Thorgeirsson, H.H. and K. Stefansson; (EPIC-NORFOLK) N.W.; (Estonian Biobank) E.M.; (FGENTCARD) D.G.; (FINCAVAS) M.K.; (FINRISK 2007/DILGOM) V.S.; (FRISCII) L.W.; (GerMIFS) T.I., C.M., K. Stark and M.E.Z.; (GLACIER) G.H.; (GoDARTS Dundee) A.S.F.D. and A.D.M.; (HPS) S.P.; (Korean GenRIC) Y.J., H.-S.K., Ji-Young Lee and J.E.P.; (LOLIPOP) S.-T.T.; (LURIC/AtheroRemo) R.L. and W. Koenig; (METSIM) J.K., M.B. and M.L.; (MIGen) R.D.; (MORGAM) K. Kuulasmaa, J.V., P.A., D. Arveiler., J.F., D.-A.T., N.K., A.P., P.B., M.M.F., A.E. and F.K.; (Ottawa Heart Genomics Study) G.A.W., S.L.H. and S.H.S.; (PennCATH/MedStar) S.E.E. and D.J.R.; (Pfizer-Broad-Malmo) D. Altshuler and D.C.; (PIVUS/Swedish Twin Registry) C.S., L.L. and N.L.P.; (PROMIS) A.R.; (SHEEP-SCARF) K.L. and U.d.F.; (THISEAS) M.D., G.K.; (Ulm-EMIL) W. Kratzer; and (WTCCC) A.J.B., P.S.B., M.S. and A.S. Hall.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

Published online at http://www.nature.com/doifinder/10.1038/ng.2480.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Berry, J.D. *et al.* Lifetime risks of cardiovascular disease. *N. Engl. J. Med.* **366**, 321–329 (2012).
2. Peden, J.F. & Farrall, M. Thirty-five common variants for coronary artery disease: the fruits of much collaborative labour. *Hum. Mol. Genet.* **20** R2, R198–R205 (2011).
3. Kingsmore, S.F., Lindquist, I.E., Mudge, J., Gessler, D.D. & Beavis, W.D. Genome-wide association studies: progress and potential for drug discovery and development. *Nat. Rev. Drug Discov.* **7**, 221–230 (2008).
4. Stein, E.A. *et al.* Effect of a monoclonal antibody to PCSK9 on LDL cholesterol. *N. Engl. J. Med.* **366**, 1108–1118 (2012).
5. Schunkert, H. *et al.* Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat. Genet.* **43**, 333–338 (2011).
6. Coronary Artery Disease (C4D) Genetics Consortium. A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. *Nat. Genet.* **43**, 339–344 (2011).
7. Clarke, R. *et al.* Genetic variants associated with Lp(a) lipoprotein level and coronary disease. *N. Engl. J. Med.* **361**, 2518–2528 (2009).
8. Samani, N.J. *et al.* Genomewide association analysis of coronary artery disease. *N. Engl. J. Med.* **357**, 443–453 (2007).
9. Erdmann, J. *et al.* New susceptibility locus for coronary artery disease on chromosome 3q22.3. *Nat. Genet.* **41**, 280–282 (2009).
10. Kathiresan, S. *et al.* Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. *Nat. Genet.* **41**, 334–341 (2009).
11. Soranzo, N. *et al.* A genome-wide meta-analysis identifies 22 loci associated with eight hematological parameters in the HaemGen consortium. *Nat. Genet.* **41**, 1182–1190 (2009).
12. Wang, F. *et al.* Genome-wide association identifies a susceptibility locus for coronary artery disease in the Chinese Han population. *Nat. Genet.* **43**, 345–349 (2011).
13. The IBC 50K CAD Consortium. Large-scale gene-centric analysis identifies novel variants for coronary artery disease. *PLoS Genet.* **7**, e1002260 (2011).
14. Yang, J. *et al.* Genome partitioning of genetic variation for complex traits using common SNPs. *Nat. Genet.* **43**, 519–525 (2011).
15. Voight, B.F. *et al.* The Metabochip, a custom genotyping array for genetic studies of metabolic, cardiovascular, and anthropometric traits. *PLoS Genet.* **8**, e1002793 (2012).
16. Teslovich, T.M. *et al.* Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* **466**, 707–713 (2010).

17. International Consortium for Blood Pressure Genome-Wide Association Studies. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature* **478**, 103–109 (2011).
18. Voight, B.F. *et al.* Twelve type 2 diabetes susceptibility loci identified through large scale association analysis. *Nat. Genet.* **42**, 579–589 (2010).
19. Dupuis, J. *et al.* New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat. Genet.* **42**, 105–116 (2010).
20. Speliotes, E.K. *et al.* Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat. Genet.* **42**, 937–948 (2010).
21. Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* **44**, 369–375 (2012).
22. Ken-Dror, G., Talmud, P.J., Humphries, S.E. & Dreno, F. *APOE/C1/C4/C2* gene cluster genotypes, haplotypes and lipid levels in prospective coronary heart disease risk among UK healthy men. *Mol. Med.* **16**, 389–399 (2010).
23. Reilly, M.P. *et al.* Identification of *ADAMTS7* as a novel locus for coronary atherosclerosis and association of *ABO* with myocardial infarction in the presence of coronary atherosclerosis: two genome-wide association studies. *Lancet* **377**, 383–392 (2011).
24. Heid, I.M. *et al.* Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. *Nat. Genet.* **42**, 949–960 (2010).
25. Storey, J.D. & Tibshirani, R. Statistical methods for identifying differentially expressed genes in DNA microarrays. *Methods Mol. Biol.* **224**, 149–157 (2003).
26. Hansson, G.K., Libby, P., Schönbeck, U. & Yan, Z.-Q. Innate and adaptive immunity in the pathogenesis of atherosclerosis. *Circ. Res.* **91**, 281–291 (2002).
27. Emerging Risk Factors Collaboration. Diabetes mellitus, fasting glucose, and risk of cause-specific death. *N. Engl. J. Med.* **364**, 829–841 (2011).
28. Mozaffarian, D. *et al.* Incidence of new-onset diabetes and impaired fasting glucose in patients with recent myocardial infarction and the effect of clinical and lifestyle risk factors. *Lancet* **370**, 667–675 (2007).
29. Turnbull, F.M. *et al.* Intensive glucose control and macrovascular outcomes in type 2 diabetes. *Diabetologia* **52**, 2288–2298 (2009).
30. Ross, R. Atherosclerosis is an inflammatory disease. *Am. Heart J.* **138**, S419–S420 (1999).

The authors of this paper are:
Panos Deloukas[1,126], Stavroula Kanoni[1,126], Christina Willenborg[2,126], Martin Farrall[3,4,126], Themistocles L Assimes[5,126], John R Thompson[6,126], Erik Ingelsson[7,126], Danish Saleheen[8–10,126], Jeanette Erdmann[2,126], Benjamin A Goldstein[5], Kathleen Stirrups[1], Inke R König[11], Jean-Baptiste Cazier[4], Åsa Johansson[12], Alistair S Hall[13], Jong-Young Lee[14], Cristen J Willer[15,16], John C Chambers[17], Tõnu Esko[18,19], Lasse Folkersen[20,21], Anuj Goel[3,4], Elin Grundberg[22], Aki S Havulinna[23], Weang K Ho[10], Jemma C Hopewell[24,25], Niclas Eriksson[12], Marcus E Kleber[26,27], Kati Kristiansson[23], Per Lundmark[28], Leo-Pekka Lyytikäinen[29,30], Suzanne Rafelt[31], Dmitry Shungin[32–34], Rona J Strawbridge[20,21], Gudmar Thorleifsson[35], Emmi Tikkanen[36,37], Natalie Van Zuydam[38], Benjamin F Voight[39], Lindsay L Waite[40], Weihua Zhang[17], Andreas Ziegler[11], Devin Absher[40], David Altshuler[41–44], Anthony J Balmforth[45], Inês Barroso[1,46], Peter S Braund[31,47], Christof Burgdorf[48], Simone Claudi-Boehm[49], David Cox[50], Maria Dimitriou[51], Ron Do[41,43], DIAGRAM Consortium[52], CARDIOGENICS Consortium[52], Alex S F Doney[38], NourEddine El Mokhtari[53], Per Eriksson[20,21], Krista Fischer[18], Pierre Fontanillas[41], Anders Franco-Cereceda[54], Bruna Gigante[55], Leif Groop[56], Stefan Gustafsson[7], Jörg Hager[57], Göran Hallmans[58], Bok-Ghee Han[14], Sarah E Hunt[1], Hyun M Kang[59], Thomas Illig[60], Thorsten Kessler[48], Joshua W Knowles[5], Genovefa Kolovou[61], Johanna Kuusisto[62], Claudia Langenberg[63], Cordelia Langford[1], Karin Leander[55], Marja-Liisa Lokki[64], Anders Lundmark[28], Mark I McCarthy[3,65,66], Christa Meisinger[67], Olle Melander[56], Evelin Mihailov[19], Seraya Maouche[68], Andrew D Morris[38], Martina Müller-Nurasyid[69–72], MuTHER Consortium[52], Kjell Nikus[73], John F Peden[3], N William Rayner[3], Asif Rasheed[9], Silke Rosinger[74], Diana Rubin[53], Moritz P Rumpf[48], Arne Schäfer[75], Mohan Sivananthan[76,77], Ci Song[7], Alexandre F R Stewart[78,79], Sian-Tsung Tan[80], Gudmundur Thorgeirsson[81,82], C Ellen van der Schoot[83], Peter J Wagner[36,37], Wellcome Trust Case Control Consortium[52], George A Wells[78,79], Philipp S Wild[84,85], Tsun-Po Yang[1], Philippe Amouyel[86], Dominique Arveiler[87], Hanneke Basart[88], Michael Boehnke[59], Eric Boerwinkle[89], Paolo Brambilla[90], Francois Cambien[68], Adrienne L Cupples[91,92], Ulf de Faire[55], Abbas Dehghan[93], Patrick Diemert[94], Stephen E Epstein[95], Alun Evans[96], Marco M Ferrario[97], Jean Ferrières[98], Dominique Gauguier[3,99], Alan S Go[100], Alison H Goodall[31,47], Villi Gudnason[81,101], Stanley L Hazen[102], Hilma Holm[35], Carlos Iribarren[100], Yangsoo Jang[103], Mika Kähönen[104], Frank Kee[105], Hyo-Soo Kim[106], Norman Klopp[60], Wolfgang Koenig[107], Wolfgang Kratzer[108], Kari Kuulasmaa[23], Markku Laakso[62], Reijo Laaksonen[108], Ji-Young Lee[14], Lars Lind[28], Willem H Ouwehand[1,109,110], Sarah Parish[24,25], Jeong E Park[111], Nancy L Pedersen[7], Annette Peters[67,112], Thomas Quertermous[5], Daniel J Rader[113], Veikko Salomaa[23], Eric Schadt[114], Svati H Shah[115,116], Juha Sinisalo[117], Klaus Stark[118], Kari Stefansson[35,81], David-Alexandre Trégouët[68], Jarmo Virtamo[23], Lars Wallentin[12], Nicholas Wareham[63], Martina E Zimmermann[118], Markku S Nieminen[117], Christian Hengstenberg[118], Manjinder S Sandhu[1,63], Tomi Pastinen[119], Ann-Christine Syvänen[28], G Kees Hovingh[88], George Dedoussis[51], Paul W Franks[32–34,120], Terho Lehtimäki[29,30], Andres Metspalu[18,19], Pierre A Zalloua[121], Agneta Siegbahn[12], Stefan Schreiber[75], Samuli Ripatti[1,37], Stefan S Blankenberg[94], Markus Perola[23], Robert Clarke[24,25], Bernhard O Boehm[74], Christopher O'Donnell[93], Muredach P Reilly[122,126], Winfried März[26,123], Rory Collins[24,25,126], Sekar Kathiresan[41,124,125,126], Anders Hamsten[20,21,126], Jaspal S Kooner[80,126], Unnur Thorsteinsdottir[35,81,126], John Danesh[9,126], Colin N A Palmer[38,126], Robert Roberts[78,79,126], Hugh Watkins[3,4,126], Heribert Schunkert[48,126] & Nilesh J Samani[31,47,126]

[1]Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK. [2]Institut für Integrative und Experimentelle Genomik, Universität zu Lübeck, Lübeck, Germany. [3]Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK. [4]Cardiovascular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford, UK. [5]Department of Medicine, Stanford University School of Medicine, Stanford, California, USA. [6]Department of Health Sciences, University of Leicester, Leicester, UK. [7]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden. [8]Center for Non-Communicable Diseases, Karachi, Pakistan. [9]Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK. [10]Department of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA. [11]Institut für Medizinische Biometrie und Statistik, Universität zu Lübeck, Lübeck, Germany. [12]Uppsala Clinical Research Center, Uppsala University, Uppsala, Sweden. [13]Division of Cardiovascular and Neuronal Remodelling, Multidisciplinary Cardiovascular Research Centre, Leeds Institute of Genetics, Health and Therapeutics, University of Leeds, Leeds, UK. [14]Center for Genome Science, Korea National Institute of Health, Korea Center for Disease Control and Prevention, Yeonje-ri, Chungwon-gun, Korea. [15]Division of Cardiovascular Medicine, Department of Internal Medicine, University of Michigan, Ann Arbor, Michigan, USA. [16]Department of Human Genetics, University of Michigan, Ann Arbor, Michigan, USA. [17]Department of Epidemiology and Biostatistics, Imperial College London, London, UK. [18]Estonian Genome Center, University of Tartu, Tartu, Estonia. [19]Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia. [20]Atherosclerosis Research Unit, Department of Medicine, Karolinska Institutet, Stockholm, Sweden. [21]Center for Molecular Medicine, Karolinska University Hospital, Stockholm, Sweden. [22]Department of Twin Research and Genetic Epidemiology, King's College London, London, UK. [23]Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki, Finland. [24]Clinical Trial Service Unit, University of Oxford, Oxford, UK. [25]Epidemiological Studies Unit, University of Oxford, Oxford, UK. [26]Mannheim Institute of Public Health, Social and Preventive Medicine, Medical Faculty of Mannheim, University of Heidelberg, Mannheim, Germany. [27]Ludwigshafen Risk and Cardiovascular Health (LURIC) Study, Freiburg, Germany. [28]Department of Medical Sciences, Uppsala University, Uppsala, Sweden. [29]Department of Clinical Chemistry, Fimlab Laboratories, Tampere University Hospital, Tampere, Finland. [30]Department of Clinical Chemistry, University of Tampere School of Medicine, Tampere, Finland. [31]Department of Cardiovascular Sciences, University of Leicester, Glenfield Hospital, Leicester, UK. [32]Genetic & Molecular Epidemiology Unit, Department of Clinical Sciences, Lund University Diabetes Center, Skåne University Hospital, Malmö, Sweden. [33]Department of Public Health & Clinical Medicine, Genetic Epidemiology & Clinical Research Group, Section for Medicine, Umeå University, Umeå, Sweden. [34]Department of Odontology, Umeå University, Umeå, Sweden. [35]deCODE Genetics, Reykjavik, Iceland. [36]Institute for Molecular Medicine FIMM, University of Helsinki, Helsinki, Finland. [37]Public Health Genomics Unit, National Institute for Health and Welfare, Helsinki, Finland. [38]Medical Research Institute, University of Dundee, Ninewells Hospital and Medical School, Dundee, UK. [39]Department of Pharmacology, University of Pennsylvania, Philadelphia, Pennsylvania, USA. [40]HudsonAlpha Institute for Biotechnology, Huntsville, Alabama, USA. [41]Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA. [42]Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts, USA. [43]Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts, USA. [44]Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA. [45]Division of Cardiovascular and Diabetes Research, Multidisciplinary Cardiovascular Research Centre, Leeds Institute of Genetics, Health and Therapeutics, University of Leeds, Leeds, UK. [46]University of Cambridge Metabolic Research Laboratories, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, UK. [47]National Institute for Health Research (NIHR) Leicester Cardiovascular Biomedical Research Unit, Glenfield Hospital, Leicester, UK. [48]Deutsches Herzzentrum München, Technische Universität München, Munich, Germany. [49]Practice of Gynecology, Ulm University Medical Centre, Ulm, Germany. [50]Biotherapeutics and Bioinnovation Center, Pfizer, South San Francisco, California, USA. [51]Department of Dietetics–Nutrition, Harokopio University, Athens, Greece. [52]A list of members and affiliations appears in the **Supplementary Note**. [53]Klinik für Innere Medizin, Kreiskrankenhaus Rendsburg, Rendsburg, Germany. [54]Cardiothoracic Surgery Unit, Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden. [55]Division of Cardiovascular Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden. [56]Department of Clinical Sciences, Diabetes and Endocrinology, Lund University, University Hospital Malmö, Malmö, Sweden. [57]Commissariat à l'Energie Atomique (CEA)–Genomics Institute, National Genotyping Centre, Paris, France. [58]Department of Public Health & Clinical Medicine, Section for Nutritional Research, Umeå University, Umeå, Sweden. [59]Department of Biostatistics, Center for Statistical Genetics, University of Michigan, Ann Arbor, Michigan, USA. [60]Hannover Unified Biobank, Hannover Medical School, Hannover, Germany. [61]First Cardiology Department, Onassis Cardiac Surgery Center 356, Athens, Greece. [62]Department of Medicine, University of Eastern Finland and Kuopio University Hospital, Kuopio, Finland. [63]Medical Research Council (MRC) Epidemiology Unit, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, UK. [64]Transplantation Laboratory, Haartman Institute, University of Helsinki, Helsinki, Finland. [65]Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, UK. [66]Oxford NIHR Biomedical Research Centre, Churchill Hospital, Oxford, UK. [67]Institute of Epidemiology II, Helmholtz Zentrum München–German Research Center for Environmental Health, Neuherberg, Germany. [68]Institut National de la Santé et la Recherche Médicale (INSERM) Unité Mixte de Recherche (UMR) S937, Institute for Cardiometabolism and Nutrition (ICAN), Pierre and Marie Curie (Paris 6) University, Paris, France. [69]Department of Medicine I, University Hospital Grosshadern, Ludwig-Maximilians-Universität, Munich, Germany. [70]Chair of Epidemiology, Institute of Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany. [71]Chair of Genetic Epidemiology, Institute of Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany. [72]Institute of Genetic Epidemiology, Helmholtz Zentrum München–German Research Center for Environmental Health, Neuherberg, Germany. [73]Heart Centre, Department of Cardiology, Tampere University Hospital, Tampere, Finland. [74]Division of Endocrinology and Diabetes, Department of Internal Medicine, Ulm University Medical Centre, Ulm, Germany. [75]Institut für Klinische Molekularbiologie, Christian-Albrechts Universität, Kiel, Germany. [76]Division of Epidemiology, Multidisciplinary Cardiovascular Research Centre (MCRC) University of Leeds, Leeds, UK. [77]Leeds Institute of Genetics, Health and Therapeutics, University of Leeds, Leeds, UK. [78]University of Ottawa Heart Institute, Cardiovascular Research Methods Centre Ontario, Ottawa, Ontario, Canada. [79]Ruddy Canadian Cardiovascular Genetics Centre, Ottawa, Ontario, Canada. [80]National Heart and Lung Institute (NHLI), Imperial College London, Hammersmith Hospital, London, UK. [81]Faculty of Medicine, University of Iceland, Reykjavik, Iceland. [82]Department of Medicine, Landspitali University Hospital, Reykjavik, Iceland. [83]Department of Experimental Immunohematology, Sanquin, Amsterdam, The Netherlands. [84]Center for Thrombosis and Hemostasis, University Medical Center Mainz, Johannes Gutenberg University Mainz, Mainz, Germany. [85]Department of Medicine 2, University Medical Center Mainz, Johannes Gutenberg University Mainz, Mainz, Germany. [86]Institut Pasteur de Lille, INSERM U744, Université Lille Nord de France, Lille, France. [87]Department of Epidemiology and Public Health, EA3430, University of Strasbourg, Strasbourg, France. [88]Department of Vascular Medicine, Academic Medical Center, Amsterdam, The Netherlands. [89]Human Genetics Center, University of Texas Health Science Center, Houston, Texas, USA. [90]Department of Experimental Medicine, University of Milano–Bicocca, Monza, Italy. [91]Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts, USA. [92]National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, Massachusetts, USA. [93]Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands. [94]Clinic for General and Interventional Cardiology, University Heart Center Hamburg, Hamburg, Germany. [95]Cardiovascular Research Institute, Washington Hospital Center, Washington, DC, USA. [96]Centre for Public Health, The Queen's University of Belfast, Belfast, UK. [97]Research Centre for Epidemiology and Preventive Medicine (EPIMED), Department of Clinical and Experimental Medicine, University of Insubria, Varese, Italy. [98]Department of Cardiology, Toulouse University School of Medicine, Rangueil Hospital, Toulouse, France. [99]INSERM UMR S872, Cordeliers Research Centre, Paris, France. [100]Division of Research, Kaiser Permanente Northern California, Oakland, California, USA. [101]Icelandic Heart Association, Kopavogur, Iceland. [102]Lerner Research Institute, Cleveland Clinic, Cleveland, Ohio, USA. [103]Cardiology Division, Department of Internal Medicine, Cardiovascular Genome Center, Yonsei University, Seoul, Korea. [104]Department of Clinical Physiology, Tampere University Hospital and University of Tampere, Tampere, Finland. [105]UK Clinical Research Collaboration (UKCRC) Centre of Excellence for Public Health (Northern Ireland), Queen's University of Belfast, Belfast, UK. [106]Department of Internal Medicine, Cardiovascular Center, Seoul National University Hospital, Seoul, Korea. [107]Department of Internal Medicine II–Cardiology, Ulm University Medical Center, Ulm, Germany. [108]Science Center, Tampere University Hospital, Tampere, Finland. [109]Department of Haematology, University of Cambridge, Cambridge, UK. [110]National Health Service (NHS) Blood and Transplant, Cambridge, UK. [111]Division of Cardiology, Samsung Medical Center, Seoul, Korea. [112]Munich Heart Alliance, Munich, Germany. [113]Division of Translational Medicine and Human Genetics, Department of Medicine, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania, USA. [114]Institute for Genomics and Multiscale Biology, Department of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, New York, USA. [115]Center for Human Genetics, Department of Medicine, Duke University Medical Center, Durham, North Carolina, USA. [116]Division of Cardiology, Department of Medicine, Duke University Medical Center, Durham, North Carolina, USA. [117]Division of Cardiology, Department of Medicine, Helsinki University Central Hospital (HUCH), Helsinki, Finland. [118]Klinik und Poliklinik für Innere Medizin II, Regensburg, Germany. [119]Department of Human Genetics, McGill University, Montréal, Québec, Canada. [120]Department of Nutrition, Harvard School of Public Health, Boston, Massachusetts, USA. [121]Lebanese American University, Chouran, Beirut, Lebanon. [122]Cardiovascular Institute, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania, USA. [123]Synlab Academy, Mannheim, Germany. [124]Cardiology Division, Center for Human Genetic Research, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts, USA. [125]Cardiovascular Research Center, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts, USA. [126]These authors contributed equally to this work. Correspondence should be addressed to P. Deloukas (panos@sanger.ac.uk) or N.J.S. (njs@leicester.ac.uk).

ONLINE METHODS

Meta-analysis and combination of evidence across stages. Analyses were performed in each study (**Supplementary Table 1a**) to test the following comparisons: all CAD cases with all controls, adjusted for sex and age; male CAD cases with male controls, adjusted for age; female CAD cases with female controls, adjusted for age; early-onset CAD cases with early age of onset (≤50 years) with all CAD controls, adjusted for sex; late-onset CAD cases (>50 years) with all controls, adjusted for sex; and all myocardial infarction cases with all controls, adjusted for age and sex. Age was defined as the recruitment age for controls and the event age for cases. We used the additive genetic model and fixed-effect inverse variance–weighted meta-analysis. SNPs were excluded from the meta-analysis if present in <17 GWAS and/or Metabochip or <13 Metabochip stage 2 studies. Heterogeneity was evaluated using the Cochran's Q and I^2 statistics. For SNPs with non-significant heterogeneity (P for $Q > 0.01$), we report fixed-effect model results. For SNPs with significant heterogeneity (P for Q <0.01), we performed an outlier test comparing the results in each study with the average of all other studies. For outliers ($P < 0.01$ or no studies with data), we excluded the most extreme study and repeated the meta-analysis. If no outliers were detected, but heterogeneity was significant, we used a random-effects model that was also used for all SNPs with significant heterogeneity in stage 1. In stage 3, we used a fixed-effect inverse variance–weighted meta-analysis.

The combination of evidence across stage 1 and stage 2 meta-analysis results was performed using Fisher's combined P-values method; using two-sided P values from stage 1 and one-sided P values from stage 2 for all SNPs with consistent direction of effect across the two stages. We estimated the stage 1 and 2 combined effect sizes for SNPs in the known loci using a fixed-effect inverse variance–weighted meta-analysis. The combination of evidence across stages 1– 3 for the replication effort was performed using a sample size–weighted meta-analysis for the selected SNPs. All participants gave written consent for participation in genetic studies, and the protocol of each study was approved by the corresponding local research ethics committee or institutional review board.

False discovery rate. FDR control is an alternative approach to experiment-wise error rate control that allows for statistical multiple testing; identifying as many significantly associated SNPs as possible with a tolerable false positive burden. FDR analysis is useful for selecting extended panels of SNPs (and genes) for experiments on the basis of multiple signals (for example, pathway or network analyses) that are robust to contamination by a small number of false positive signals. However, given the specific design of the Metabochip array to include selected SNPs (for replication) with significant P values and several high-density regions (for fine mapping) associated with CAD and the other cardiometabolic traits, the number of SNPs significantly associated with CAD, as well as high LD, could bias the FDR analysis. Therefore, we excluded from the 79,138 SNPs with available stage 1 and 2 data all SNPs falling in a high-density region associated with CAD (**Tables 1** and **2**), as well as CAD risk SNPs associated at $P < 5 \times 10^{-8}$. Furthermore, we performed an LD-based SNP pruning of the remaining high-density regions ($r^2 < 0.2$). In total, 54,806 SNPs were included in the FDR analysis.

We combined stage 1 and 2 data as an inverse variance–weighted average, and P values were calculated by Wald test. SNPs selected because their stage 1 P values were below 0.01 had their combined P values adjusted. If p_0 is the P value used as the criterion for selection in stage 1, z_{12} is the standardized test statistic obtained by combining the stages (arbitrarily assumed to be positive) and s_1 and s_2 are the standard errors for the two stages, then the adjusted P value is the sum of two integrals representing the two tails in which the stage 1 result might fall. The first is:

$$I_1 = \int_{z_{12}}^{\infty} \int_{\Phi^{-1}(1-\frac{p_0}{2})}^{\infty} \frac{\Phi(u)\Phi\left(s_2\left[v\sqrt{\frac{1}{s_1^2}+\frac{1}{s_2^2}}-\frac{u}{s_1}\right]\right)}{p_0} s_2\sqrt{\frac{1}{s_1^2}+\frac{1}{s_2^2}}\,\partial u\,\partial v$$

and the second has the same form but is integrated from $-\infty$ to $\Phi^{-1}(p_0/2)$, where Φ is the cumulative normal function. To test the adjusted P values, a simulation was performed in which null SNPs were generated and selected in stage 1 on the basis of their P values. These were combined with random second-stage data simulated again assuming a null effect. The adjusted P values had the expected uniform distribution between zero and one, suitable for use in the FDR analysis.

FDR analysis was performed using QVALUE software. A natural cubic spline (with 4 degrees of freedom) was fitted to provide a smoothed estimate of the proportion of null P values ($\hat{\pi}_0$). A density histogram of the P values for the 54,806 SNPs is shown in **Supplementary Figure 7**. At FDR = 0.05, we obtained 138 SNPs that were combined with 73 independent SNPs from fine-mapping regions associated with CAD. The selection included the SNP with the lowest combined P value per fine-mapping region and all SNPs within these regions that met the 5% FDR criterion in a separate analysis and were unlinked ($r^2 < 0.2$). Finally, all SNPs reported in **Tables 1** and **2** were added to the set of 211 SNPs (5% FDR results and CAD fine-mapped regions), resulting in 153 independent SNPs (104 identified through the FDR analysis) at $r^2 < 0.2$, which were used for heritability calculations (**Supplementary Table 9**).

Heritability. Heritability estimates were calculated locus by locus using the multifactorial liability threshold model based on OR estimates that assume that the lead SNP at a locus accurately tags the disease-causing variant, as described in ref. 12. The calculations are based on a disease prevalence estimate of 5% and an estimate of 40% for the total heritability of coronary disease.

Expression analyses. We interrogated the 16 new (or proxy; $r^2 > 0.8$) CAD risk SNPs for *cis* eQTL expression in multiple tissues: the ASAP study[31] used tissue biopsies taken from patients undergoing carotid endartectomy (plaque $n = 117$) or valve surgery (liver $n = 152$, aorta media $n = 117$, aorta adventitia $n = 103$ and mammary artery $n = 88$). Expression data were generated using the Affymetrix HG-U133 plus 2.0 array (plaque) or the Affymetrix ST1.0 Exon array (liver, aorta and mammary artery); in the MuTHER study[32], RNA levels were measured in LCLs ($n = 826$), skin ($n = 705$) and fat biopsies ($n = 825$) from 850 female twins (one-third monozygotic and two-thirds dizygotic) from the TwinsUK resource using the Illumina HumanHT-12v3 array. We assessed genotype with gene expression associations, using an additive linear model (within a 1-Mb window); in Cardiogenics[5], monocytes and macrophages were collected from healthy subjects and individuals with CAD, and RNA was profiled with the Illumina Human Ref-8 array. eQTL analysis was undertaken in 459 healthy individuals from Cambridge, UK, using an additive linear model (1-Mb window); in the Massachusetts General Hospital study[33] of liver, omentum and subcutaneous adipose tissue among subjects undergoing Roux-en-Y gastric bypass surgery, eQTL analysis was performed with a linear regression model using a 1-Mb window.

In loci with significant *cis*-eQTL signal(s) ($P < 1 \times 10^{-4}$), we also identified the most strongly associated *cis*-eQTL SNP (eSNP) for the corresponding transcript and then performed conditional analyses, including in the regression model, with either the lead eSNP or the lead CAD-associated SNP. On the basis of the conditional analysis, we determined whether the same variant underlies both gene expression regulation and disease.

Finally, we interrogated the lead SNPs in the 16 new CAD susceptibility loci for allelic expression imbalance effects in LCLs, fibroblasts and monocytes ($n = 188$; Cardiogenics), as described in ref. 34.

Network analysis. Genes for network analysis were selected using 310 SNPs (88 SNPs in known and new CAD risk loci and 222 SNPs at FDR <10% and LD pruned to $r^2 \leq 0.7$). We first selected genes with an eQTL ($P \leq 1 \times 10^{-6}$) and then on the basis of physical proximity (included overlapping genes on opposite strands or at equal distance from the SNP; genes were considered within a 40-kb window centered on the SNP). Spliced ESTs and putative transcripts were not included. Network analysis was performed using the Ingenuity Pathway Analysis software tool (IPA; Ingenuity Systems). We considered molecules and/or relationships available in The IPA Knowledge Base for human, mouse or rat and set the confidence filter to experimentally observed or high (predicted). Networks were generated with a maximum size of 70 genes, allowing up to 10 networks. Molecules in the query set with recorded interactions were 'eligible' for network construction using the IPA algorithm. Networks were ranked according to their degree of relevance to the eligible molecules in the query data set. The score takes into account the number of eligible molecules in the network and its size, as well as the total number of eligible molecules

analyzed and the total number of molecules in the Ingenuity Knowledge Base that could potentially be included in the networks. The network score is based on the hypergeometric distribution and is calculated by right-tailed Fisher's exact test. The significance *P* value associated with enrichment of functional processes was calculated using the right-tailed Fisher's exact test by considering the number of query molecules that participate in that function and the total number of molecules that are known to be associated with that function in the Ingenuity Knowledge Base.

31. Folkersen, L. *et al.* Association of genetic risk variants with expression of proximal genes identifies novel susceptibility genes for cardiovascular disease. *Circ. Cardiovasc. Genet.* **3**, 365–373 (2010).
32. Grundberg, E. *et al.* Mapping *cis*- and *trans*-regulatory effects across multiple tissues in twins: the MuTHER Study. *Nat. Genet.* **44**, 1084–1089 (2012).
33. Schadt, E.E. *et al.* Mapping the genetic architecture of gene expression in human liver. *PLoS Biol.* **6**, e107 (2008).
34. Ge, B. *et al.* Global patterns of *cis* variation in human cells revealed by high-density allelic expression analysis. *Nat. Genet.* **41**, 1216–1222 (2009).

doi:10.1038/ng.2480 NATURE GENETICS

American Journal of Epidemiology Advance Access published December 19, 2012

American Journal of Epidemiology

DOI: 10.1093/aje/kws297

Systematic Reviews and Meta- and Pooled Analyses

Meta-Analysis Investigating Associations Between Healthy Diet and Fasting Glucose and Insulin Levels and Modification by Loci Associated With Glucose Homeostasis in Data From 15 Cohorts

Jennifer A. Nettleton*, Marie-France Hivert, Rozenn N. Lemaitre, Nicola M. McKeown, Dariush Mozaffarian, Toshiko Tanaka, Mary K. Wojczynski, Adela Hruby, Luc Djoussé, Julius S. Ngwa, Jack L. Follis, Maria Dimitriou, Andrea Ganna, Denise K. Houston, Stavroula Kanoni, Vera Mikkilä, Ani Manichaikul, Ioanna Ntalla, Frida Renström, Emily Sonestedt, Frank J. A. van Rooij, Stefania Bandinelli, Lawrence de Koning, Ulrika Ericson, Neelam Hassanali, Jessica C. Kiefte-de Jong, Kurt K. Lohman, Olli Raitakari, Constantina Papoutsakis, Per Sjogren, Kathleen Stirrups, Erika Ax, Panos Deloukas, Christopher J. Groves, Paul F. Jacques, Ingegerd Johansson, Yongmei Liu, Mark I. McCarthy, Kari North, Jorma Viikari, M. Carola Zillikens, Josée Dupuis, Albert Hofman, Genovefa Kolovou, Kenneth Mukamal, Inga Prokopenko, Olov Rolandsson, Ilkka Seppälä, L. Adrienne Cupples, Frank B. Hu, Mika Kähönen, André G. Uitterlinden, Ingrid B. Borecki, Luigi Ferrucci, David R. Jacobs, Jr., Stephen B. Kritchevsky, Marju Orho-Melander, James S. Pankow, Terho Lehtimäki, Jacqueline C. M. Witteman, Erik Ingelsson, David S. Siscovick, George Dedoussis, James B. Meigs, and Paul W. Franks

* Correspondence to Dr. Jennifer A. Nettleton, Division of Epidemiology, Human Genetics and Environmental Sciences, School of Public Health, University of Texas Health Science Center at Houston, 1200 Herman Pressler Drive, Suite E-641, Houston, TX 77030 (e-mail: jennifer.a.nettleton@uth.tmc.edu).

Initially submitted January 11, 2012; accepted for publication June 5, 2012.

Whether loci that influence fasting glucose (FG) and fasting insulin (FI) levels, as identified by genome-wide association studies, modify associations of diet with FG or FI is unknown. We utilized data from 15 US and European cohort studies comprising 51,289 persons without diabetes to test whether genotype and diet interact to influence FG or FI concentration. We constructed a diet score using study-specific quartile rankings for intakes of whole grains, fish, fruits, vegetables, and nuts/seeds (favorable) and red/processed meats, sweets, sugared beverages, and fried potatoes (unfavorable). We used linear regression within studies, followed by inverse-variance-weighted meta-analysis, to quantify 1) associations of diet score with FG and FI levels and 2) interactions of diet score with 16 FG-associated loci and 2 FI-associated loci. Diet score (per unit increase) was inversely associated with FG ($\beta = -0.004$ mmol/L, 95% confidence interval: −0.005, −0.003) and FI ($\beta = -0.008$ ln-pmol/L, 95% confidence interval: −0.009, −0.007) levels after adjustment for demographic factors, lifestyle, and body mass index. Genotype variation at the studied loci did not modify these associations. Healthier diets were associated with lower FG and FI concentrations regardless of genotype at previously replicated FG- and FI-associated loci. Studies focusing on genomic regions that do not yield highly statistically significant associations from main-effect genome-wide association studies may be more fruitful in identifying diet-gene interactions.

diabetes; dietary pattern; gene-environment interaction; glucose; insulin

Abbrevations; ARIC, Atherosclerosis Risk in Communities; CHARGE, Cohorts for Heart and Aging Research in Genomic Epidemiology; CHS, Cardiovascular Health Study; FG, fasting glucose; FG-GRS, FG-associated genetic risk score; FHS, Family Heart Study; FI, fasting insulin; Framingham, Framingham Generation 5 and Offspring Studies; GENDAI, Gene-Diet Attica Investigation on Childhood Obesity; GHRAS, Greek Health Randomized Aging Study; GLACIER, Gene-Lifestyle Interactions and Complex Traits in Elevated Disease Risk; GWAS, genome-wide association studies; Health ABC, Health, Aging and Body

Composition; InCHIANTI, Invecchiare in Chianti [Aging in the Chianti Area]; Malmö, Malmö Diet and Cancer Study (cardiovascular cohort); ln, natural log; MESA, Multi-Ethnic Study of Atherosclerosis; Rotterdam, Rotterdam Study; SD, standard deviation; SNP, single nucleotide polymorphism; THISEAS, The Hellenic Study of Interactions between SNPs and Eating in Atherosclerosis Susceptibility; ULSAM, Uppsala Longitudinal Study of Adult Men; Young Finns, Young Finns Study.

Recent technological and methodological advances have led to multiple genome-wide association studies (GWAS) of complex human diseases such as type 2 diabetes (1, 2) and related quantitative traits (3). While results of these meta-analyses have identified multiple loci with modest effect estimates, much of the heritability of the phenotypic traits remains unexplained. Thus, the utility of using genotypes at these loci to improve clinical practice remains unknown. Nevertheless, clinical and lay public awareness of health-related genomic technologies is growing (4–6). This raises important clinical and public health questions: Do lifestyle choices, like adhering to a healthier diet, offset genetic risks (7–10)? Does genetic variation within populations necessitate individualized health-promoting dietary recommendations?

A "healthy diet" can be characterized using many different approaches. One popular approach is to create a composite score that ranks individuals on the basis of their intakes of foods or nutrients that have been favorably or unfavorably associated with diseases or risk factors (11–13). The resulting scores, or "dietary patterns," capture the highly complex nature of diet, where multiple foods and their nutrient constituents are consumed—none in isolation (14–16). Public health recommendations based on dietary patterns are also more easily understood than nutrient-based recommendations, since they can be placed in context with a person's behavior. Although there are many methods for characterizing dietary patterns, healthier diets share several common characteristics that are correspondingly reflected in dietary recommendations across countries. Diets associated with lower risk of type 2 diabetes and lower numbers of cardiometabolic risk factors (17, 18) comprise largely plant foods (e.g., whole grains, fruits, vegetables) and plant and marine sources of fat (e.g., nuts and seeds, fatty fish) in exchange for red and processed meats, foods high in sugar and salt, and highly refined grains.

The purpose of this study was to 1) evaluate associations of a dietary pattern score with fasting glucose (FG) and fasting insulin (FI) levels and 2) evaluate whether genotypes at known loci associated with FG and FI (3) modify the associations of dietary pattern with FG and FI, using data from multiple US and European cohort studies.

MATERIALS AND METHODS

Cohorts

The present work is a collaboration of investigators from US and European epidemiologic cohort studies included in the Nutrition Working Group of the CHARGE (Cohorts for Heart and Aging Research in Genomic Epidemiology) Consortium (19). Table 1 provides descriptive information about the 15 studies included in this investigation (additional details have been published previously (19) and are given in Web Table 1 (available at http://aje.oxfordjournals.org/)). The analyses included the following cohort studies: the Atherosclerosis Risk in Communities (ARIC) Study, the Framingham Generation 5 and Offspring Studies (Framingham), the Rotterdam Study (Rotterdam), the Cardiovascular Health Study (CHS), the Gene-Diet Attica Investigation on Childhood Obesity (GENDAI), the Greek Health Randomized Aging Study (GHRAS), the Uppsala Longitudinal Study of Adult Men (ULSAM), Gene-Lifestyle Interactions and Complex Traits in Elevated Disease Risk (GLACIER), the Family Heart Study (FHS), the Health, Aging and Body Composition (Health ABC) Study, the Malmö Diet and Cancer Study (cardiovascular cohort) (Malmö), Invecchiare in Chianti [Aging in the Chianti Area] (InCHIANTI), the Multi-Ethnic Study of Atherosclerosis (MESA), the Young Finns Study (Young Finns), and the Hellenic Study of Interactions between SNPs and Eating in Atherosclerosis Susceptibility (THISEAS). All persons studied were free of type 2 diabetes (defined by self-reported diabetes, pharmacologic treatment for diabetes, or FG concentrations ≥7 mmol/L) when FG and FI levels were measured, consented to genetic research, and provided written informed consent. All studies' examination protocols were approved by local institutional review boards, and the procedures followed in each were in accordance with the ethical standards of the responsible institutional or regional committee on human experimentation or with the Helsinki Declaration of 1975 as revised in 1983.

Diet score

Diet was assessed via food frequency questionnaire (in 13 of 15 cohorts), food records kept for 7 consecutive days (in ULSAM), or two 1-day dietary recalls (in GENDAI) (19) (see also Web Table 2). The diet score was based on intakes of 9 food groups defined consistently for all studies. Whole grains, fish, fruits, vegetables, and nuts/seeds were designated as favorable foods, whereas red and processed meats, sweets, sugared beverages, and fried potatoes were designated as unfavorable (additional details are given in Web Table 2). Intakes of foods/beverages were estimated in servings per day for all cohorts, with the exception of the ULSAM cohort, where grams per day were used. Intake of each food group was categorized into quartiles and assigned ascending values (0, 1, 2, 3) for favorable foods and descending values (3, 2, 1, 0) for unfavorable foods. These values were summed to generate a diet score (range, 0–27 points), with higher scores representing healthier diets.

We selected food groups for inclusion in the score based on 1) country-specific dietary guidelines; 2) results of investigations of the associations of specific dietary factors

Table 1. Characteristics of Participants From 15 Cohort Studies Included in a Meta-Analysis of the Influence of Diet and Genotype on Fasting Glucose and Insulin Concentrations, CHARGE Consortium

Cohort	First Author, Year (Reference No.)	Region	Maximum Sample Size (*n*)[a]	Mean Age, years (SD)	% Female	Mean Fasting Glucose Level, mmol/L (SD)	Mean Fasting Insulin Level, ln-pmol/L (SD)
ARIC Study	ARIC Investigators, 1989 (33)	United States	8,591	54.2 (5.7)	53.7	5.47 (0.50)	4.07 (0.66)
CHS	Fried, 1991 (34)	United States	2,745	72.3 (5.4)	62.3	5.53 (0.52)	4.44 (0.43)
FHS	Higgins, 1996 (35)	United States	3,187	51.4 (13.6)	53.6	5.22 (0.54)	4.10 (0.57)
Framingham	Feinleib, 1975 (36) and Splansky, 2007 (37)	United States	5,795	46.0 (11.5)	54.7	5.19 (0.48)	3.30 (0.41)
GENDAI	Papoutsakis, 2007 (38)	Mediterranean	1,087	11.2 (0.7)	53.2	4.75 (0.48)	3.69 (0.54)
GHRAS	Kanoni, 2008 (39)	Mediterranean	856	71.8 (5.7)	71.2	5.83 (1.63)	3.76 (0.56)[b]
GLACIER	Renström, 2011 (40)	Northern Europe	15,146	52.0 (8.8)	60.7	5.37 (0.62)	3.72 (0.64)[b]
Health ABC Study[b]	Houston, 2008 (41)	United States	1,281	74.8 (2.9)	50.2	5.16 (0.55)	3.81 (0.53)
InCHIANTI	Ferrucci, 2000 (42)	Mediterranean	1,071	67.7 (15.8)	56.3	4.84 (0.61)	4.18 (0.53)
Malmö	Berglund, 1993 (43)	Northern Europe	3,679	57.8 (6.0)	59.0	5.53 (0.52)	3.62 (0.53)
MESA	Bild, 2002 (44)	United States	2,271	62.4 (10.3)	51.9	4.85 (0.56)	3.48 (0.61)
Rotterdam	Hofman, 2011 (45)	Northern Europe	2,303	71.9 (6.6)	58.7	5.50 (0.53)	4.10 (0.52)
THISEAS	Theodoraki, 2010 (46)	Mediterranean	598	55.9 (13.6)	48.5	5.31 (0.64)	3.96 (0.58)[b]
ULSAM	Hedstrand, 1975 (47)	Northern Europe	933	71.0 (19.2)	0.0	5.37 (0.56)	4.31 (0.53)
Young Finns	Raitakari, 2008 (48)	Northern Europe	1,746	37.7 (5.0)	56.2	5.25 (0.48)	3.70 (0.77)

Abbreviations: ARIC, Atherosclerosis Risk in Communities; CHARGE, Cohorts for Heart and Aging Research in Genomic Epidemiology; CHS, Cardiovascular Health Study; FHS, Family Heart Study; Framingham, Framingham Generation 5 and Offspring Studies; GENDAI, Gene-Diet Attica Investigation on Childhood Obesity; GHRAS, Greek Health Randomized Aging Study; GLACIER, Gene-Lifestyle Interactions and Complex Traits in Elevated Disease Risk; Health ABC, Health, Aging and Body Composition; InCHIANTI, Invecchiare in Chianti [Aging in the Chianti Area]; Malmö, Malmö Diet and Cancer Study (cardiovascular cohort); MESA, Multi-Ethnic Study of Atherosclerosis; Rotterdam, Rotterdam Study; SD, standard deviation; SNP, single nucleotide polymorphism; THISEAS, The Hellenic Study of Interactions between SNPs and Eating in Atherosclerosis Susceptibility; ULSAM, Uppsala Longitudinal Study of Adult Men; Young Finns, Young Finns Study.

[a] Maximum sample sizes for both outcomes are provided; exceptions include GHRAS, GLACIER, and THISEAS, where fewer observations were available for fasting insulin: 670, 917, and 258, respectively.

[b] Fasting glucose and fasting insulin levels were measured from baseline (year 1) samples; diet score variables were collected during year 2; and covariates (see Web Table 3) were measured at baseline (year 1).

and patterns with diabetes and its risk factors; 3) data availability in participating cohorts; 4) regional food usage patterns; and 5) food product consistency across cohorts. We also inspected correlations among these food groups and others not included in the final score in order to evaluate whether the consumption patterns (correlation matrices) were similar across cohorts. We did not include dairy foods in the score, despite the fact that dairy foods, particularly reduced-fat dairy foods, are endorsed by most national dietary guidelines. Evidence regarding beneficial effects of dairy food consumption (reduced-fat vs. whole-fat dairy foods) on the risk of impaired glucose control or type 2 diabetes is inconclusive (11, 20–23), and product composition (particularly in terms of fat percentage, fermentation/culturing processes, and amounts of added sugar—all important factors in this context) is highly variable across regions represented in the present work (criterion 5 above). When information about preparation was available, we excluded fried fish from the fish food group. We did not include baked, boiled, or mashed white potatoes in the total for vegetables, since high intake of white potatoes has been associated with greater risk of type 2 diabetes (24) and may reflect a Western animal-based diet in some cohorts. However, we also did not include white potatoes as an unfavorable food, since white potatoes represent an important component of a healthy Mediterranean-style diet in some cohorts. While some guidelines emphasize replacing animal sources of protein with plant sources of protein such as legumes, we did not include legumes in our diet score calculation because 1) intakes were low and 2) legumes are commonly consumed with meat products (e.g., pork and beans, meat chili), particularly in the United States, rather than as a meat substitute, as is more common in Mediterranean regions (11, 25). Lastly, we did not include poultry in the score because of the absence of compelling evidence linking poultry intake to glucose regulation or diabetes and the positive correlation between poultry and meat consumption in most cohorts.

Genetic loci

We selected the 16 loci associated with FG and the 2 loci associated with FI that met the criteria for genome-wide significance in a previous meta-analysis of GWAS (3) (allele frequencies and effect sizes are shown in Web Table 3, genotyping methods in Web Table 4). We also created an FG-associated genetic risk score (FG-GRS) by summing the number of risk alleles for each participant across the FG-associated single nucleotide polymorphisms (SNPs), theoretically ranging in most cohorts from 0 (no FG-raising alleles) to 32 (homozygous for the FG-raising allele at each of the 16 SNPs). In GENDAI and GHRAS, we calculated the FG-GRS for 14 of the 16 FG-associated SNPs, since neither rs11558471 (*SLC30A8*) nor rs4506565 (*TCF7L2*) was genotyped in these cohorts. In Malmö, we included persons with more than 60% of the SNPs genotyped and then imputed the missing genotypes (with the most common genotype). Results were similar when the 3 cohorts with missing genotype information were not included in FG-GRS-related analyses; thus, to preserve sample size, we included all 15 cohorts.

FG and FI concentrations and other characteristics

FG and FI concentrations were quantified for each cohort using enzymatic methods and radioimmunoassays, respectively (cohort-specific methods are shown in Web Table 4). We statistically analyzed FG values without transformation; because FI data were not normally distributed, FI values were natural log (ln)-transformed before statistical analysis. We present beta coefficients from regression analyses for (ln)FI. Measurement methods for other relevant covariates (listed below) are described in Web Table 5.

Statistical analyses

Cohort-specific analyses. For each cohort, we calculated associations between diet score and FG and FI concentrations using multivariable linear regression, with the diet score modeled as a continuous variable. Analyses were performed at each study center according to a standardized analytic plan. Model 1 adjusted for energy intake (kcal/day), age, sex, field center (in Health ABC, CHS, ARIC, FHS, InCHIANTI, and MESA), and population and/or family substructure (using principal components analysis for relevant cohorts, in CHS, FHS, Framingham, MESA, and Young Finns). Model 2 included further adjustment for smoking, physical activity, education, and alcohol consumption (defined within each cohort; described in Web Table 5). Model 3 further adjusted for body mass index (weight (kg)/height $(m)^2$). Regression coefficients from these models represent the predicted difference in FG (mmol/L) or FI (ln-pmol/L) concentration per 1-unit increase in diet score. For each cohort, we also assessed associations of the FG-GRS, the 16 FG-associated SNPs, and the 2 FI-associated SNPs with respective FG and FI outcomes, adjusting for age, sex, field center, and/or population substructure. An additive genetic model was used, consistent with the association pattern for these loci (3).

Our primary interaction tests of interest were between diet score and the FG-GRS (FG outcome) and the 2 FI-associated SNPs (FI outcome); interactions between diet score and the 16 individual SNPs making up the FG-GRS were secondary, exploratory analyses. To test these interactions, we included a diet score × FG-GRS (or SNP) cross-product term along with model 1 covariates. The resulting interaction regression coefficients represent the difference in the magnitude of the healthy diet association (per 1-unit increase in score) with FG (mmol/L) or FI (ln-pmol/L) concentration per copy of an FG- or FI-raising allele.

Meta-analyses. We used an inverse-variance-weighted meta-analysis with fixed effects, employing the rmeta package (version 2.16) in R 2.13.1 (http://www.R-project.org/), for diet score-outcome associations and METAL (http://www.sph.umich.edu/csg/abecasis/Metal/index.html) for SNP-outcome associations and diet score × SNP interactions. Heterogeneity was assessed by means of the I^2 index (26). Figures were generated with Stata 11.0 (Stata Corporation, College Station, Texas).

Sample sizes for the associations of diet score with FG concentration ranged from 48,787 to 51,289 (in models 3 and 1, respectively); and for FI, sample sizes ranged from

Table 2. Distribution of Diet Score and Food Group Components Across the 15 Cohort Studies Included in a Meta-Analysis of the Influence of Diet and Genotype on Fasting Glucose and Insulin Concentrations, CHARGE Consortium

Cohort	Diet Score[a]		Diet Score Favorable Food Groups, servings/day[b] (SD)					Diet Score Unfavorable Food Groups, servings/day[b] (SD)			
	Mean	Range	Whole Grains	Fish	Fruit	Vegetables	Nuts and Seeds	Red and Processed Meats	Desserts and Sweets	Sugar-sweetened Beverages	Fried Potatoes
ARIC Study	13.7	1–27	1.37 (1.25)	0.26 (0.28)	1.50 (1.25)	1.72 (1.15)	0.42 (0.58)	1.00 (0.72)	1.45 (1.37)	0.47 (0.88)	0.11 (0.16)
CHS	13.7	1–27	1.02 (0.65)	0.32 (0.30)	2.74 (1.47)	2.83 (1.49)	0.20 (0.25)	0.69 (0.58)	0.85 (0.39)	0.14 (0.26)	0.09 (0.14)
FHS	13.1	1–26	1.58 (1.55)	0.20 (0.22)	1.56 (1.38)	1.61 (1.27)	0.34 (0.57)	2.46 (1.21)	1.64 (1.44)	0.67 (1.1)	0.13 (0.2)
Framingham	13.2	1–26	1.17 (1.13)	0.27 (0.27)	1.26 (1.16)	2.80 (1.92)	0.37 (0.53)	0.77 (0.59)	1.41 (1.32)	0.48 (0.86)	0.10 (0.13)
GENDAI	9.5	1–20	0.41 (0.77)	0.21 (0.60)	1.21 (1.35)	1.08 (1.21)	NA	1.69 (1.36)	0.85 (0.85)	0.43 (0.62)	NA
GHRAS	11.0	2–19	1.06 (1.53)	0.33 (0.20)	2.15 (1.38)	1.53 (0.57)	NA	0.43 (0.34)	1.24 (1.09)	0.10 (0.17)	NA
GLACIER	11.9	0–24	2.71 (1.47)	0.17 (0.13)	1.59 (1.17)	1.61 (1.21)	NA	0.62 (0.31)	1.52 (1.34)	0.32 (0.49)	0.10 (0.13)
Health ABC Study[c]	15.7	3–27	1.01 (0.71)	0.16 (0.15)	1.88 (1.10)	2.00 (0.99)	0.32 (0.38)	0.71 (0.49)	1.40 (0.86)	0.05 (0.18)	0.07 (0.10)
InCHIANTI	10.6	2–20	0.20 (0.68)	0.22 (0.17)	2.83 (1.38)	2.58 (1.34)	0.02 (0.06)	1.03 (0.52)	2.41 (1.57)	0.08 (0.26)	NA
Malmö	13.7	1–26	1.92 (1.83)	0.55 (0.40)	2.02 (1.23)	2.50 (1.32)	0.07 (0.18)	1.45 (0.73)	3.30 (1.95)	0.29 (0.54)	0.18 (0.29)
MESA	13.6	1–27	0.69 (0.58)	0.15 (0.19)	1.89 (1.48)	2.23 (1.32)	0.35 (0.50)	0.53 (0.46)	1.37 (1.46)	0.41 (0.82)	0.11 (0.15)
Rotterdam	10.2	1–20	2.88 (1.50)	0.21 (0.18)	2.11 (1.16)	2.15 (0.94)	0.29 (0.51)	1.38 (0.63)	1.60 (1.13)	0.51 (0.66)	NA
THISEAS	11.9	0–26	1.37 (1.60)	0.50 (0.47)	1.66 (1.51)	3.53 (3.10)	0.59 (1.03)	1.20 (1.18)	0.94 (1.18)	0.34 (0.71)	0.26 (0.47)
ULSAM[d]	13.4	3–24	19.8 (13.3)	18.7 (13.7)	116.4 (99.2)	68.8 (53.9)	0.34 (5.27)	72.0 (27.9)	63.8 (53.5)	39.4 (88.2)	12.0 (16.6)
Young Finns	13.6	2–27	3.25 (1.88)	0.40 (0.32)	2.05 (2.10)	3.39 (2.26)	0.04 (0.08)	1.22 (0.78)	1.47 (1.26)	0.54 (0.90)	0.17 (0.19)

Abbreviations: ARIC, Atherosclerosis Risk in Communities; CHARGE, Cohorts for Heart and Aging Research in Genomic Epidemiology; CHS, Cardiovascular Health Study; FHS, Family Heart Study; Framingham, Framingham Generation 5 and Offspring Studies; GENDAI, Gene-Diet Attica Investigation on Childhood Obesity; GHRAS, Greek Health Randomized Aging Study; GLACIER, Gene-Lifestyle Interactions and Complex Traits in Elevated Disease Risk; Health ABC, Health, Aging and Body Composition; InCHIANTI, Invecchiare in Chianti [Aging in the Chianti Area]; Malmö, Malmö Diet and Cancer Study (cardiovascular cohort); MESA, Multi-Ethnic Study of Atherosclerosis; NA, not applicable; Rotterdam, Rotterdam Study; SD, standard deviation; SNP, single nucleotide polymorphism; THISEAS, The Hellenic Study of Interactions between SNPs and Eating in Atherosclerosis Susceptibility; ULSAM, Uppsala Longitudinal Study of Adult Men; Young Finns, Young Finns Study.

[a] Summed quartile ranks of favorable food groups (0, 1, 2, and 3, for lowest to highest quartiles, respectively) and reversed quartile ranks of unfavorable food groups (3, 2, 1, and 0, for lowest to highest quartiles, respectively). The theoretical range is 0–27 points, with a high score representing the healthiest diet based on the selected parameters.

[b] Servings/day except for ULSAM (see footnote d).

[c] Fasting glucose and fasting insulin levels were measured from baseline (year 1) samples; diet score variables were collected during year 2; and covariates (see Web Table 3) were measured at baseline (year 1).

[d] In ULSAM, data were collected by 7-day food record and are characterized in g/day. See Web Table 1 for individual foods and beverages included within each food group.

Table 3. Associations of Healthy Diet Score With Fasting Glucose and Fasting Insulin Concentrations in a Meta-Analysis, CHARGE Consortium

	No. of Persons	β (SE)[a]	I^2, %	95% Confidence Interval
Fasting glucose, mmol/L				
Model 1[b]	51,289	−0.005 (0.0005)*	62.6	34.6, 78.6
Model 2[c]	48,902	−0.004 (0.0005)*	54.2	18, 74.5
Model 3[d]	48,787	−0.004 (0.0005)*	22.1	0, 57.8
Fasting insulin, ln-pmol/L				
Model 1	35,907	−0.010 (0.0006)*	71.3	51.7, 83
Model 2	34,415	−0.009 (0.0007)*	45.2	0, 70
Model 3	34,305	−0.008 (0.0005)*	12.3	0, 50.2

Abbreviations: CHARGE, Cohorts for Heart and Aging Research in Genomic Epidemiology; SE, standard error.

* $P < 0.0001$.

[a] Beta coefficient and standard error for the estimated difference in fasting glucose (mmol/L) or fasting insulin (ln-pmol/L) concentration per 1-unit increase in diet score.

[b] Model 1 adjusted for age, sex, energy intake, field center (in the Health, Aging and Body Composition Study, the Cardiovascular Health Study, the Atherosclerosis Risk in Communities Study, the Family Heart Study, Invecchiare in Chianti, and the Multi-Ethnic Study of Atherosclerosis), and population substructure (by principal components in the Cardiovascular Heath Study, the Family Heart Study, the Framingham Generation 5 and Offspring Studies, the Multi-Ethnic Study of Atherosclerosis, and the Young Finns Study).

[c] Model 2 adjusted for model 1 covariates plus highest attained educational level, smoking status, physical activity level, and alcohol intake (see Web Table 4 for cohort-specific definitions).

[d] Model 3 adjusted for the model 2 covariates plus body mass index.

34,305 to 35,907 (in models 3 and 1, respectively). Sample sizes for the interaction analyses for FG ranged from 48,872 for rs10830963 (*MTNR1B*) to 51,377 for rs2191349 (*KB/TMEM*), with sample sizes for the other 14 SNPs falling between those values. The sample size for interaction with the FG-GRS was 51,063. Sample sizes for the interaction analyses for FI were 35,739 for rs35767 (*IGF1*) and 35,991 for rs78094 (*GCKR*).

We defined the level of statistical significance on the basis of a Bonferroni correction: $P < 0.025$ for associations between the diet score and the two outcomes of interest; $P < 0.017$ for primary tests of interaction with the FG-GRS and the 2 FI-associated SNPs; and $P < 0.003$ for exploratory tests of interaction with each of the 16 individual FG-associated SNPs. Estimates of statistical power for various sample- and effect-size combinations are published elsewhere (19).

RESULTS

Mean values for the food groups comprising the diet score within each study are shown in Table 2. Variation in values followed expected regional differences in food consumption (19) but did not appear to relate to type of dietary assessment tool, age of the cohort, or chronologic years of dietary assessment.

Diet score was inversely associated with both FG and FI concentrations (Table 3); that is, healthier diets were associated with lower FG and FI concentrations. While the associations were not statistically significant within all cohorts, 12 of 15 cohorts showed inverse associations between diet score and FG (Figure 1), and all 15 cohorts showed inverse associations between diet score and FI (Figure 2). The meta-analyzed associations were robust; adjustment for demographic factors (model 1), lifestyle factors (model 2), and body mass index (model 3), a potential mediator of the relation between diet and health outcomes, had no material impact on the strength or magnitude of the effect estimates ($P < 0.0001$ for all) (Table 3). For each 5-unit change (approximately equal to the mean standard deviation (SD)) in diet score (pointing towards a healthier diet), FG concentrations were 0.03 mmol/L lower and FI concentrations were 0.05 ln-pmol/L lower (results were derived from model 3 regression coefficients; see Table 3 and Figures 3 and 4).

The previously published associations between 16 SNPs and FG and between 2 SNPs and FI were observed in the present collection of cohorts (10 of our 15 cohorts (or 53% of our sample size based on individuals) contributed to the original 54-cohort collaboration (3)) (Table 4). Effect sizes for FG ranged from a 0.01-mmol/L greater to a 0.08-mmol/L greater FG concentration per FG-raising allele, and the effect for FI was a 0.02-(ln)pmol/L increase per FI-raising allele, similar to values reported in the earlier meta-analysis (Table 4) (3). The FG-GRS was also significantly associated with FG concentrations in the present meta-analysis: For each additional FG-GRS unit (risk allele), FG concentrations were 0.03 mmol/L greater ($\beta = 0.03$ (standard error, 0.001), $P < 0.0001$; Table 4).

We observed no interactions between diet score and the FG-GRS, the 16 individual FG-associated SNPs, or the 2 FI-associated SNPs in our meta-analysis (Table 5; Web Figures 1–19). Within some cohorts, there were statistically

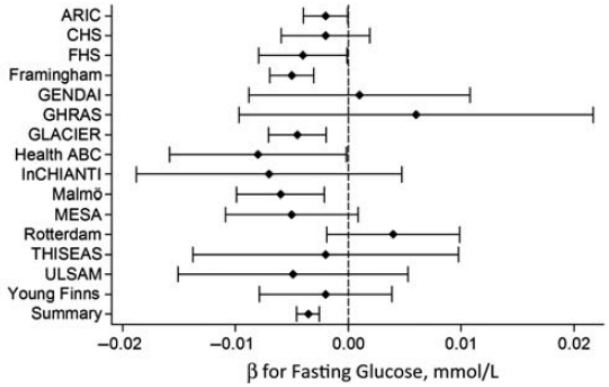


Figure 1. Associations between diet score and fasting glucose concentration in a meta-analysis of data from 15 cohort studies, CHARGE (Cohorts for Heart and Aging Research in Genomic Epidemiology) Consortium. Regression coefficients (β) for each of the 15 cohorts and the total association, summarized across all 15 cohorts, represent the difference in fasting glucose level (mmol/L) per 1-unit increase in diet score after adjustment for the model 3 covariates: energy intake, age, sex, field center (in Health ABC, CHS, ARIC, FHS, and InCHIANTI), population substructure (by principal components in CHS, FHS, Framingham, MESA, and Young Finns), smoking, physical activity level, highest attained educational level, alcohol consumption, and body mass index. Bars, 95% confidence interval. (ARIC, Atherosclerosis Risk in Communities Study; CHS, Cardiovascular Health Study, FHS, Family Heart Study; Framingham, Framingham Generation 5 and Offspring Studies; GENDAI, Gene-Diet Attica Investigation on Childhood Obesity; GHRAS, Greek Health Randomized Aging Study; GLACIER, Gene-Lifestyle Interactions and Complex Traits in Elevated Disease Risk; Health ABC, Health, Aging and Body Composition Study; InCHIANTI, Invecchiare in Chianti [Aging in the Chianti Area]; Malmö, Malmö Diet and Cancer Study (cardiovascular cohort); MESA, Multi-Ethnic Study of Atherosclerosis; Rotterdam, Rotterdam Study; SNP, single nucleotide polymorphism; THISEAS, The Hellenic Study of Interactions between SNPs and Eating in Atherosclerosis Susceptibility; ULSAM, Uppsala Longitudinal Study of Adult Men; Young Finns, Young Finns Study).

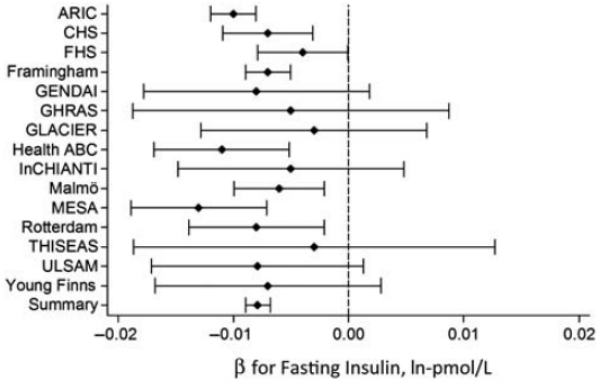


Figure 2. Associations between diet score and fasting insulin concentration in a meta-analysis of data from 15 cohort studies, CHARGE (Cohorts for Heart and Aging Research in Genomic Epidemiology) Consortium. Regression coefficients (β) for each of the 15 cohorts and the total association, summarized across all 15 cohorts, represent the difference in fasting insulin level (ln-pmol/L) per 1-unit increase in diet score after adjustment for the model 3 covariates: energy intake, age, sex, field center (in Health ABC, CHS, ARIC, FHS, and InCHIANTI), population substructure (by principal components in CHS, FHS, Framingham, MESA, and Young Finns), smoking, physical activity level, highest attained educational level, alcohol consumption, and body mass index. Bars, 95% confidence interval. (ARIC, Atherosclerosis Risk in Communities Study; CHS, Cardiovascular Health Study, FHS, Family Heart Study; Framingham, Framingham Generation 5 and Offspring Studies; GENDAI, Gene-Diet Attica Investigation on Childhood Obesity; GHRAS, Greek Health Randomized Aging Study; GLACIER, Gene-Lifestyle Interactions and Complex Traits in Elevated Disease Risk; Health ABC, Health, Aging and Body Composition Study; InCHIANTI, Invecchiare in Chianti [Aging in the Chianti Area]; Malmö, Malmö Diet and Cancer Study (cardiovascular cohort); MESA, Multi-Ethnic Study of Atherosclerosis; Rotterdam, Rotterdam Study; SNP, single nucleotide polymorphism; THISEAS, The Hellenic Study of Interactions between SNPs and Eating in Atherosclerosis Susceptibility; ULSAM, Uppsala Longitudinal Study of Adult Men; Young Finns, Young Finns Study).

significant interactions ($P < 0.05$), but these were inconsistent across cohorts with respect to loci (Web Figures 2–19) and were probably false-positive findings owing to the number of tests performed. Removing the youngest cohort (GENDAI, where the mean age was 11.2 years) did not change these conclusions. Inspection of the data according to the mean age of the cohorts (e.g., <70 years (10 cohorts) vs. ≥70 years (5 cohorts) at assessment) also revealed no consistent differences in the direction of interaction regression coefficients (Web Figures 1–19). Results from random-effects meta-analyses conducted on all of these data were not materially different from those of the fixed-effects meta-analysis (data not shown).

DISCUSSION

Using data from 15 well-characterized epidemiologic cohorts comprising US and European subjects without known diabetes, we observed favorable associations between adherence to a healthy diet, as reflected by the diet score, and FG and FI concentrations. These associations were not modified by genotype at loci previously shown to be reliably associated with glucose homeostasis, such that the association between a healthy diet and FG homeostasis was maintained independently of genotype at these loci. Thus, these data suggest that adhering to a healthy diet is important for everyone, regardless of genotype at these loci.

Although diet quality, as reflected by the diet score, did not modify the associations of the selected loci with FG and FI per se, a risk-allele carrier who adheres to a healthier diet would have lower FG and FI levels than a risk-allele carrier with a less healthy diet. Moreover, our data raise the possibility that modest differences in diet quality (towards a healthier diet) might offset the small genetic risk associated with common variants related to glucose homeostasis. For example, the average effect size across all 16 FG-raising alleles was approximately 0.03 mmol/L (a 0.03-mmol/L greater FG level) per allele, which compares in magnitude to an approximately 1.5-SD greater diet score (i.e., towards a healthier diet). More strikingly, the average 0.02-(ln)pmol/L

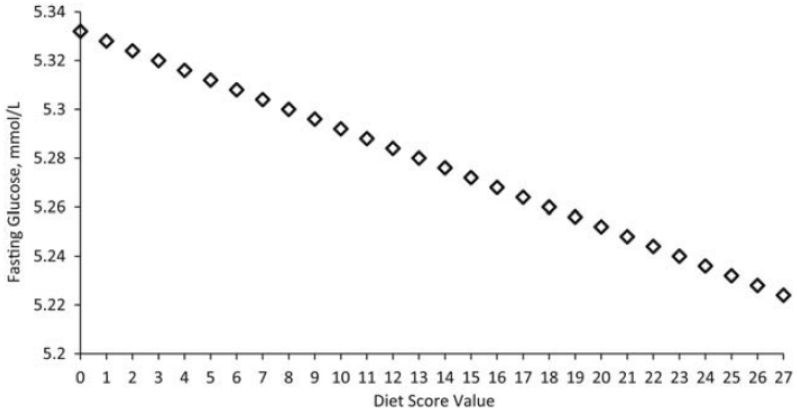


Figure 3. Predicted fasting glucose concentration according to diet score in a meta-analysis of data from 15 cohort studies, CHARGE (Cohorts for Heart and Aging Research in Genomic Epidemiology) Consortium. The graph shows predicted fasting glucose concentrations across the spectrum of possible diet score values (0–27), where a diet score of 13 is set to the across-cohorts mean fasting glucose level (5.28 mmol/L), fasting glucose concentrations are 0.004 mmol/L (the regression coefficient generated from model 3) lower per successively higher diet score value, and fasting glucose concentrations are 0.004 mmol/L higher per successively lower diet score value. The model 3 covariates included energy intake, age, sex, field center (in the Health, Aging and Body Composition Study, the Cardiovascular Health Study, the Atherosclerosis Risk in Communities Study, the Family Heart Study, and Invecchiare in Chianti), population substructure (by principal components in the Cardiovascular Heath Study, the Family Heart Study, the Framingham Generation 5 and Offspring Studies, the Multi-Ethnic Study of Atherosclerosis, and the Young Finns Study), smoking, physical activity level, highest attained educational level, alcohol consumption, and body mass index.

greater FI level per FI-raising allele compares in magnitude to a ½-SD greater diet score.

Our observation that a healthy diet was cross-sectionally associated with lower FG and FI levels is consistent with previous studies of dietary pattern indexes and specific food groups comprising our healthy diet score (18, 19, 27–29). Our work demonstrates that dietary data can be coordinated sufficiently across studies from diverse regions to create a

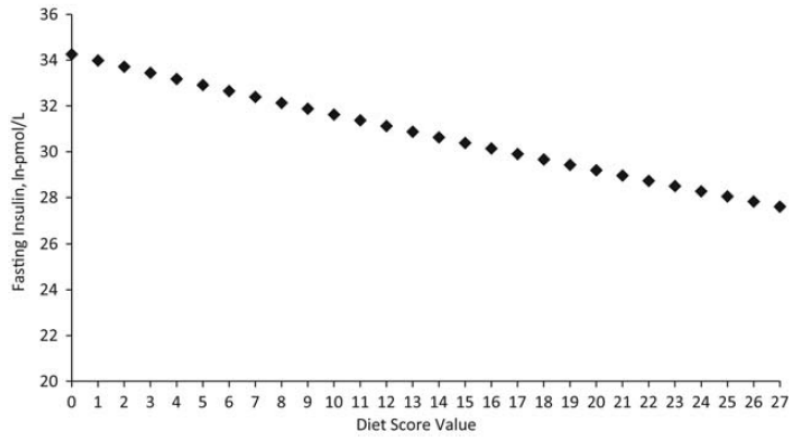


Figure 4. Predicted fasting insulin concentration according to diet score in a meta-analysis of data from 15 cohort studies, CHARGE (Cohorts for Heart and Aging Research in Genomic Epidemiology) Consortium. The graph shows predicted fasting insulin concentrations across the spectrum of possible diet score values (0–27), where a diet score of 13 is set to the across-cohorts mean fasting insulin level (3.43 ln-pmol/L), fasting insulin concentrations are 0.008 ln-pmol/L (the regression coefficient generated from model 3) lower per successively higher diet score value, and fasting insulin concentrations are 0.008 ln-pmol/L higher per successively lower diet score value. The plotted values are the result of exponentiating the ln-pmol/L estimates. The model 3 covariates included energy intake, age, sex, field center (in the Health, Aging and Body Composition Study, the Cardiovascular Health Study, the Atherosclerosis Risk in Communities Study, the Family Heart Study, and Invecchiare in Chianti), population substructure (by principal components in the Cardiovascular Heath Study, the Family Heart Study, the Framingham Generation 5 and Offspring Studies, the Multi-Ethnic Study of Atherosclerosis, and the Young Finns Study), smoking, physical activity level, highest attained educational level, alcohol consumption, and body mass index.

Table 4. Associations[a] of Single Nucleotide Polymorphisms and Fasting Glucose Genetic Risk Score With Fasting Glucose and Fasting Insulin Concentrations in a Meta-Analysis, CHARGE Consortium

SNP	Chromosome	Nearest Gene	Effect Allele[b]	Other Allele	No. of Cohorts	Summary Association	
						β (SE)[c]	*P* Value
Fasting glucose-related loci							
rs340874	1	*PROX1*	C	T	15	0.0198 (0.0034)	8.31×10^{-09}
rs560887	2	*G6PC2*	C	T	15	0.0735 (0.0036)	3.05×10^{-90}
rs780094	2	*GCKR*	C	T	15	0.0298 (0.0034)	3.64×10^{-18}
rs11708067	3	*ADCY5*	C	T	15	0.0305 (0.0041)	1.28×10^{-13}
rs11920090	3	*SLC2A2*	T	A	15	0.0318 (0.0048)	4.88×10^{-11}
rs4607517	7	*GCK*	A	G	15	0.0612 (0.0045)	1.71×10^{-42}
rs2191349	7	*DGKB/TMEM195*	T	G	15	0.0248 (0.0033)	1.25×10^{-13}
rs11558471	8	*SLC30A8*	A	G	13[d]	0.0360 (0.0039)	2.88×10^{-20}
rs7034200	9	*GLIS3*	A	C	15	0.0180 (0.0033)	6.66×10^{-08}
rs10885122	10	*ADRA2A*	G	T	15	0.0200 (0.0053)	1.60×10^{-04}
rs4506565	10	*TCF7L2*	C	T	13[d]	0.0276 (0.0038)	4.79×10^{-13}
rs10830963	11	*MTNR1B*	G	C	15	0.0801 (0.0040)	2.80×10^{-91}
rs7944584	11	*MADD*	A	T	15	0.0249 (0.0038)	6.87×10^{-11}
rs11605924	11	*CRY2*	A	C	15	0.0239 (0.0034)	1.03×10^{-12}
rs174550	11	*FADS1*	T	C	15	0.0189 (0.0035)	8.52×10^{-08}
rs11071657	15	*FAM148B*	A	G	15	0.0071 (0.0035)	4.59×10^{-02}
FG-GRS					15	0.0283 (0.0009)	<0.0001
Fasting insulin-related loci							
rs780094	2	*GCKR*	C	T	15	0.0168 (0.0036)	3.70×10^{-06}
rs35767	12	*IGF1*	G	A	15	0.0153 (0.0048)	1.44×10^{-03}

Abbreviations: CHARGE, Cohorts for Heart and Aging Research in Genomic Epidemiology; FG-GRS, fasting glucose genetic risk score; SE, standard error; SNP, single nucleotide polymorphism.

[a] Adjusted for age, sex, field center (in the Health, Aging and Body Composition Study, the Cardiovascular Health Study, the Atherosclerosis Risk in Communities Study, the Family Heart Study, Invecchiare in Chianti, and the Multi-Ethnic Study of Atherosclerosis), and population substructure (by principal components in the Cardiovascular Heath Study, the Family Heart Study, the Framingham Generation 5 and Offspring Studies, the Multi-Ethnic Study of Atherosclerosis, and the Young Finns Study).

[b] Coded uniformly in all cohorts. Allele frequencies for each cohort can be found in Web Table 2.

[c] Beta coefficient and standard error for the estimated difference in fasting glucose (mmol/L) or fasting insulin (ln-pmol/L) concentration per 1-unit increase in the effect allele (the fasting glucose- or fasting insulin-raising allele), assuming an additive genetic model.

[d] SNP not genotyped in this cohort.

meaningful, predictive dietary score reflecting healthy food consumption.

We chose to test for interaction between diet and a select set of loci from a previous meta-analysis of GWAS for 3 main reasons: 1) the loci (or the functional SNPs these variants tag) from such studies are those with the strongest evidence of biologic relevance to glucose and insulin homeostasis; 2) these loci are the ones most likely to be used by medical practitioners (and consumers) for prognostic purposes (4–6); and 3) given the penalty for multiple testing and the probably small effect sizes for diet-gene interaction, an analysis focused on fewer SNPs would be more statistically powerful. Thus, we aimed to assess whether the effects of GWAS-identified loci altered the favorable associations of dietary factors with glucose and insulin homeostasis. Our finding of no interaction is important, as it suggests that the favorable influences attributed to dietary factors are likely to be conveyed regardless of genotype at the selected loci. Thus, from a public health viewpoint, population-based dietary recommendations are of benefit to everyone regardless of genetic variation, at least on the basis of the loci studied here. However, there may be other regions that do interact with diet that we overlooked in our focus on top-ranked GWAS loci, since, in order to reach the extremely low *P* values required for statistical significance, such loci necessarily show little heterogeneity in phenotypic effects (30). Future work focused on genome-wide interaction may uncover regions of the genome that influence biologic response to diet (30). A focus on a reduced number of loci remains important for preserving statistical power. However, other strategies that identify regions that are more likely to interact with environmental factors

Table 5. Meta-Analyzed Effect of Interactions Between Healthy Diet Score, Single Nucleotide Polymorphisms, and Fasting Glucose Genetic Risk Score on Fasting Glucose and Fasting Insulin Concentrations, CHARGE Consortium[a]

SNP	No. of Persons	Diet Score × SNP Interaction	
		β (SE)[b]	*P* Value
Fasting glucose-related loci			
rs340874	51,063	0.0010 (0.0008)	0.22
rs560887	51,117	0.0001 (0.0008)	0.91
rs780094	50,810	0.0007 (0.0008)	0.35
rs11708067	51,403	−0.0003 (0.0009)	0.77
rs11920090	50,828	0.0000 (0.0011)	0.99
rs4607517	51,172	0.0003 (0.0010)	0.76
rs2191349	51,377	−0.0009 (0.0008)	0.23
rs11558471	51,149	0.0002 (0.0009)	0.80
rs7034200	49,146	−0.0002 (0.0008)	0.83
rs10885122	51,126	0.0001 (0.0012)	0.93
rs4506565	51,145	0.0003 (0.0009)	0.77
rs10830963	48,872	−0.0003 (0.0009)	0.78
rs7944584	50,644	0.0006 (0.0009)	0.48
rs11605924	50,720	−0.0005 (0.0008)	0.56
rs174550	51,163	−0.0007 (0.0008)	0.40
rs11071657	51,273	0.0006 (0.0008)	0.44
FG-GRS	51,120	0.0001 (0.0002)	0.67
Fasting insulin-related loci			
rs780094	35,991	−0.0011 (0.0009)	0.25
rs35767	35,739	−0.0011 (0.0012)	0.38

Abbreviations: CHARGE, Cohorts for Heart and Aging Research in Genomic Epidemiology; FG-GRS, fasting glucose genetic risk score; SE, standard error; SNP, single nucleotide polymorphism.

[a] Results were adjusted for age, sex, energy intake, field center (in the Health, Aging and Body Composition Study, the Cardiovascular Health Study, the Atherosclerosis Risk in Communities Study, the Family Heart Study, Invecchiare in Chianti, and the Multi-Ethnic Study of Atherosclerosis), and population substructure (by principal components in the Cardiovascular Heath Study, the Family Heart Study, the Framingham Generation 5 and Offspring Studies, the Multi-Ethnic Study of Atherosclerosis, and the Young Finns Study).

[b] Beta coefficient and standard error for the estimated difference in fasting glucose (mmol/L) or fasting insulin (ln-pmol/L) concentration per 1-unit increase in the effect allele (the fasting glucose- or fasting insulin-raising allele), assuming an additive genetic model, interacting with a 1-point increase in the diet score.

show promise in helping to uncover gene-environment interactions while keeping penalties for multiple testing at a minimum (31).

Some caveats are warranted when interpreting the present data. First, we assumed that the dietary factors most relevant for glucose and insulin homeostasis were included in our diet score and that these factors were measured well. For the majority of our cohorts, we had evidence of successful estimation of dietary intake from conventional validation or reliability studies (described elsewhere (19)). Moreover, our observed associations between the healthy diet score and FG and FI levels are biologically plausible and consistent with other major studies on this topic, thus serving, to some degree, as evidence of construct validity. A second potential limitation of this study is that the global nature of our healthy diet score, which takes into account multiple food choices, may have overwhelmed the biologic influences of individual foods or food components. For example, we previously studied interactions between these same genetic loci and intake of whole-grain foods and observed evidence of interactions between whole grain intake and variation at the *GCKR* locus (19). In contrast, in the present study, we observed no evidence of interaction between the diet score and *GCKR* or any of the other studied loci, suggesting that the signal may be specific to whole grains (or a constituent) and that inclusion of other aspects of a healthy diet diluted this signal. Third, the magnitude of the association between the diet score and FG and FI was modest. Although this is consistent with most reports of associations between dietary factors (which are measured with known random error) and disease-related outcomes, results should be interpreted with consideration of their clinical relevance. Fourth, the selected loci explain only a small fraction of the variation in FG and FI levels (3). Fifth, while our study was uniquely large, it did not have sufficient power to detect very small interaction effects; however, such small interaction effects may be of limited clinical relevance. Lastly, observational studies are prone to residual confounding and causal inference is difficult, particularly in cross-sectional studies such as ours (32). Moreover, such data cannot inform us about the impact of *changing* dietary quality in the short term, a question that is of key importance in designing preventive interventions and that requires intervention studies to be adequately addressed.

Our meta-analysis of data from 15 cohort studies had several strengths. We were able to achieve a large sample size that far exceeded almost all previous studies of gene-diet interactions, while also using a standardized analytic plan and uniformly defined dietary exposures. Furthermore, we were able to take advantage of existing observational data which captured habitual dietary intake, perhaps most significant in the context of gene-environment interactions (32). Additionally, such studies possess information on important confounders, effect modifiers, or mediators of exposure-outcome relations which can be used in analyses. Much of the dietary data used in the current study came from long-standing, well-designed studies with appropriate assessments of data quality and a history of published nutritional epidemiologic research.

Determining whether genetic loci that have been reliably associated with complex disease traits modify associations attributed to protective lifestyle behaviors is important because the presence of such interactions might guide further research and targeted disease prevention. Determining that interactions do not exist between these loci and lifestyle behaviors is important, not least because direct-to-consumer personal genome profiling is now widely available, but data

concerning the utility of the information provided by these products and companies are not. Thus, studies such as ours may help dispel misunderstanding about the way common genetic variants affect disease risk and whether knowledge of one's genomic profile should motivate specific changes in lifestyle, as suggested by some personal genome product manufacturers (4). Based on the evidence reported here, we conclude that the importance of adhering to a healthy diet per se in maintaining glucose and insulin homeostasis is not influenced by one's genotypes at the loci we have studied. Although the present study suggests that the published variants for glucose and insulin traits do not interact with dietary patterns, it remains possible that future studies will discover novel loci that do interact with dietary factors. Future studies focusing on regions of the genome other than those that emerge as the most statistically significant main-effect signals from GWAS may be more fruitful in identifying diet-gene interactions.

ACKNOWLEDGMENTS

Author affiliations: Division of Epidemiology, Human Genetics and Environmental Sciences, School of Public Health, University of Texas Health Science Center at Houston, Houston, Texas (Jennifer A. Nettleton); Department of Medicine, Faculty of Medicine and Health Sciences, University of Sherbrooke, Sherbrooke, Quebec, Canada (Marie-France Hivert); Cardiovascular Health Research Unit, Department of Medicine, School of Medicine, University of Washington, Seattle, Washington (Rozenn N. Lemaitre, David S. Siscovick); Jean Mayer USDA Human Nutrition Research Center on Aging at Tufts University, Boston, Massachusetts (Nicola M. McKeown, Adela Hruby, Paul F. Jacques); Division of Cardiovascular Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts (Dariush Mozaffarian); Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts (Dariush Mozaffarian); Clinical Research Branch, National Institute on Aging, Baltimore, Maryland (Toshiko Tanaka, Luigi Ferrucci); Department of Genetics, Division of Statistical Genomics, School of Medicine, Washington University, St. Louis, Missouri (Mary K. Wojczynski, Ingrid B. Borecki); Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts (Luc Djoussé); Department of Biostatistics, School of Public Health, Boston University, Boston, Massachusetts (Julius S. Ngwa, Josée Dupuis, L. Adrienne Cupples); Division of Biostatistics, School of Public Health, University of Texas Health Science Center at Houston, Houston, Texas (Jack L. Follis); Department of Mathematics, School of Arts and Sciences, University of St. Thomas, Houston, Texas (Jack L. Follis); Department of Nutrition and Dietetics, Harokopio University, Athens, Greece (Maria Dimitriou, Ioanna Ntalla, Constantina Papoutsakis, George Dedoussis); Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden (Andrea Ganna, Erik Ingelsson); Sticht Center on Aging, Department of Internal Medicine, Section on Gerontology and Geriatric Medicine, Wake Forest School of Medicine, Winston-Salem, North Carolina (Denise K. Houston, Stephen B. Kritchevsky); Welcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, United Kingdom (Stavroula Kanoni, Kathleen Stirrups, Panos Deloukas); Division of Nutrition, Department of Food and Environmental Sciences, Faculty of Agriculture and Forestry, University of Helsinki, Helsinki, Finland (Vera Mikkilä); Center for Public Health Genomics and Division of Biostatistics and Epidemiology, University of Virginia, Charlottesville, Virginia (Ani Manichaikul); Department of Nutrition, Harvard School of Public Health, Boston, Massachusetts (Frida Renström, Paul W. Franks, Lawrence de Koning, Frank B. Hu); Department of Clinical Sciences, Faculty of Medicine, Lund University, Malmö, Sweden (Frida Renstrom, Ulrika Ericson, Emily Sonestedt, Marju Orho-Melander, Paul W. Franks); Department of Public Health and Clinical Medicine, Umeå University Hospital, Umeå, Sweden (Frida Renström, Ingegerd Johansson, Olov Rolandsson, Paul W. Franks); Department of Epidemiology, Erasmus MC, Rotterdam, the Netherlands (Frank J. A. vanRooij, Albert Hofman, Jacqueline C. M. Witteman, M. Carola Zillikens, André G. Uitterlinden); Netherlands Genomics Initiative-sponsored Netherlands Consortium for Healthy Aging, Leiden, the Netherlands (Frank J. A. vanRooij, Albert Hofman, Jessica C. Kiefte-de Jong, M. Carola Zillikens, André G. Uitterlinden); Geriatric Unit, Azienda Sanitaria di Firenze, Florence, Italy (Stefania Bandinelli); Oxford Centre for Diabetes, Endocrinology and Metabolism, Radcliffe Department of Medicine, University of Oxford, Oxford, United Kingdom (Neelam Hassanali, Christopher J. Groves, Mark I. McCarthy, Inga Prokopenko); Department of the Generation R Study Group, Erasmus MC, Rotterdam, the Netherlands (Jessica C. Kiefte-de Jong); Department of Biostatistical Sciences, Division of Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, North Carolina (Kurt K. Lohman); Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland (Olli Raitakari); Department of Clinical Physiology, Turku University Hospital, Turku, Finland (Olli Raitakari); Division of Clinical Nutrition and Metabolism, Department of Public Health and Caring Sciences, Uppsala University, Uppsala, Sweden (Per Sjogren, Erika Ax); Department of Odontology, Faculty of Medicine, Umeå University, Umeå, Sweden (Ingegerd Johansson); Department of Epidemiology and Prevention, Division of Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, North Carolina (Yongmei Liu); Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom (Mark I. McCarthy); Oxford NIHR Biomedical Research Centre, Churchill Hospital, Oxford, United Kingdom (Mark I. McCarthy); Department of Epidemiology, School of Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina (Kari North); Carolina Center for Genome Sciences, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina (Kari North); Department of Medicine, University of Turku and Turku University Hospital, Turku, Finland (Jorma Viikari); Department of Internal Medicine, Erasmus

MC, Rotterdam, the Netherlands (M. Carola Zillikens, André G. Uitterlinden); Cardiology Department, Onassis Cardiac Surgery Center, Athens, Greece (Genovefa Kolovou); Division of General Medicine and Primary Care, Beth Israel Deaconess Medical Center, Boston, Massachusetts (Kenneth Mukamal); Unit of Family Medicine, Department of Clinical Chemistry, University of Tampere and Tampere University Hospital, Tampere, Finland (Ilkka Seppälä, Terho Lehtimäki); Department of Biostatistics, School of Public Health, Boston University, Boston, Massachusetts; Department of Clinical Physiology, University of Tampere and Tampere University Hospital, Tampere, Finland (Mika Kähönen); Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, Minnesota (David R. Jacobs, Jr., James S. Pankow); and General Medicine Division, Clinical Epidemiology Unit, and Diabetes Research Unit, Massachusetts General Hospital, Boston, Massachusetts (James B. Meigs).

Because of the large number of cohort studies involved in this analysis, sources of support for each study are given in Web Table 6 (http://aje.oxfordjournals.org/). Cohort-specific acknowledgements can also be found in Web Table 6.

Conflict of interest: none declared.

REFERENCES

1. Frayling TM. Genome-wide association studies provide new insights into type 2 diabetes aetiology. *Nat Rev Genet*. 2007; 8(9):657–662.
2. Grant RW, Moore AF, Florez JC. Genetic architecture of type 2 diabetes: recent progress and clinical implications. *Diabetes Care*. 2009;32(6):1107–1114.
3. Dupuis J, Langenberg C, Prokopenko I, et al. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat Genet*. 2010;42(2): 105–116.
4. Pearson H. Genetic testing for everyone. *Nature*. 2008; 453(7195):570–571.
5. Hunter DJ, Khoury MJ, Drazen JM. Letting the genome out of the bottle—will we get our wish? *N Engl J Med*. 2008; 358(2):105–107.
6. Positively disruptive [editorial]. *Nat Genet*. 2008;40(2):119.
7. Burke W, Psaty BM. Personalized medicine in the era of genomics. *JAMA*. 2007;298(14):1682–1684.
8. Venkat Narayan KM, Weber MB. Clinical risk factors, DNA variants, and the development of type 2 diabetes [letter]. *N Engl J Med*. 2009;360(13):1360.
9. Gulcher J, Stefansson K. Clinical risk factors, DNA variants, and the development of type 2 diabetes [letter]. *N Engl J Med*. 2009;360(13):1360–1361.
10. Lyssenko V, Nilsson P, Groop L. Clinical risk factors, DNA variants, and the development of type 2 diabetes [letter]. *N Engl J Med*. 2009;360(13):1361.
11. Nettleton JA, Steffen LM, Ni H, et al. Dietary patterns and risk of incident type 2 diabetes in the Multi-Ethnic Study of Atherosclerosis (MESA). *Diabetes Care*. 2008;31(9): 1777–1782.
12. Lockheart MS, Steffen LM, Rebnord HM, et al. Dietary patterns, food groups and myocardial infarction: a case-control study. *Br J Nutr*. 2007;98(2):380–387.
13. Nettleton JA, Schulze MB, Jiang R, et al. A priori-defined dietary patterns and markers of cardiovascular disease risk in the Multi-Ethnic Study of Atherosclerosis (MESA). *Am J Clin Nutr*. 2008;88(1):185–194.
14. Hu FB. Dietary pattern analysis: a new direction in nutritional epidemiology. *Curr Opin Lipidol*. 2002;13(1):3–9.
15. Jacobs DR Jr, Gross MD, Tapsell LC. Food synergy: an operational concept for understanding nutrition. *Am J Clin Nutr*. 2009;89(5 suppl):1543S–1548S.
16. Jacobs DR Jr, Steffen LM. Nutrients, foods, dietary patterns as exposures in research: a framework for food synergy. *Am J Clin Nutr*. 2003;78(3 suppl):508S–513S.
17. Mozaffarian D, Appel LJ, Van Horn L. Components of a cardioprotective diet: new insights. *Circulation*. 2011; 123(24):2870–2891.
18. Esposito K, Kastorini CM, Panagiotakos DB, et al. Prevention of type 2 diabetes by dietary patterns: a systematic review of prospective studies and meta-analysis. *Metab Syndr Relat Disord*. 2010;8(6):471–476.
19. Nettleton JA, McKeown NM, Kanoni S, et al. Interactions of dietary whole-grain intake with fasting glucose- and insulin-related genetic loci in individuals of European descent: a meta-analysis of 14 cohort studies. *Diabetes Care*. 2010; 33(12):2684–2691.
20. Choi HK, Willett WC, Stampfer MJ, et al. Dairy consumption and risk of type 2 diabetes mellitus in men: a prospective study. *Arch Intern Med*. 2005;165(9):997–1003.
21. Liu S, Choi HK, Ford E, et al. A prospective study of dairy intake and the risk of type 2 diabetes in women. *Diabetes Care*. 2006;29(7):1579–1584.
22. Pereira MA, Jacobs DR Jr, Van Horn L, et al. Dairy consumption, obesity, and the insulin resistance syndrome in young adults: the CARDIA Study. *JAMA*. 2002;287(16): 2081–2089.
23. Mozaffarian D, Cao H, King IB, et al. Trans-palmitoleic acid, metabolic risk factors, and new-onset diabetes in U.S. adults: a cohort study. *Ann Intern Med*. 2010;153(12):790–799.
24. Halton TL, Willett WC, Liu S, et al. Potato and French fry consumption and risk of type 2 diabetes in women. *Am J Clin Nutr*. 2006;83(2):284–290.
25. Liese AD, Schulz M, Moore CG, et al. Dietary patterns, insulin sensitivity and adiposity in the multi-ethnic Insulin Resistance Atherosclerosis Study population. *Br J Nutr*. 2004;92(6):973–984.
26. Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. *BMJ*. 2003;327(7414):557–560.
27. Pan A, Sun Q, Bernstein AM, et al. Red meat consumption and risk of type 2 diabetes: 3 cohorts of US adults and an updated meta-analysis. *Am J Clin Nutr*. 2011;94(4): 1088–1096.
28. Villegas R, Xiang YB, Elasy T, et al. Fish, shellfish, and long-chain n-3 fatty acid consumption and risk of incident type 2 diabetes in middle-aged Chinese men and women. *Am J Clin Nutr*. 2011;94(2):543–551.
29. Nanri A, Mizoue T, Noda M, et al. Fish intake and type 2 diabetes in Japanese men and women: the Japan Public Health Center-based Prospective Study. *Am J Clin Nutr*. 2011;94(3):884–891.
30. Franks PW. Gene × environment interactions in type 2 diabetes. *Curr Diab Rep*. 2011;11(6):552–561.
31. Paré G, Cook NR, Ridker PM, et al. On the use of variance per genotype as a tool to identify quantitative trait interaction effects: a report from the Women's Genome Health Study. *PLoS Genet*. 2010;6(6):e1000981. (doi:10.1371/journal.pgen.1000981).

32. Franks PW, Nettleton JA. Gene × lifestyle interactions and complex disease traits—inferring cause and effect from observational data, sine qua non. *Am J Epidemiol*. 2010; 172(9):992–999.
33. The ARIC investigators. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. *Am J Epidemiol*. 1989;129(4):687–702.
34. Fried LP, Borhani NO, Enright P, et al. The Cardiovascular Health Study: design and rationale. *Ann Epidemiol*. 1991; 1(3):263–276.
35. Higgins M, Province M, Heiss G, et al. NHLBI Family Heart Study: objectives and design. *Am J Epidemiol*. 1996;143(12):1219–1228.
36. Feinleib M, Kannel WB, Garrison RJ, et al. The Framingham Offspring Study: design and preliminary data. *Prev Med*. 1975;4(4):518–525.
37. Splansky GL, Corey D, Yang Q, et al. The Third Generation Cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination. *Am J Epidemiol*. 2007;165(11): 1328–1335.
38. Papoutsakis C, Vidra NV, Hatzopoulou I, et al. The Gene-Diet Attica Investigation on Childhood Obesity (GENDAI): overview of the study design. *Clin Chem Lab Med*. 2007; 45(3):309–315.
39. Kanoni S, Dedoussis GV. Design and descriptive characteristics of the GHRAS: the Greek Health Randomized Aging Study. *Med Sci Monit*. 2008;14(4):CR204–CR212.
40. Renström F, Shungin D, Johansson I, et al. Genetic predisposition to long-term nondiabetic deteriorations in glucose homeostasis: ten-year follow-up of the GLACIER Study. *Diabetes*. 2011;160(1):345–354.
41. Houston DK, Nicklas BJ, Ding J, et al. Dietary protein intake is associated with lean mass change in older, community-dwelling adults: the Health, Aging, and Body Composition (Health ABC) Study. *Am J Clin Nutr*. 2008;87(1):150–155.
42. Ferrucci L, Bandinelli S, Benvenuti E, et al. Subsystems contributing to the decline in ability to walk: bridging the gap between epidemiology and geriatric practice in the InCHIANTI study. *J Am Geriatr Soc*. 2000;48(12): 1618–1625.
43. Berglund G, Elmstähl S, Janzon L, et al. The Malmö Diet and Cancer Study: design and feasibility. *J Intern Med*. 1993;233(1):45–51.
44. Bild DE. Multi-Ethnic Study of Atherosclerosis: objectives and design. *Am J Epidemiol*. 2002;156(9):871–881.
45. Hofman A, van Duijn CM, Franco OH, et al. The Rotterdam Study: 2012 objectives and design update. *Eur J Epidemiol*. 2011;26(8):657–686.
46. Theodoraki EV, Nikopensius T, Suhorutšenko J, et al. Fibrinogen beta variants confer protection against coronary artery disease in a Greek case-control study. *BMC Med Genet*. 2010;11:p28. (doi:10.1186/1471-2350-11-28).
47. Hedstrand H. A study of middle-aged men with particular reference to risk factors for cardiovascular disease. *Ups J Med Sci Suppl*. 1975;19:1–61.
48. Raitakari OT, Juonala M, Rönnemaa T, et al. Cohort profile: the Cardiovascular Risk in Young Finns Study. *Int J Epidemiol*. 2008;37(6):1220–1226.

Chapter 4

CONCLUSIONS

4 CONCLUSIONS

The results of the current PhD thesis are supportive of the high prevalence of risk factors in the case group with established CAD. Furthermore, exclusive olive oil consumption in cooking seems to have a beneficial impact on CAD risk. The participation of the THISEAS study in the CARDIOoGRAMplusC4D Consortium contributed to the identification of novel loci for CAD, with data from a Greek sample. The findings revealed novel loci robustly associated with CAD. In addition, the construction of a GRS based on a panel of 53 variants was associated with an increased CAD risk. THISEAS remain the first study in Greece that used the Metabochip and Omni-express arrays for genetic analysis. With regard to dietary components and among eight generated dietary factors, the Western type diet showed a modest effect on CAD risk. Apart from the CARDIOoGRAMplusC4D Consortium, the THISEAS study also participated in CHARGE Consortium through data sharing. A meta-analysis from CHARGE, demonstrated an inverse association between an a priori diet score with FG and FI levels regardless of genotypic information reflected by a GRS or genotype of previously replicated loci. Results also demonstrated that a GPS was associated with decreasing glucose levels in the presence of a GRS, based on a panel of 20 glucose-raising variants. Last but not least, the association of the GRS with glucose levels was weakened after interaction with the GPS.

During the last decade, the GWAs approach was useful in identifying novel loci for CAD and other complex traits. Before the GWAs experimental design, the vast majority of loci published and robustly associated with complex traits were unknown. This study design provided researchers with an unbiased option to detect association with common variants within the population, even in the absence of knowledge or assumption regarding the location or function of the variant. GWAs discoveries about genes have extrapolated human genetic research. They have also provided new biological leads that promise the generation of additional genes and the elucidation of genetic mechanisms and biological pathways. The CARDIOoGRAMplusC4D consortium remains the largest GWAs that investigated genetic variation in CAD and the THISEAS study contributed to the generation of important scientific knowledge with Greek genotypic data.

Limitations of GWA analyses include the risk of false positive results, misclassification of cases and controls, biases from incomplete phenotyping of cases and controls and

confounders due to population stratification. In addition, most variants identified have no functional relevance to disease and the clinical utility for prognosis are yet to be determined. Furthermore, given the results of CARDIOoGRAMplusC4D analysis, along with others collaborative efforts, it is obvious that GWAs have been able to account for a small proportion of genetic variation in CAD or other complex traits (typically less than 10%). Professor Sir Alec Jeffreys, ESHG Award Lecturer 2010, pointed out in an interview that "the fact remains that the bulk of heritability in these conditions cannot be ascribed to loci that have emerged from GWAs, which clearly isn't going to be the answer to everything". The outcomes of these studies have tested the "common disease-common variant" hypothesis. Therefore, the debate regarding the allele frequency of variants contributing to common disease, including CAD, remains open with the other side arguing that rare variants may have a larger effect on common diseases. However, the GWA approach cannot detect loci with rare alleles because of lack of statistical significance and it cannot explain all of the familial clustering.

Sequencing provides a solution to this obstacle. Advances in sequencing technology since 1990s enabled the massive increase of data output, while sequencing cost has fallen dramatically over the last decades. This progress enables the scientists to analyze an even larger set of samples in an attempt to discover all related disease-causing genes and to be able to establish personalized genomic medicine in clinical settings. Exome sequencing is the most widely used method, which is a cost-effective technique that enables the sequencing of the subset of DNA that encodes proteins (exons). Although the exome constitutes less than 2% of the human genome, it encloses disease-causing variants. Therefore, it is an approach of great importance giving the capacity to identify variants responsible for common diseases or diseases that follow the Mendelian inheritance.

Also, the majority of GWA studies have been limited to European Caucasian populations. Allele frequencies may differ between different ethnic groups, thus it is important to extent association studies in non-European ancestry populations in an attempt to detect even more rare variants . Isolated populations could also be an option of great value for the discovery of novel rare variants.

Currently, a "weighted" or "unweighted" construction of a GRS, based on a panel of CAD associated loci, is a promising approach to examine the cumulative predictive ability of genetic diversities on CAD outcome. Up to date, there is a limited number of studies that have constructed and used GRSs, each based on a panel of CAD associated SNPs, published

by consortium meta-analyses. The findings of the studies demonstrated a modest association between the GRS under examination and CAD outcome.

However, evidence is not yet supportive to implement GRSs in clinical practice in order to adequately improve CAD risk and/or predict future CAD events in secondary prevention. Two future perspectives are the identification of additional CAD associated variants and more predictive risk prediction models.

Regarding an environmental parameter, that is diet, research using DPs highlights the importance of food diversity to health outcomes. Although dietary assessment is approached in various ways, many studies have consistent results regarding DPs and CAD. These results are important and may contribute to the implementation of programs and services supporting for overall healthy patterns against disease prevention and development instead of the avoidance of certain foods or nutrients. Policy health promotion can encourage the improvement on eating behaviors at an individual and population level.

Despite the flaws, GWAs have been a valuable tool in genetic epidemiology and have been successful in identifying common variants with small or modest effect on CAD. Their results represent a wealth of data and provide new insights to the genetic architecture of the disease. Also, GWAs have demonstrated that common traits cannot be explained by a number of common variants with moderate effect. Thus, research in the near future should consider to: determine allele association to non-European samples; investigate gene-gene interactions and gene-environment interactions; expand GWAs to family studies and monozygotic twins; to improve phenotyping in order to reduce heterogeneity; to conduct large prospective studies with micro-array genotyping. Successful yielding DNA's secrets and gene-environment interactions will enable the prosperous design of risk stratification algorithms to identify patients at high risk of developing certain diseases, based on their genetic profile, clinical indices and behavioral habits.

The combination of epidemiology with rapid advances in technology will unfold future discoveries, which will bring us closer to the ultimate goal of personalized medicine: recommendations, prognosis and treatment that can be personalized at an individual level and tailored to the specific needs of an individual, taking into consideration information from the genetic imprint as interacts with the environment.

Chapter 5

REFERENCES

5 REFERENCES

Alexander JK. Obesity and coronary heart disease. Am J Med Sci 2001;321(4):215-224.

American Diabetes Association. Standards of medical care in diabetes--2013. Diabetes Care 2013;36:S11-S66.

Andresdottir MB, Sigurdsson G, Sigvaldason H, Gudnason V. Fifteen percent of myocardial infarctions and coronary revascularizations explained by family history unrelated to conventional risk factors: the Reykjavik Cohort Study. Eur Heart J 2002; 23:1655–1663

Artinian NT, Fletcher GF, Mozaffarian D, et al. Interventions to promote physical activity and dietary lifestyle changes for cardiovascular risk factor reductionin adults: a scientific statement from the American Heart Association. Circulation. 2010;122:406-441.

Assmann G, Cullen P, Schulte H. Simple scoring scheme for calculating the risk of acute coronary events based on the 10-year follow-up of the prospective cardiovascular Münster (PROCAM) study. Circulation 2002;105:310-315.

Assmann G, Schulte H. The Prospective Cardiovascular Münster (PROCAM) study: prevalence of hyperlipidemia in persons with hypertension and/or diabetes mellitus and the relationship to coronary heart disease. Am Heart J 1988;116:1713-1724.

Astrup A, Dyerberg J, Elwood P, et al. The role of reducing intakes of saturated fat in the prevention of cardiovascular disease: where does the evidence stand in 2010? Am J Clin Nutr 2011;93:684–688.

Bachmann JM, Willis BL, Ayers CR, Khera A, Berry JD. Association between family history and coronary heart disease death across long-term follow-up in men: the Cooper Center Longitudinal Study. Circulation 2012, 26;125:3092-3098

Bailar JC 3rd. Passive smoking, coronary heart disease, and meta-analysis. N Engl J Med 1999;340:958–959

Bild DE , Bluemke DA, Burke GL, et al. Multi-Ethnic Study of Atherosclerosis: objectives and design. Am J Epidemiol 2002;156:871-881.

Bronte-Stewart B. Cigarette smoking and ischemic heart disease. Brit Med J 1961;1:379–384.

Buckland G, Travier N, Barricarte A, et al. Olive oil intake and CHD in the European Prospective Investigation into Cancer and NutritionSpanish cohort. Br J Nutr 2012;108:2075-2082.

Budoff MJ, Nasir K, Katz R, et al. Thoracic aortic calcification and coronary heart disease events: the multi-ethnic study of atherosclerosis (MESA). Atherosclerosis 2011;215:196-202.

Buechley RW, Drake RM, Breslow L. Relationship of amount of cigarette smoking to coronary heart disease mortality rates in men. Circulation 1958;18:1085–1090.

Cambien F, Poirier O, Lecerf L, et al. Deletion polymorphism in the gene for angiotensin-converting enzyme is a potent risk factor for myocardial infarction. Nature 1992;359:641–644.

CARDIoGRAMplusC4D Consortium, Deloukas P, Kanoni S, et al. Large-scale association analysis identifies new risk loci for coronary artery disease. Nat Genet 2013; 45, 25-33.

Carnethon MR. Physical activity and cardiovascular disease: How much is enough? Am J Lifestyle Med 2009;3:44S–49S .

Chaer RA, Billeh R, Massad MG. Genetics and gene manipulation therapy of premature coronary artery disease. Cardiology 2004;101:122–130.

Chobanian AV, Bakris GL, Black HR, et al; Joint National Committee on Prevention, Detection,Evaluation, and Treatment of High Blood Pressure. National Heart, Lung, and Blood Institute;National High Blood Pressure Education Program Coordinating Committee. Seventh report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure. Hypertension 2003;42:1206-1252.

Chrysohoou C, Panagiotakos DB, Pitsavos C, et al. Gender differences on the risk evaluation of acute coronary syndromes: the CARDIO2000 study. Prev Cardiol 2003;6:71-77.

Davies RW, Dandona S, Stewart AF, et al. Improved prediction of cardiovascular disease based on a panel of single nucleotide polymorphisms identified through genome-wide association studies. Circ Cardiovasc Genet 2010;3:468-474.

Davies SW. Clinical presentation and diagnosis of coronary artery disease: stable angina. Br Med Bull 2001;59:17-27.

Delaneau O, Marchini J, Zagury JF. A linear complexity phasing method for thousands of genomes. Nat Methods 2012;9:179-181.

Detrano R, Guerci AD, Carr JJ, et al. Coronary calcium as a predictor of coronary events in four racial or ethnic groups. N Engl J Med 2008;358:1336–1345.

Devaraj S, Jialal I. Failure of vitamin E in clinical trials: is gamma-tocopherol the answer? Nutr Rev 2005;63:290-293.

Dimitriou M, Rallidis LS, Theodoraki EV, Kalafati IP, Kolovou G, Dedoussis GV. Exclusive olive oil consumption has a protective effect on coronary artery disease; overview of the THISEAS study. Public Health Nutr 2015;30:1-7.

Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to smoking: 50 years' observations on male British doctors. BMJ 2004;328:1519.

Doyle JT, Dawber TR, Kannel WB, Kinch SH, Kahn HA. The relationship of cigarette smoking to coronary heart disease; The second report of the combined experience of the Albany, NY and Framingham, Mass, Studies. JAMA 1964 Dec 7;190:886-890.

Ganna A, Magnusson PK, Pedersen NL, et al. Multilocus genetic risk scores for coronary heart disease prediction. Arterioscler Thromb Vasc Biol 2013; 33:2267-2272.

Goldstein DJ. Beneficial health effects of modest weight loss. Int J Obes Relat Metab Disord 1992;16:397-415.

Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults . Executive summary of the Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) . JAMA 2001;285:2486–2497 .

Centers for Disease Control and Prevention (US), National Center for Chronic Disease Prevention and Health Promotion (US), Office on Smoking and Health (US). How Tobacco Smoke Causes Disease: The Biology and Behavioral Basis for Smoking-Attributable Disease: A Report of the Surgeon General. Atlanta (GA): Centers for Disease Control and Prevention (US) 2010. Available from: https://www.ncbi.nlm.nih.gov/books/NBK53017/.

Cholesterol Treatment Trialists' (CTT) Collaborators, Mihaylova B, Emberson J, et al. The effects of lowering LDL cholesterol with statin therapy in people at low risk of vascular disease: meta-analysis of individual data from 27 randomised trials. Lancet 2012;380:581-590.

Global Burden of Metabolic Risk Factors for Chronic Diseases Collaboration (BMI Mediated Effects), Lu Y, Hajifathalian K, et al. Metabolic mediators of the effects of body-mass index, overweight, and obesity on coronary heart disease and stroke: a pooled analysis of 97 prospective cohorts with 1·8 million participants. Lancet 2014 15;383:970-983.

Colditz GA, Stampfer MJ, Willett WC, Rosner B, Speizer FE, Hennekens CH. A prospective study of parental history of myocardial infarction and coronary heart disease in women. Am J Epidemiol 1986 Jan;123:48-58.

Estruch R, Ros E, Martínez-González MA. Mediterranean diet for primary prevention of cardiovascular disease. N Engl J Med 2013;369:676-677.

Fan J, Watanabe T. Inflammatory reactions in the pathogenesis of atherosclerosis. J Atheroscler Thromb 2003;10:63-71.

Finegold JA, Asaria P, Francis DP. Mortality from ischaemic heart disease by country, region, and age: statistics from World Health Organisation and United Nations. Int J Cardiol 2013;168:934-945.

Franklin SS, Larson MG, Khan SA, et al. Does the relation of blood pressure to coronary heart disease risk change with aging? The Framingham HeartStudy. Circulation 2001;103:1245-1249.

Frazier-Wood AC. Dietary Patterns, Genes, and Health: Challenges and Obstacles to be Overcome. Curr Nutr Rep 2015;4:82-87.

Friedewald WT, Levi RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. Clin Chem 1972;18:499–502.

Ford ES, Ajani UA, Croft JB, et al. Explaining the decrease in U .S . deaths from coronary disease, 1980–2000. N Engl J Med 2007;356:2388–2398.

Fox CS, Coady S, Sorlie PD, et al. Increasing cardiovascular disease burden due to diabetes mellitus: the Framingham Heart Study. Circulation 2007;115:1544-1550.

Fung TT, Chiuve SE, McCullough ML, Rexrode KM, Logroscino G, Hu FB. Adherence to a DASH-style diet and risk of coronary heart disease and stroke in women. Arch Intern Med 2008;168:713-720.

Fung TT, Malik V, Rexrode KM, Manson JE, Willett WC, Hu FB. Sweetened beverage consumption and risk of coronary heart disease in women. Am J Clin Nutr 2009;89:1037-1042.

Gordon T, Castelli WP, Hjortland MC, et al. Predicting coronary heart disease in middle-aged and older persons. The Framington study. JAMA 1977;238:497-499.

Grech ED. Pathophysiology and investigation of coronary artery disease. BMJ 2003;326:1027-1030.

Griffin SJ, Borch-Johnsen K, Davies MJ, et al. Effect of early intensive multifactorial therapy on 5-year cardiovascular outcomes in individuals with type 2diabetes detected by screening (ADDITION-Europe): a cluster-randomised trial. Lancet 2011;378:156-167.

Guasch-Ferré M, Hu FB, Martínez-González MA, et al. Olive oil intake and risk of cardiovascular disease and mortality in the PREDIMED Study. BMC Med 2014;12:78.

Guenther PM, Casavale KO, Reedy J, et al. Update of the Healthy Eating Index: HEI-2010. J Acad Nutr Diet 2013;113:569-580.

Hegglin R, Keiser G. [Smoking and coronary disease]. Schweiz Med Wochenschr 1955; 85:53-55.

Helgadottir A, Thorleifsson G, Manolescu A, et al. A common variant on chromosome 9p21 affects the risk of myocardial infarction. Science 2007;316:1491-3.

Hill JA. In vino veritas: alcohol and heart disease. Am J Med Sci. 2005;329:124-35.

Hivert MF, Jablonski KA, Perreault L, et al. Updated genetic score based on 34 confirmed type 2 diabetes Loci is associated with diabetes incidence and regression to normoglycemia in the diabetes prevention program. Diabetes. 2011;60:1340-1348.

Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet 2009;5:e1000529.

Hu FB. Diet and lifestyle influences on risk of coronary artery disease. Curr Atheroscler rep 2009;11:257-263.

Icks A, Claessen H, Kirchberger I, et al. Mortality after first myocardial infarction in diabetic and non-diabetic people between 1985 and 2009. The MONICA/KORA registry. Eur J Epidemiol 2014;29:899-909.

Iqbal R, Anand S, Ounpuu S, et al. Dietary patterns and the risk of acute myocardial infarction in 52 countries: results of the INTERHEART study. Circulation 2008;118:1929–1937.

Jeppesen J, Schaaf P, Jones C, Zhou MY, Chen YD,Reaven GM. Effects of low-fat, high-carbohydrate diets on risk factors for ischemic heart disease in postmenopausal women. AmJClinNutr 1997;65:1027-1033.

Jhamnani S, Patel D, Heimlich L, King F., Walitt B, Lindsay J. Meta-analysis of the effects of lifestyle modifications on coronary and carotid atherosclerotic burden. Am J Cardiol 2015;115:268-275.

Kalandidi A, Tzonou A, Toupadaki N, et al. A case-control study of coronary heart disease in Athens, Greece. Int J Epidemiol 1992;21:1074-1080.

Kalra S, Duggal S, Valdez G et al. Review of acute coronary syndrome diagnosis and management. Postgrad Med 2008;120: 18-27

Kannel WB, McGee DL. Diabetes and cardiovascular disease: the Framingham Study. JAMA 1979;241:2035-2038.

Kannel WB, McGee DL. Diabetes and glucose tolerance as risk factors for cardiovascular disease: the Framingham Study. Diabetes Care 1979;2:120-126.

Kannel WB. Some lessons in cardiovascular epidemiology from Framingham. Am J Cardiol 1976;37:269-282.

Kannel WB, McGee D, Gordon T. A general cardiovascular risk profile: the Framingham Study. Am J Cardiol 1976;38:46-51.

Kannel WB, Dawber TR, Kagan A, Revotskie N, Stokes J 3rd. Factors of risk in the development of coronary heart disease--six year follow-up experience. The FraminghamStudy. Ann Intern Med 1961;55:33-50.

Keys A, Taylor HL, Blackburn H, Brozek J, Anderson JT, Simonson E. Coronary heart disease among Minnesota business and professional men followed 15 years. Circulation 1963;28:381-395.

Keys A, Menotti A, Aravanis C, et al. The seven countries study: 2,289 deaths in 15 years. Prev Med 1984;13:141–154

Klein S, Burke LE, Bray GA, et al; on behalf of the American Heart Association Council on Nutrition, Physical Activity, and Metabolism . Clinical implications of obesity with specific focus on cardiovascular disease: a statement for professionals from the American Heart Association Council on Nutrition, Physical Activity, and Metabolism. Circulation 2004;110:2952–2967.

Kontogianni MD, Panagiotakos DB, Chrysohoou C, Pitsavos C, Zampelas A, Stefanadis C. The impact of olive oil consumption pattern on the risk of acute coronary syndromes: The CARDIO2000 case-control study. Clin Cardiol 2007;30:125–129.

Kromhout D, Keys A, Aravanis C, et al. Food consumption patterns in the 1960s in seven countries. Am J Clin Nutr 1989;49:889-894.

Lewington S, Clarke R, Qizilbash N, Peto R, Collins R; Prospective Studies Collaboration. Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. Lancet 2002;360:1903-1913.

Lim SS, Vos T, Flaxman AD, Danaei G, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 201215;380:2224-2226.

Liu S, Lee IM, Ajani U, et al. Intake of vegetables rich in carotenoids and risk of coronary heart disease in men: The Physicians' Health Study. Int J Epidemiol 2001;30:130–135.

Lloyd-Jones DM, Hong Y, Labarthe D, et al. Defining and setting national goals for cardiovascular health Promotion and disease reduction The American Heart Association's strategic Impact Goal through 2020 and beyond, Circulation 2010;121:586-613.

Lloyd-Jones DM, Nam BH, D'Agostino RB Sr, et al. Parental cardiovascular disease as a risk factor for cardiovascular disease in middle-aged adults: a prospective study of parents and offspring. JAMA 2004;291:2204-2211.

Lozano R, Naghani M, Foreman K, Foreman, Ket al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010. A systematic analysis for the Global burden of disease study 2010. Lancet 2012; 380:2095-2128.

Luo C, Zhang Y, Ding Y, et al. Nut consumption and risk of type 2 diabetes, cardiovascular disease, and all-cause mortality: a systematic review and meta-analysis. Am J Clin Nutr 2014;100:256–269.

Maas AH, Appelman YE. Gender differences in coronary heart disease. Neth Heart J 2010;18:598-602.

Maraki M, Kollia A, Gioxari A, Sidossis LS, Kavouras SA. Development, reliability and validity of the Athens Physical Activity Questionnaire. Clin Nutr 20010;S5:80.

Marenberg ME, Risch N, Berkman LF, Floderus B, de Faire U. Genetic susceptibility to death from coronary heart disease in a study of twins. N Engl J Med 1994;330:1041-1046.

Martinez-González MA, Bes-Rastrollo M. Dietary patterns, Mediterranean diet, and cardiovascular disease. Curr Opin Lipidol 2014;25:20-26.

Martínez-González MÁ, Corella D, Salas-Salvadó J, et al; PREDIMED Study Investigators. Cohort profile: design and methods of the PREDIMED study. Int J Epidemiol 2012;41:377-385.

Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. Diabetologia. 1985;28:412-419.

McGeachie M, Ramoni RL, Mychaleckyj JC, et al. Integrative predictive model of coronary artery calcification in atherosclerosis. Circulation 2009;120:2448-2454.

McPherson R, Pertsemlidis A, Kavaslar N, Stewart A, Roberts R, Cox DR. A common allele on chromosome 9 associated with coronary heart disease. Science 2007;316:1488-1491.

Medical Expenditure Panel Survey: household component summary tables. Table 1: Total Health Services: Median and Mean Expenses per Person With Expense and Distribution of Expenses by Source of Payment: United States, 2011. Agency for Healthcare Research and Quality Website. Available from: http://meps.ahrq.gov/mepsweb/data_stats/tables_compendia_hh_interactive.jsp?_SERVICE=MEPSSocket0&_PROGRAM=MEPSPGM.TC.SAS&File=HCFY2011&Table=HCFY2011%5FPLEXP%5F%40&VAR1=AGE&VAR2=SEX&VAR3=RACETH5C&VAR4=INSURCOV&VAR5=POVCAT11&VAR6=MSA&VAR7=REGION&VAR8=HEALTH&VAR01=4+17+44+64&VAR02=1&VAR03=1&VAR04=1&VAR05=1&VAR06=1&VAR07=1&VAR08=1&_Debug

Mendis S, Puska P, Norrving B editors. Global Atlas on Cardiovascular Disease Prevention and Control. World Health Organization, Geneva, WHO, 2011. Available from: http://www.who.int/cardiovascular_diseases/publications/atlas_cvd/en/.

Menotti A, Keys A, Aravanis C, et al. Seven Countries Study. First 20-year mortality data in 12 cohorts of six countries. Ann Med 1989;21:175-179.

Mente A, de Koning L, Shannon HS, Anand SS. A systematic review of the evidence supporting a causal link between dietary factors and coronary heart disease. Arch Intern Med 2009;169:659-669

Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 1998;16:1215.

Miura K, Daviglus ML, Dyer AR, et al. Relationship of blood pressure to 25-year mortality due to coronary heart disease, cardiovascular diseases, and all causes in young adult men: the Chicago Heart Association Detection Project in Industry. Arch Intern Med 2001;161:1501-1508.

Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, et al. Heart Disease and Stroke Statistics-2015 Update: A Report From the American Heart Association. Circulation 2016 ;133:e38-60.

Mozaffarian D, Katan MB, Ascherio A, Stampfer MJ, Willett WC. Trans fatty acids and cardiovascular disease. N Engl J Med 2006;354:1601–1613.

Moulopoulos SD, Adamopoulos PN, Diamantopoulos EI, Nanas SN, Anthopoulos LN, Iliadi-Alexandrou M. Coronary heart disease risk factors in a random sample of Athenian adults. The Athens Study. Am J Epidemiol. 1987;126:882-892.

Murabito JM, Pencina MJ, Nam BH, et al. Sibling cardiovascular disease as a risk factor for cardiovascular disease in middle-aged adults. JAMA 2005;294:3117-3123.

Myers RH, Kiely DK, Cupples LA, Kannel WB. Parental history is an independent risk factor for coronary artery disease: the Framingham Study. Am Heart J 1990;120:963-969.

Myocardial Infarction Genetics Consortium, Kathiresan S, Voight BF, et al. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. Nat Genet 2009;41:334-341.

Napoli C, D'Armiento FP, Mancini FP, et al.Fatty streak formation occurs in human fetal aortas and is greatly enhanced by maternal hypercholesterolemia: intimal accumulation of low density lipoprotein and its oxidation precede monocyte recruitment into early atherosclerotic lesions. J Clin Invest 1997;100:2680-2690.

National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, andTreatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report. Circulation 2002 ;106:3143-421.

Nettleton JA, Schulze MB, Jiang R, Jenny NS, Burke GL, Jacobs DR Jr. A priori-defined dietary patterns and markers of cardiovascular disease risk in the Multi-Ethnic Study of Atherosclerosis (MESA). Am J Clin Nutr 2008;88:185-194.

Newby PK, Tucker KL. Empirically derived eating patterns using factor or cluster analysis: a review. Nutr Rev 2004;62:177-203.

Ng M, Fleming T, Robinson M, et al. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2014;384:766-781.

Nikpay M, Goel A, Won HH, et al. A comprehensive 1000 Genomes-based genome-wide association meta-analysis of coronary artery disease. Nat Genet 2015; 47:1121-1123.

Omenn GS, Goodman GE, Thornquist MD, et al. Effects of a combination of beta carotene and vitamin A on lung cancer and cardiovascular disease. N Engl J Med 1996;334:1150-1155.

Ouyang P, Michos ED, Karas RH. Hormone replacement therapy and the cardiovascular system: Lessons learned and unanswered questions. J Am Coll Cardiol 2006;47:1741–1753.

Ozaki K, Ohnishi Y, Iida A, et al. Functional SNPs in the lymphotoxin-alpha gene that are associated with susceptibility to myocardial infarction. Nat Genet 2002;32:650-654.

Panagiotakos DB, Pitsavos C, Arvaniti F et al. Adherence to the Mediterranean food pattern predicts the prevalence of hypertension, hypercholesterolemia, diabetes and obesity, among healthy adults; the accuracy of the MedDietScore. Prev Med 2007;44:335–340.

Panagiotakos DB, Georgousopoulou EN, Pitsavos C, et al. Ten-year (2002-2012) cardiovascular disease incidence and all-cause mortality, in urban Greek population: the ATTICA Study. Int J Cardiol 2015;180:178-184.

Panagiotakos DB, Pitsavos C, Skoumas Y, Stefanadis C. The association between food patterns and the metabolic syndrome using principal components analysis: The ATTICA Study. J Am Diet Assoc 2007;107:979-987

Panagiotakos DB, Pitsavos C, Chrysohoou C, Skoumas I, Stefanadis C. Prevalence and five-year incidence (2001-2006) of cardiovascular disease risk factors in a Greek sample: the ATTICA study Hellenic J Cardiol 2009;50:388-395.

Panagiotakos DB, Pitsavos C, Chrysohoou C, Skoumas I, Stefanadis C; ATTICA Study. Five-year incidence of cardiovascular disease and its predictors in Greece: the ATTICA study. Vasc Med 2008;13:113-121.

Panagiotakos DB, Pitsavos C, Stefanadis C. Dietary patterns: a Mediterranean diet score and its relation to clinical and biological markers of cardiovascular disease risk. Nutr Metab Cardiovasc Dis 2006;16:559-68.

Panagiotakos DB, Pitsavos C, Chrysohoou C, Stefanadis C, Toutouzas P. Risk stratification of coronary heart disease through established and emerging lifestyle factors in a Mediterranean population: CARDIO2000 epidemiological study. J Cardiovasc Risk 2001;8:329-335.

Park SK, Auchincloss AH, O'Neill MS, et al. Particulate air pollution, metabolic syndrome, and heart rate variability: the multi-ethnic study of atherosclerosis (MESA). Environ Health Perspect 2010;118:1406-1411.

Paynter NP, Chasman DI, Paré G, et al. Association between a literature-based genetic risk score and cardiovascular events in women. JAMA 2010: 303: 631-637.

Peden JF, Farrall M. Thirty-five common variants for coronary artery disease: the fruits of much collaborative labour. Hum Mol Genet 2011;20:198–205.

Peterson LR, Waggoner AD, Schechtman KB, et al. Alterations in left ventricular structure and function in young healthy obese women: assessment by echocardiography and tissue Doppler imaging. J Am Coll Cardiol 2004;43:1399–1404.

Physical Activity Guidelines Advisory Committee. Physical Activity Guidelines Advisory Committee Report, 2008. Washington, DC: U.S. Department of Health and Human Services, 2008. Available from: https://health.gov/paguidelines/report

Piepoli MF, Hoes AW, Agewall S, et al. 2016 European Guidelines on cardiovascular disease prevention in clinical practice: The Sixth Joint Task Force of the European Society of Cardiology and Other Societies onCardiovascular Disease Prevention in Clinical Practice (constituted by representatives of 10 societies and by invited experts): Developed with the special contribution of the EuropeanAssociation for Cardiovascular Prevention & Rehabilitation (EACPR). Eur J Prev Cardiol 2016;23:NP1-NP96

Pitsavos C, Panagiotakos DB, Chrysohoou C, Stefanadis C. Epidemiology of cardiovascular risk factors in Greece: aims, design and baseline characteristics of the ATTICA study. BMC Public Health 2003;3:32.

Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007: 81;559-575.

Rasmussen-Torvik LJ, Li M, Kao WH, et al. Association of a fasting glucose genetic risk score with subclinical atherosclerosis: The Atherosclerosis Risk in Communities (ARIC) study. Diabetes 2011;60:331-335.

Ronksley PE, Brien SE, Turner BJ, Mukamal KJ, Ghali WA. Association of alcohol consumption with selected cardiovascular disease outcomes: a systematic review and meta-analysis. BMJ 2011;342:d671.

Rosendorff C, Black HR, Cannon CP, et al. Treatment of hypertension in the prevention and management of ischemic heart disease: a scientific statement from the American Heart Association Council for High Blood Pressure Research and the Councils on Clinical Cardiology and Epidemiology and Prevention. Circulation 2007;115:2761-2788.

Pepine CJ, Nichols WW. The Pathophysiology of Chronic Ischemic Heart Disease. Clinical Cardiology 2007;30:4–9

Porta A, Rienzo MD, Wessel N, Kurths J. Addressing the complexity of cardiovascular regulation A. Philos Trans A Math Phys Eng Sci. 2009;367:1215-1218.

Preuss M, König IR, Thompson JR, et al. Design of the Coronary ARtery DIsease Genome-Wide Replication And Meta-Analysis (CARDIoGRAM) Study: A Genome-wide association meta-analysis involving more than 22000 cases and 60000 controls. Circ Cardiovasc Genet 2010;3:475-483.

Rosengren A, Hawken S, Ounpuu S, et al. Association of psychosocial risk factors with risk of acute myocardial infarction in 11119 cases and 13648 controls from 52 countries (the INTERHEART study): case-control study. Lancet 2004;364:953-962.

Samani NJ, Erdmann J, Hall AS, et al. Genome wide association analysis of coronary artery disease. N Engl J Med 2007;357:443-453.

Sidney S, Rosamond WD, Howard VJ, Luepker, RV, National Forum for Heart Disease and Stroke Prevention. The "heart disease and stroke statistics--2013 update" and the need for a national cardiovascular surveillance system. Circulation 2013; 127:21-23.

Sofi F, Macchi C, Abbate R, Gensini GF, Casini A. Mediterranean diet and health status: an updated meta-analysis and a proposal for a literature-based adherence score. Public Health Nutr 2014;17:2769-2782.

Schildkraut JM, Myers RH, Cupples LA, Kiely DK, Kannel WB. Coronary risk associated with age and sex of parental heart disease in the Framingham Study. Am J Cardiol 1989;64:555-559.

Schoenborn CA, Stommel M. Adherence to the 2008 adult physical activity guidelines and mortality risk. Am J Prev Med . 2011;40:514–521.

Schunkert H, König IR, Kathiresan S, et al. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. Nat Genet 2011;43:333-338.

Song Y, Stampfer MJ, Liu S. Meta-analysis: apolipoprotein E genotypes and risk for coronary heart disease. Ann Intern Med 2004;141:137–147.

Stallones RA. The association between tobacco smoking and coronary heart disease. Int J Epidemiol 2015;44:735-743.

Stary HC, Chandler AB, Glagov S, et al. A definition of initial, fatty streak, and intermediate lesions of atherosclerosis. A report from the Committee on Vascular Lesions of the Council on Arteriosclerosis, American Heart Association. Circulation 1994;89:2462-2478.

Scarborough Peter, Wickramasinghe Kremlin, Bhatnagar Prachi, Mike Rayner. Trends in coronary heart disease, 1961-2011. British Heart Foundation: London 2011. Available from: https://www.bhf.org.uk/publications/statistics/trends-in-coronary-heart-disease-1961-2011

Snowden CB, McNamara PM, Garnson RJ, Feinleib M, Kannel WB, Epstein FH. Predicting coronary heart disease in siblings--a multivariate assessment: : the Framingham Heart Study. Am J Epidemiol 1982;115:217-222.

Stewart ST, Cutler DM, Rosen AB. Forecasting the effects of obesity and smoking on U.S. life expectancy. N Engl J Med 2009;361:2252–2260.

Supreme Scientific Health Council & Ministry of Health and Welfare of Greece. Dietary guidelines for adults in Greece. Archives of Hellenic Medicine 1999;16:516-524.

Tada H, Melander O, Louie JZ, et al. Risk prediction by genetic risk scores for coronary heart disease is independent of self-reported family history. Eur Heart J 2015;37:561-567.

ten Kate LP, Boman H, Daiger SP, Motulsky AG. Familial aggregation of coronary heart disease and its relation to known genetic risk factors. Am J Cardiol 1982;50:945-953.

Thanassoulis G, Peloso GM, Pencina MJ, et al. A genetic risk score is associated with incident cardiovascular disease and coronary artery calcium: the Framingham Heart Study. Circ Cardiovasc Genet 2012;5:113-121.

The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. Am J Epidemiol 1989;129:687-702.

Thompson PD, Buchner D, Pina IL, et al. Exercise and physical activity in the prevention and treatment of atherosclerotic cardiovascular disease: a statement from the Council on Clinical Cardiology (Subcommittee on Exercise, Rehabilitation, and Prevention) and the Council on Nutrition, Physical Activity, and Metabolism (Subcommittee on Physical Activity). Circulation 2003;107:3109-3116.

Tikkanen E, Havulinna AS, Palotie A, Salomaa V, Ripatti S. Genetic risk prediction and a 2-stage risk screening strategy for coronary heart disease. Arterioscler Thromb Vasc Biol 2013;33:2261-2266.

Towfighi A, Zheng L, Ovbiagele B. Sex-specific trends in midlife coronary heart disease risk and prevalence. Arch Intern Med 2009;169:1762-1766.

Threapleton DE, Greenwood DC, Evans CE, et al. Dietary fibre intake and risk of cardiovascular disease: systematic review and meta-analysis. BMJ 2013;347:f6879.

Tucker K. Dietary patterns, approaches, and multicultural perspective. Appl Physiol Nutr Metab 2010; 35: 211-218.

US Department of Health and Human Services . How Tobacco Smoke Causes Disease: The Biology and Behavioral Basis for Smoking-Attributable Disease: A Report of the Surgeon General . Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 2010 .

Vasan RS, Larson MG, Leip EP, et al. Impact of high-normal blood pressure on the risk of cardiovascular disease. N Engl J Med 2001 ;345:1291-1297.

Weijmans M, van der Graaf Y, Reitsma JB, Visseren FL. Paternal or maternal history of cardiovascular disease and the risk of cardiovascular disease in offspring. A systematic review and meta-analysis. Int J Cardiol 2015, 20;179:409-416.

Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007;447:661-678.

Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. Circulation 1998;97:1837-1847.

Wong BW, Meredith A, Lin D, McManus BM. The biological role of inflammation in atherosclerosis. Can J Cardiol 2012;28:631-641.

Wong ND. Epidemiological studies of CHD and the evolution of preventive cardiology. Nat Rev Cardiol 2014;11: 276-289.

World Health Organization. Obesity and Overweight. Geneva 2016 Available from: http://www.who.int/mediacentre/factsheets/fs311/en/

World Health Organization. Guidelines: Sugars Intake for Adults and Children. Geneva 2015. Available from: http://www.who.int/nutrition/publications/guidelines/sugars_intake/en/

World Health Organization. Global Atlas on Cardiovascular Diseases Prevention and Control. Geneva 2011. Available from: http://www.who.int/cardiovascular_diseases/publications/atlas_cvd/en/

World Health Organization. Global health risks: mortality and burden of disease attributable to major risks. Geneva 2009. Available from: http://www.who.int/healthinfo/global_burden_disease/global_health_risks/en/

World Health Organization. Waist Circumference and Waist-Hip Ratio, Report of a WHO Expert Consultation. Geneva 2008. Available from: http://apps.who.int/iris/bitstream/10665/44583/1/9789241501491_eng.pdf

World Health Organisation. Global health risks 2009. Geneva 2009. Available from: http://www.who.int/healthinfo/global_burden_disease/global_ health_risks/en/index.html, Accessed December 2009.

Yatsuya H, Li Y, Hilawe EH, et al. Global trend in overweight and obesity and its association with cardiovascular disease incidence. Circ J 2014;78:2807-2818.

Yiannakouris N, Katsoulis, M, Trichopoulou A, Ordovas JM and Trichopoulos D. Additive influence of genetic predisposition and conventional risk factors in the incidence of coronary heart disease: a population-based study in Greece. BMJ Open 2014; 5:e004387.

Yusuf S, Hawken S, Ounpuu S, et al. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. Lancet 2004;364:937-952.

Zhao Y, Wu K, Zheng J, Zuo R, Li D. Association of coffee drinking with all-cause mortality: a systematic review and meta-analysis. Public Health Nutr 2015;18:1282-1291

Zheng J, Huang T, Yu Y, Hu X, Yang B, Li D. Fish consumption and CHD mortality: an updated meta-analysis of seventeen cohort studies. Public Health Nutr 2012;15: 725–737.

Zdravkovic S, Wienke A, Pedersen NL, Marenberg ME, Yashin, AI, De Faire U. Heritability of death from coronary heart disease: a 36-year follow-up of 20 966 Swedish twins. J Intern Med 2002:252;247-254.

Chapter 6

Supplementary Material

6 SUPPLEMENTARY MATERIAL

6.1. Paper 2

Lifestyle may modify the glucose-raising effect of genetic loci. A study in the Greek population

Supplementary Methods

Study population

All participants were informed verbally as well as in a written form before they consent to participate in the study [1]. The study was approved by the Ethics Committee of Harokopio University of Athens and the Greek Ministry of Education. Depending on the data availability of the parameters used in each analysis with respect to glucose levels, the sample number ranges from 533 to 1132 individuals. For the GPS association with glucose levels 552 individuals contributed, for the associations of the wGRS with glucose levels the available sample comprised of 1132 subjects and for the interaction tests between the GPS and the wGRS 533 subjects were included.

Adiposity and Biochemical measurements

Adiposity measurements included height, weight, waist circumference and hip circumference; obtained from trained personnel. BMI was calculated as weight (kg) / height squared (m^2) and Waist-to-Hip Ratio (WHR: a measure of fat distribution) was also calculated. Fasting venous serum was used for the glycemic measurements. Glucose levels were measured using the enzymatic colorimetric assay (ACE analyzer). Serum insulin was measured via immunofluorescence on an automatic analyzer with direct chemiluminescence immunometric assay, sandwich type, of two points, utilizing constant quantities of two antibodies according to the manufacturer's instructions (ST AIA pack IRI, Tosoh AIA System Analyzers, San Francisco, CA). Insulin resistance was estimated using the homeostasis model assessment (HOMA-IR) with the following equation: HOMA-IR = fasting insulin (μIU/ml) × fasting glucose (mmol/L)/22.5 [2].

Genotyping

Genomic DNA was extracted from whole blood using the salting-out method [3]. DNA samples were genotyped using the Illumina Metabochip at the Wellcome Trust Sanger Institute, Hinxton, UK. The genotype calling algorithm was GenoSNP. The exclusion criteria

included: sample call rate < 95%, SNP call rate 98%, sex discrepancies, ethnic outliers as assessed by performing multidimensional scaling (MDS) in PLINK [3] and samples with genome-wide heterozygosity higher than ± 3SD. Three SNPs (rs2657879, rs2302593, rs7708285) failed the 98% call rate threshold and were replaced by imputed data. Imputation was performed using the cosmopolitan 1000 Genome panel and IMPUTE2 [4]. Prior phasing was done using SHAPEIT [5]. Estimated probabilities for these 3 SNPs were converted to dosages using gtool (http://www.well.ox.ac.uk/~cfreeman/software/gwas/gtool.html). In the imputed format each SNP is represented as a set of three probabilities which correspond to allele pairs. Using a threshold cutoff for the probability values (0.9), genotypes were expressed as pairs of 1, 2, 0 (Supplemental table S4).

Assessment of dietary patterns and physical activity

Dietary intake was estimated by the use of a food frequency questionnaire. The questionnaires evaluate food and drink consumption for the last one year and consisted of 172 food items representative of the major food groups (red meat, poultry, fish, seafood, dairy products, refined and non-refined grains, raw and cooked vegetables, fruits, eggs, sweets, beverages). Data from the questionnaires were analysed for their macro and micronutrient content using the Nutritionist Pro, version 2.2 software (Axxya Systems-Nutritionist Pro, Stafford, TX, USA). Mean energy, macro and micronutrient intakes were estimated for all participants.

There was also an enrichment of The Nutritionist Pro food database by adding analyses of traditional Greek foods, recipes and nutrient information[6]. The consumption of all food items was recorded in terms of frequency (never or less than once a month, occasionally, sometimes, daily consumption) and quantity (no, small, medium abundant consumption) by the use of images representing standard portion sizes.

Food groups assessed for the present analysis and Glucose Preventive Score calculation are those associated with glucose levels in our sample (fruits and fresh juice, vegetables, soft drinks and beverages intake). Participants completed a physical activity questionnaire recall for the last seven days [7]. The degree of physical activity was assessed via the amount of activity recorded at work.

Modelling of Glucose Preventive Score (GPS)

The parameters we selected for the GPS calculation are proven to be associated with glucose levels in the literature or indicated a trend of association in our study.

Prospective studies show that higher intake of sugar-sweetened fruit juice may increase type 2 diabetes risk [8, 9]. Women consuming ≥ 1 sugar sweetened beverages (SSB) per day had a greater risk of developing type 2 diabetes compared to those consuming <1 SSB per month [10, 11]. SSB contribute to a high glycaemic load leading to insulin resistance and impaired β-cell function. In our study higher consumption of soft drinks and beverages with sugars was associated with increased glucose levels. There is evidence supporting that higher consumption of fruits such as blueberries, apples and grapes could be linked to lowering glucose levels [8]. It is consistently reported that fruit and vegetable consumption is an important dietary aspect, associated with a lower risk for glucose concentrations [12-14]. Several meta-analysis results support that an increase in the daily consumption of green leafy vegetables could have a glucose decreasing effect [15-17]. In addition to diet, physical activity plays a crucial role in the management of glucose homeostasis and higher exercise levels could result in significant improvements of glycaemic control. Even a moderate increase in physical activity could result in improvement of insulin sensitivity [18-20]. There is evidence in the Greek population supporting that physical activity plays a significant role in insulin sensitivity and could improve the obesity consequences on insulin resistance [21].

Statistical methods and analysis

We evaluated the association between the preventive score and glycemic traits concentrations using multivariate linear regression, with the preventive score modeled as a continuous variable. For energy improbable intake, individuals with ± 3SD intake were excluded. Models were adjusted for age, sex, BMI and total energy intake.

Linear regression models assuming an additive genetic model of inheritance, were used to test for the association of the 20 SNPs with blood glucose levels in our study. The reported effect sizes, SE and p-values result after adjustment for age and sex (Supplementary Table S4).

Regression coefficients from the linear regression models we applied to test the association between the scores and the traits tested, represent the difference in glucose (mmol/L) or insulin (ln-pmol/L) concentration per 1-unit increase in the preventive score. We also assessed the associations of the wGRS with the glycemic traits, adjusting for age, sex and BMI.

The interaction regression coefficients represent the difference in glucose (mmol/L) levels, per 1 unit increase in the preventive lifestyle score, per weighted copy of a glucose raising allele, estimated through the wGRS.

To further characterize the direction of the interaction we conducted stratified analysis in which the study individuals were stratified into two groups according to the median of the GPS score. Linear regression models were used to test for the associations between the wGRS for each group with glucose levels, adjusting for age, sex and BMI.

For the environmental score the coefficient of determination (R^2) was used as a measure to express the proportion of total variation explained by the model. We used 'Anova' to produce the incremental sums of squares table, to get the incremental variance explained. The proportion of variance explained by the GPS was calculated after scaling the sum squares by 100 divided by their sum.

Genotype by environment (GxE) variance contribution to the total variance of glucose levels was estimated using a tool for genome-wide Complex trait Analysis (GCTA) [22]. GxE interaction and the main effects of the genetic factors are treated as random effects in the model. The main effects of the E factors are treated as fixed effects, while using adjustments for confounding factors (age, sex, BMI). The "-gxe" option was used for the estimation of variance explained by the GxE interaction. The E factor which is the environmental term in GxE interaction, was the GPS score. GxE heritability was calculated as the GxE variance divided by the total phenotypic variance.

We used Quanto v1.2.4 for power calculations (http://hydra.usc.edu/gxe/). Our model has 80% power to detect a 0.05 mmol/L change in glucose levels under the effect of the preventive score, 80% power to detect a 0.06 mmol/L change in glucose levels under the effect of the wGRS and 80% power to detect a 0.06 mmol/L change in glucose levels under the interaction effect of the environmental and genetic risk score. Power calculations are provided in Supplemental Figure S3.

Supplemental Table S1: Determinants of Glucose Preventive Score

GPS parameters	Impact	1st tertile[a]	2nd tertile[a]	3rd tertile[a]
Moving (hours during work/day)	+	0.257± 0.573	3.341±0.575	7.701±2.914
Fruits and fresh juice (servings/day)	+	0.227±0.130	1.647±0.963	6.038±2.270
Vegetables (servings/day)	+	0.615±0.337	2.996±1.241	7.825±1.921
Soft drinks and beverages with sugar (servings/day)	-	0.124±0.258	1.792± 0.291	3.538±0.896

Data present tertile ranks of determinants comprising the GPS score (0, 1 and 2 points, for lowest to highest tertiles, respectively) and reversed tertile ranks of unfavourable determinant (2, 1, and 0, for lowest to highest tertile respectively).

[a]Tertile values are presented as means ± SD.

Supplemental Table S2: Descriptive characteristics of participants

Descriptive characteristics					
	Total Sample	**Males**	**Females**	**P**	**N**
Age (years)	57.768 ± 14.121	56.67 ± 13.24	59.12 ± 15.08	1.3×10^{-3}	
Energy Intake (kcal/d)	1912.646 ± 1042.727	2061.719 ± 1100.749	1749.963 ± 952.7023	1.7×10^{-5} [d]	685
GPS	3.729 ± 1.351	3.791 ± 1.457	3.652 ± 1.216	0.458 [c]	585
wGRS	22.285 ± 2.351	22.384 ± 2.366	22.199 ± 2.324	0.150 [d]	1132
GRS	19.670 ± 2.613	19.747 ± 2.613	19.621 ± 2.509	0.417 [d]	1132
Glycaemic Traits					
Glucose (mmol/L)	5.318 ± 0.645	5.402 ± 0.650	5.221 ± 0.617	1.9×10^{-9} [c]	647
Insulin (pmol/L) [b]	71.582 ± 40.585	78.270 ± 41.809	66.116 ± 38.753	0.210 [c]	305
Insulin (ln-pmol/L)	4.123 ± 0.552	4.219 ± 0.551	4.045 ± 0.541	0.210 [c]	
HOMA.IR	2.470 ± 1.498	2.76± 1.63	2.23 ± 1.34	0.050 [c]	302
GPS parameters					
Moving (hours during work/day)	1.127 ± 2.386	1.588 ± 2.886	0.5905 ± 1.443	3.2×10^{-6} [c]	646
Fruits and fresh juice (servings/day) [a]	1.000 (1.160)	1.000 (1.710)	1.170 (2.170)	5.5×10^{-7} [c]	605
Vegetables (servings/day) [a]	1.900 (2.740)	2.200 (3.820)	1.720 (2.620)	0.230 [c]	606
Soft drinks and beverages with sugar (servings/day)	0.308 ± 0.731	0.369 ± 0.773	0.240 ± 0.678	0.170 [c]	606

Glucose Preventive Score, weighted Genetic Risk Score, glycaemic traits and GPS components. Data are presented as means ± SD, median (intertertile range)[a], for the Total sample and males/females. [b]Insulin was analysed on the natural log scale and back transformed to the geometric scale for presentation. P values [c]adjusted for age, BMI and energy intake, [d]P values adjusted for age and BMI.

Supplemental Table S3: Linear regression models of each component of the Glucose Preventive Score with Glucose levels.

GPS parameters	Beta [a]	SE	P	N
Moving (hours during work/day)	-0.011	0.011	0.030	612
Fruits and fresh juice (servings/day)	-0.033	0.014	0.023	570
Vegetables (servings/day)	-0.035	0.013	0.008	571
Soft drinks and beverages with sugar (servings/day)	0.085	0.034	0.013	571

[a]Beta coefficient and standard error for the estimated difference in fasting glucose (mmol/L) concentration for each determinant comprising the preventive score, adjusted for age, sex, BMI and total energy intake.

Supplemental Table S4: THISEAS association summary statistics for glucose levels associated loci.

						MAGIC				THISEAS				
SNP	Candidate Gene	Chr	Position (bp)	Effect allele	Other Allele	EAF	beta[a]	SE	P-value	EAF	HWE	beta[a]	SE	P-value
rs10747083	P2RX2	12	132000000	A	G	0.663	0.013	0.002	7.57E-09	0.681	0.440	0.023	0.028	0.418
rs10811661	CDKN2A	9	22124094	T	C	0.82	0.024	0.003	5.65E-18	0.805	0.818	0.023	0.033	0.485
rs11603334	CENTD2	11	72110633	G	A	0.833	0.019	0.003	1.12E-11	0.891	0.224	0.018	0.042	0.675
rs11619319	PDX1	13	27385599	G	A	0.226	0.02	0.002	1.33E-15	0.213	0.433	-0.081	0.032	**0.011**
rs11715915	PRKAR2A	3	49430334	C	T	0.675	0.012	0.002	4.90E-08	0.655	0.749	0.074	0.027	**0.007**
rs16913693	CTNNAL1	9	111000000	T	G	0.973	0.043	0.007	3.51E-11	0.984	1.000	0.168	0.095	0.079
rs17762454	RREB1	6	7158199	T	C	0.262	0.012	0.002	1.88E-07	0.246	0.742	0.014	0.031	0.650
rs2302593b	GIPR	19	50888474	C	G	0.503	0.014	0.002	9.26E-10	0.506	NA	0.008	0.026	0.772
rs2657879b	GLS2	12	55151605	G	A	0.182	0.012	0.003	5.69E-06	0.201	NA	-0.044	0.035	0.206
rs3783347	C14orf68	14	99909014	G	T	0.789	0.017	0.003	1.32E-10	0.828	0.611	0.010	0.035	0.776
rs3829109	LHX3	9	138000000	G	A	0.707	0.017	0.003	1.13E-10	0.804	0.280	0.004	0.033	0.894
rs4869272	PCSK1	5	95565204	T	C	0.689	0.018	0.002	1.02E-15	0.665	0.587	0.025	0.028	0.358
rs576674	KL	13	32452302	G	A	0.154	0.017	0.003	2.26E-08	0.191	0.050	0.075	0.035	**0.031**
rs6072275	LPIN3	20	39177319	A	G	0.165	0.016	0.003	1.66E-08	0.148	0.019	0.068	0.038	0.071
rs6113722	FOXA2	20	22505099	G	A	0.957	0.035	0.005	2.49E-11	0.920	0.246	0.075	0.049	0.128
rs6943153	GRB10	7	50759073	T	C	0.335	0.015	0.002	1.63E-12	0.238	0.385	-0.007	0.031	0.818
rs7651090	IGF2BP2	3	187000000	G	A	0.306	0.013	0.002	1.75E-08	0.320	0.084	-0.017	0.028	0.536
rs7708285b	ZBED3	5	76461623	G	A	0.271	0.011	0.003	4.89E-06	0.284	NA	0.011	0.029	0.699
rs9368222	CDKAL1	6	20794975	A	C	0.283	0.014	0.002	1.00E-09	0.282	0.342	0.057	0.029	**0.048**
rs983309	PPP1R3B	8	9215142	T	G	0.117	0.026	0.003	6.29E-15	0.088	0.655	-0.056	0.047	0.238

Supplemental table S4:
[a]Effect sizes (beta) and SE are given for Glucose (mmol/L), [b] Imputation quality: info > 0.97.
Chr = chromosome, bp = base pairs, EAF = effect allele frequency, SE = standard error
Allelic test p, beta and SE are shown for each single SNP.
Effect sizes (beta) are reported for the effect allele. Bold high-lighted loci indicated nominal evidence for association with glucose levels.
Results were obtained using linear regression models assuming an additive effect. Adjustments: age and sex.
EAF: Effect Allele Frequency, HWE: Hardy Weinberg Equilibrium

Supplemental Table S5: THISEAS association summary statistics for glucose levels associated loci.

						MAGIC				THISEAS				
SNP	Candidate Gene	Chr	Position (bp)	Effect allele	Other Allele	EAF	beta[a]	SE	P-value	EAF	HWE	beta[a]	SE	P-value
rs10747083	P2RX2	12	132000000	A	G	0.663	0.013	0.002	7.57E-09	0.681	0.440	0.008	0.028	0.766
rs10811661	CDKN2A	9	22124094	T	C	0.82	0.024	0.003	5.65E-18	0.805	0.818	0.022	0.033	0.494
rs11603334	CENTD2	11	72110633	G	A	0.833	0.019	0.003	1.12E-11	0.891	0.224	0.024	0.043	0.576
rs11619319	PDX1	13	27385599	G	A	0.226	0.02	0.002	1.33E-15	0.213	0.433	-0.047	0.032	0.143
rs11715915	PRKAR2A	3	49430334	C	T	0.675	0.012	0.002	4.90E-08	0.655	0.749	0.079	0.028	**0.005**
rs16913693	CTNNAL1	9	111000000	T	G	0.973	0.043	0.007	3.51E-11	0.984	1.000	0.185	0.099	**0.063**
rs17762454	RREB1	6	7158199	T	C	0.262	0.012	0.002	1.88E-07	0.246	0.742	0.023	0.031	0.444
rs2302593b	GIPR	19	50888474	C	G	0.503	0.014	0.002	9.26E-10	0.506	NA	0.016	0.026	0.534
rs2657879b	GLS2	12	55151605	G	A	0.182	0.012	0.003	5.69E-06	0.201	NA	-0.038	0.035	0.284
rs3783347	C14orf68	14	99909014	G	T	0.789	0.017	0.003	1.32E-10	0.828	0.611	0.034	0.034	0.318
rs3829109	LHX3	9	138000000	G	A	0.707	0.017	0.003	1.13E-10	0.804	0.280	0.010	0.033	0.768
rs4869272	PCSK1	5	95565204	T	C	0.689	0.018	0.002	1.02E-15	0.665	0.587	0.024	0.027	0.386
rs576674	KL	13	32452302	G	A	0.154	0.017	0.003	2.26E-08	0.191	0.050	0.064	0.034	**0.064**
rs6072275	LPIN3	20	39177319	A	G	0.165	0.016	0.003	1.66E-08	0.148	0.019	0.052	0.038	0.168
rs6113722	FOXA2	20	22505099	G	A	0.957	0.035	0.005	2.49E-11	0.920	0.246	0.062	0.049	0.209
rs6943153	GRB10	7	50759073	T	C	0.335	0.015	0.002	1.63E-12	0.238	0.385	-0.012	0.031	0.696
rs7651090	IGF2BP2	3	187000000	G	A	0.306	0.013	0.002	1.75E-08	0.320	0.084	-0.027	0.028	0.332
rs7708285b	ZBED3	5	76461623	G	A	0.271	0.011	0.003	4.89E-06	0.284	NA	0.008	0.029	0.781
rs9368222	CDKAL1	6	20794975	A	C	0.283	0.014	0.002	1.00E-09	0.282	0.342	0.068	0.029	**0.018**
rs983309	PPP1R3B	8	9215142	T	G	0.117	0.026	0.003	6.29E-15	0.088	0.655	-0.065	0.047	0.169

Supplemental Table S5:

[a]Effect sizes (beta) and SE are given for Glucose (mmol/L), [b] Imputation quality: info > 0.97.

Chr = chromosome, bp = base pairs, EAF = effect allele frequency, SE = standard error; EAF: Effect Allele Frequency, HWE: Hardy Weinberg Equilibrium

Allelic test p, beta and SE are shown for each single SNP.

Effect sizes (beta) are reported for the effect allele. Bold high-lighted loci indicated nominal evidence for association with glucose levels.

Results were obtained using linear regression models assuming an additive effect. Adjustments: age, sex and BMI.

Supplemental Table S6: Associations[a] of the Genetic Risk Scores with glucose levels in the subsample used in the interaction analyses.

A. Associations[a] of the weighted Genetic Risk Score with glucose levels in the subsample used in the interaction analyses.

	Beta[b]	SE[b]	P	N[b]
Glycemic Traits				
Glucose (mmol/L)	0.071	0.029	0.016	533

B. Associations[a] of the unweighted Genetic Risk Score with glucose levels in the subsample used in the interaction analyses.

	Beta[b]	SE[b]	P	N[b]
Glycemic Traits				
Glucose (mmol/L)	0.057	0.027	0.039	533

[a]Adjusted for age, sex and BMI

Supplemental Figure S1: Scatterplot of the negative association between the Glucose Preventive Score (GPS) and glucose levels (mmol/L) in individuals without type 2 diabetes, after controlling for age, sex, BMI and total energy intake.

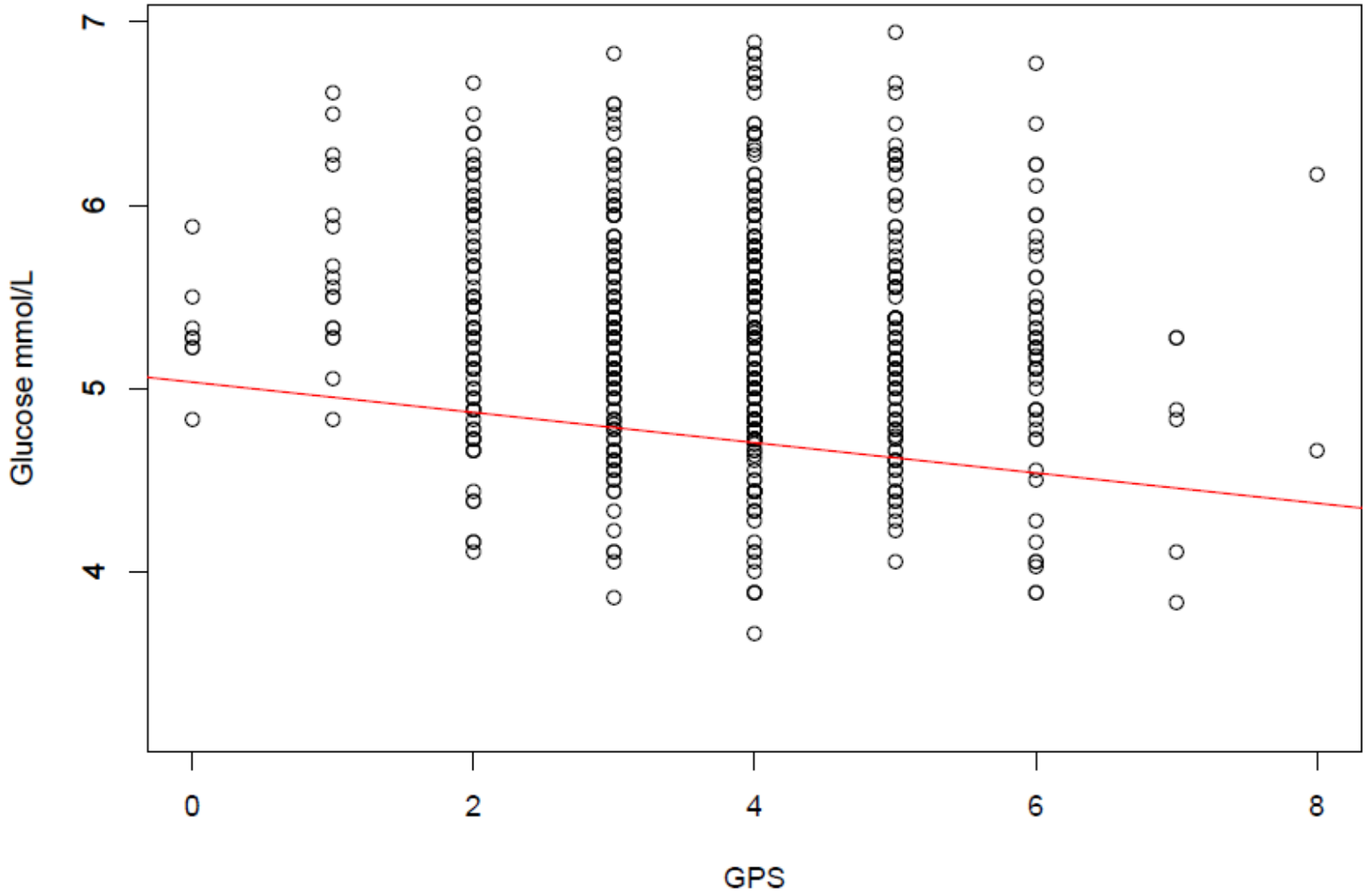
Supplemental Figure S2: Scatterplot of the positive correlation between the weighted genetic risk score (wGRS) and glucose levels (mmol/L) in individuals without type 2 diabetes, after controlling for age, sex and BMI.

Supplemental Figure S3: Estimated power of the analyses for alpha 0.05. Calculations were performed using Quanto v1.2.4. Lines represent the different power values to detect the effect: **A.** in the model for the association of the glucose preventive score (GPS) with the trait, **B.** in the model for the association of the weighted genetic risk score (wGRS) with the trait and **C.** in the interaction model for association of GPS x wGRS effect with the trait.

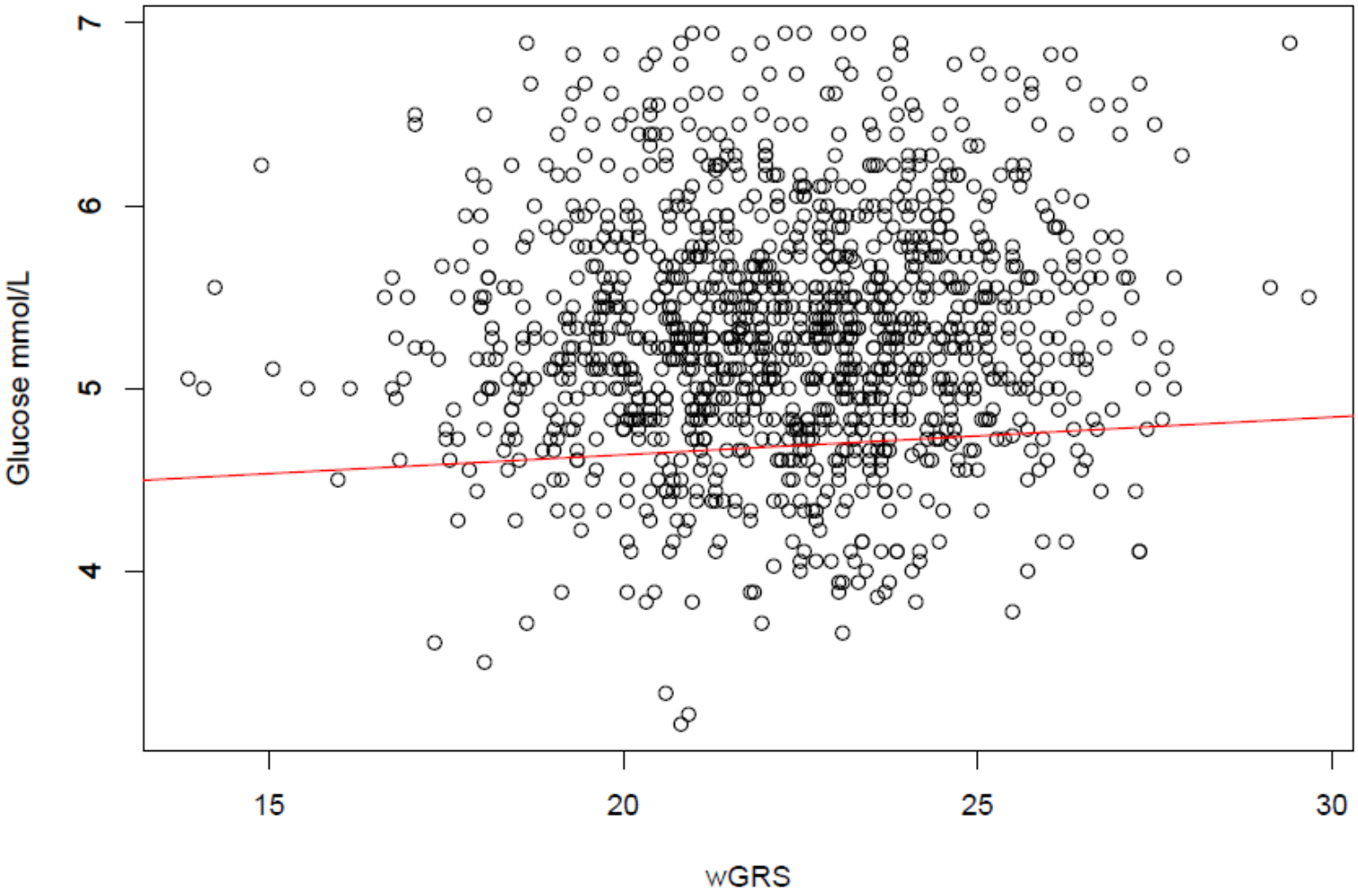
References

[1] Theodoraki EV, Nikopensius T, Suhorutsenko J, Peppes V, Fili P, Kolovou G, et al. Fibrinogen beta variants confer protection against coronary artery disease in a Greek case-control study. BMC Med Genet. 2010;11:28.
[2] Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. Diabetologia. 1985;28:412-9.
[3] Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81:559-75.
[4] Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet. 2009;5:e1000529.
[5] Delaneau O, Marchini J, Zagury JF. A linear complexity phasing method for thousands of genomes. Nat Methods. 2012;9:179-81.
[6] Kafatos A, Verhagen H, Moschandreas J, Apostolaki I, Van Westerop JJ. Mediterranean diet of Crete: foods and nutrient content. J Am Diet Assoc. 2000;100:1487-93.
[7] Liese AD, Ma X, Maahs DM, Trilk JL. Review. Journal of Sport and Health Science. 2013;2:21-38.
[8] Muraki I, Imamura F, Manson JE, Hu FB, Willett WC, van Dam RM, et al. Fruit consumption and risk of type 2 diabetes: results from three prospective longitudinal cohort studies. BMJ (Clinical research ed). 2013;347:f5001.
[9] Montonen J, Jarvinen R, Knekt P, Heliovaara M, Reunanen A. Consumption of sweetened beverages and intakes of fructose and glucose predict type 2 diabetes occurrence. J Nutr. 2007;137:1447-54.
[10] Schulze MB, Manson JE, Ludwig DS, Colditz GA, Stampfer MJ, Willett WC, et al. Sugar-sweetened beverages, weight gain, and incidence of type 2 diabetes in young and middle-aged women. Jama. 2004;292:927-34.
[11] Bazzano LA, Li TY, Joshipura KJ, Hu FB. Intake of fruit, vegetables, and fruit juices and risk of diabetes in women. Diabetes Care. 2008;31:1311-7.
[12] Hodge AM, English DR, O'Dea K, Giles GG. Dietary patterns and diabetes incidence in the Melbourne Collaborative Cohort Study. Am J Epidemiol. 2007;165:603-10.
[13] van Dam RM, Rimm EB, Willett WC, Stampfer MJ, Hu FB. Dietary patterns and risk for type 2 diabetes mellitus in U.S. men. Ann Intern Med. 2002;136:201-9.
[14] Montonen J, Knekt P, Harkanen T, Jarvinen R, Heliovaara M, Aromaa A, et al. Dietary patterns and the incidence of type 2 diabetes. Am J Epidemiol. 2005;161:219-27.
[15] Carter P, Gray LJ, Troughton J, Khunti K, Davies MJ. Fruit and vegetable intake and incidence of type 2 diabetes mellitus: systematic review and meta-analysis. BMJ (Clinical research ed). 2010;341:c4229.
[16] Cooper AJ, Forouhi NG, Ye Z, Buijsse B, Arriola L, Balkau B, et al. Fruit and vegetable intake and type 2 diabetes: EPIC-InterAct prospective study and meta-analysis. Eur J Clin Nutr. 2012;66:1082-92.
[17] Hamer M, Chida Y. Intake of fruit, vegetables, and antioxidants and risk of type 2 diabetes: systematic review and meta-analysis. J Hypertens. 2007;25:2361-9.
[18] Chudyk A, Petrella RJ. Effects of exercise on cardiovascular risk factors in type 2 diabetes: a meta-analysis. Diabetes Care. 2011;34:1228-37.
[19] Colberg SR, Sigal RJ, Fernhall B, Regensteiner JG, Blissmer BJ, Rubin RR, et al. Exercise and type 2 diabetes: the American College of Sports Medicine and the American Diabetes Association: joint position statement executive summary. Diabetes Care. 2010;33:2692-6.
[20] Thomas DE, Elliott EJ, Naughton GA. Exercise for type 2 diabetes mellitus. Cochrane Database Syst Rev. 2006:Cd002968.
[21] Kavouras SA, Panagiotakos DB, Pitsavos C, Chrysohoou C, Anastasiou CA, Lentzas Y, et al. Physical activity, obesity status, and glycemic control: The ATTICA study. Med Sci Sports Exerc. 2007;39:606-11.
[22] Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet.88:76-82.

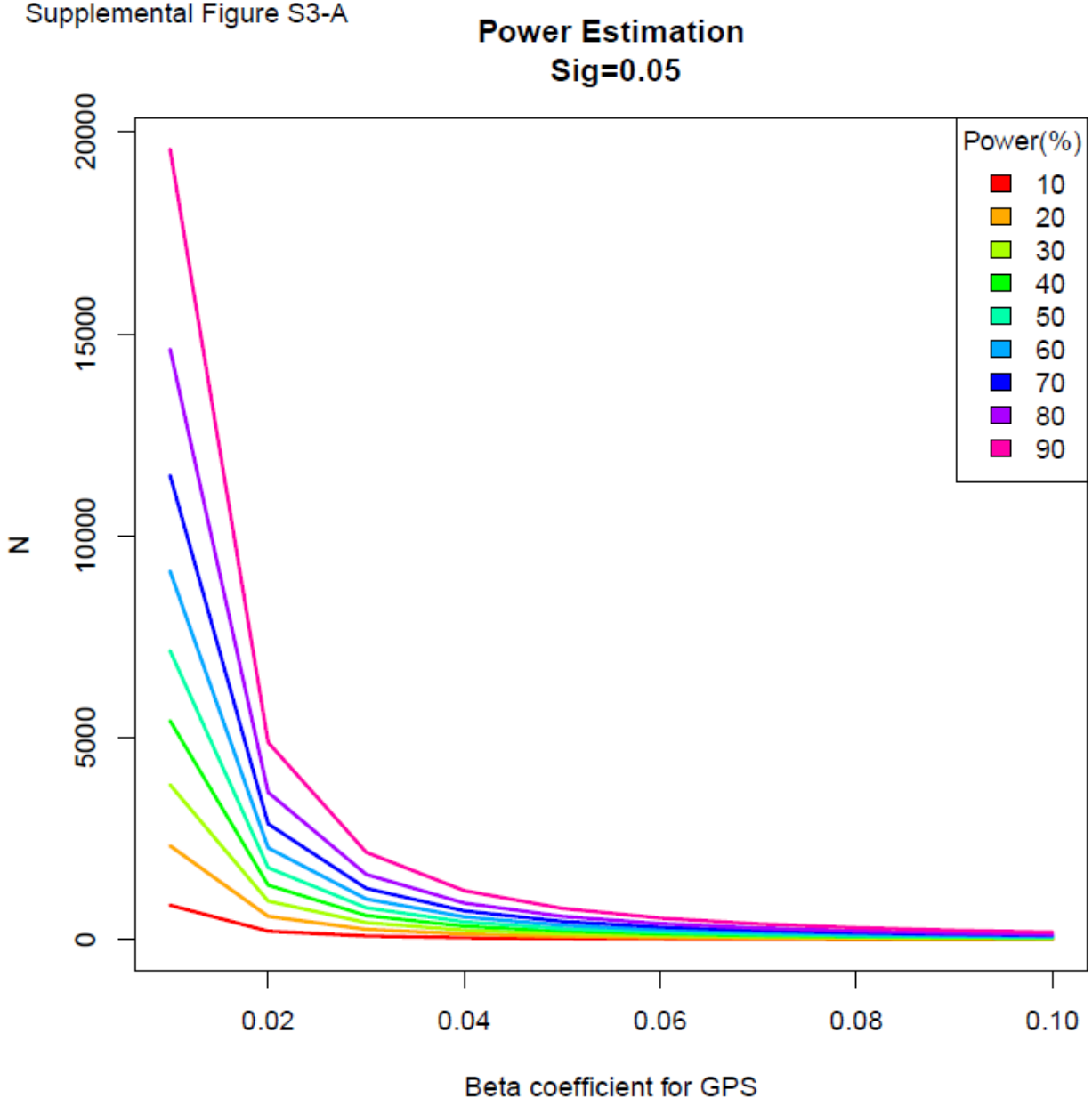
Supplemental Figure S1



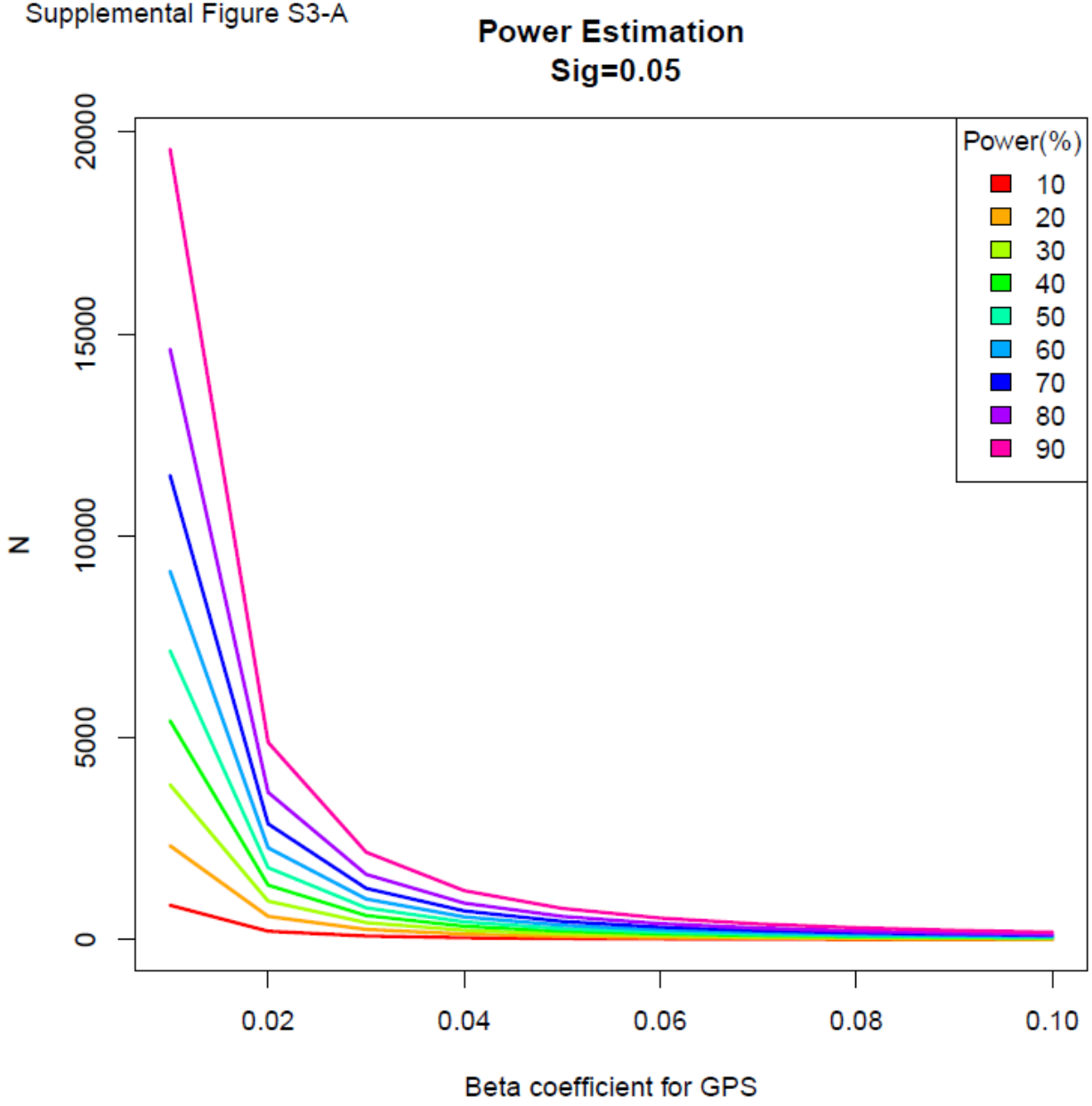
Supplemental Figure S2



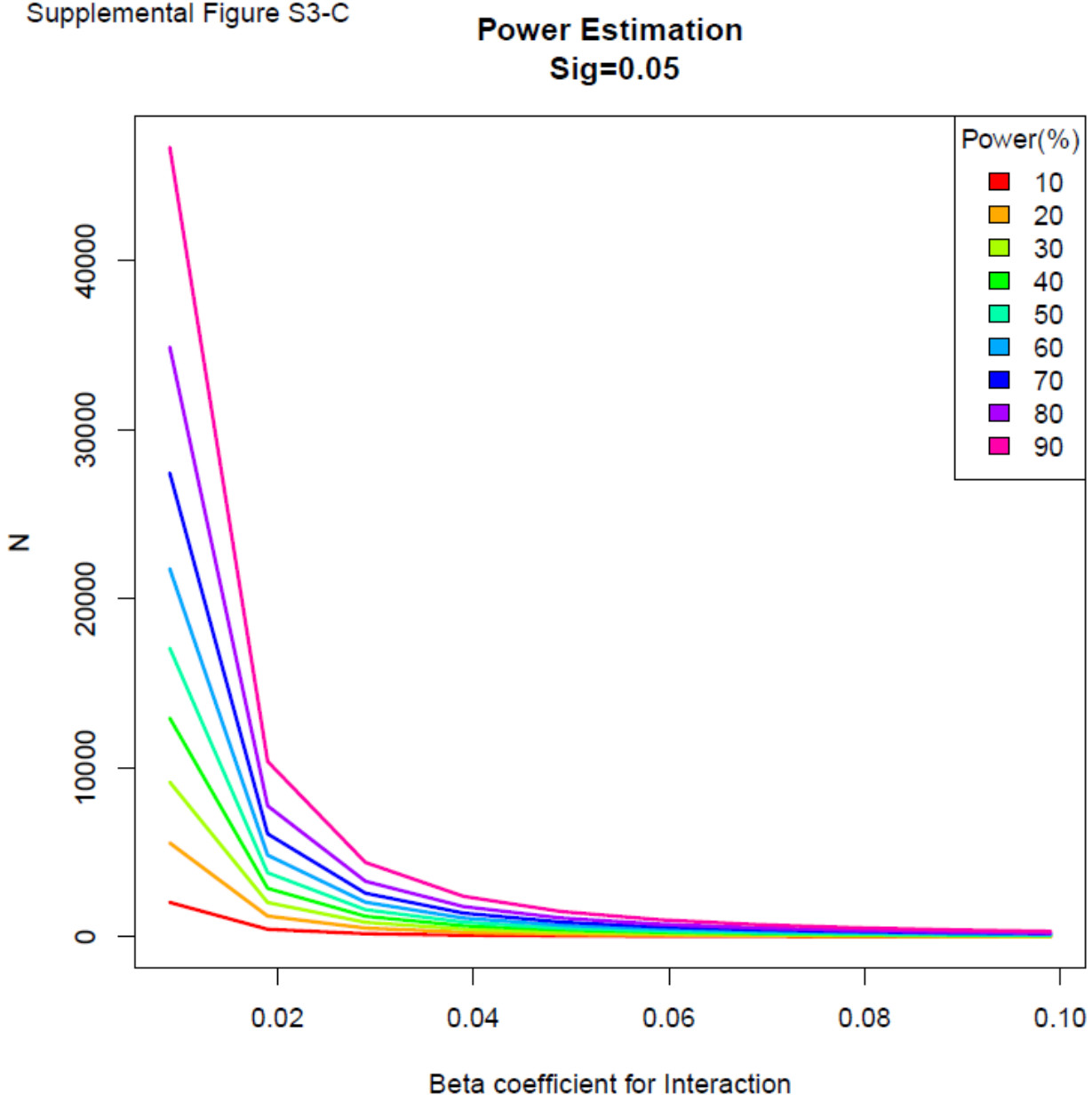
Supplemental Figure S3-A



Supplemental Figure S3-A



Supplemental Figure S3-C



6.2. Paper 3

Additive influence of genetic variants increases the risk of coronary artery disease in a Greek sample: The THISEAS study

Supplementary Table 1. Study-specific genotyping information

				Genotyping				Imputation			
SNP	Candidate Gene	Chr	Effect/ Non effect allele	SNP call rate (%)	p for HWE	Platform	Genotype calling algorithm	SNPs that met QC criteria (N)	Haplotyping software (version)	Imputation software (version)	Imputation nfo
rs602633	SORT1	1	C/A	N/A	N/A	N/A	N/A	652.329	SHAPEIT (V2)	IMPUTE (V2)	0,98910
rs11206510	PCSK9	1	T/C	100	1,0000	Illumina Omni Express	Illuminus	N/A	N/A	N/A	N/A
rs6725887	WDR12	2	C/T	N/A	N/A	N/A	N/A	652.329	SHAPEIT (V2)	IMPUTE (V2)	1,00000
rs9818870	MRAS	3	T/C	100	0,3571	Illumina Omni Express	Illuminus	N/A	N/A	N/A	N/A
rs12190287	TCF21	6	C/G	N/A	N/A	N/A	N/A	652.329	SHAPEIT (V2)	IMPUTE (V2)	0,98348
rs2048327	SLC22A3-LPAL2-LPA	6	G/A	100	0,3365	Illumina Omni Express	Illuminus	N/A	N/A	N/A	N/A
rs11556924	ZC3HC1	7	C/T	100	0,0552	Illumina Omni Express	Illuminus	N/A	N/A	N/A	N/A
rs1333049	CDKN2BAS1	9	C/G	N/A	N/A	N/A	N/A	652.329	SHAPEIT (V2)	IMPUTE (V2)	0,89464
rs579459	ABO	9	C/T	100	0,2517	Illumina Omni Express	Illuminus	N/A	N/A	N/A	N/A
rs12413409	CYP17A1-CNNM2-NT5C2	10	G/A	N/A	N/A	N/A	N/A	652.329	SHAPEIT (V2)	IMPUTE (V2)	0,99472
rs2505083	KIAA1462	10	C/T	N/A	N/A	N/A	N/A	652.329	SHAPEIT (V2)	IMPUTE (V2)	0,87851
rs974819	PDGFD	11	A/G	100	0,4565	Illumina Omni Express	Illuminus	N/A	N/A	N/A	N/A
rs3184504	SH2B3	12	T/C	100	0,03231	Illumina Omni Express	Illuminus	N/A	N/A	N/A	N/A
rs9515203	COL4A1-COL4A2	13	T/C	100	0,8454	Illumina Omni Express	Illuminus	N/A	N/A	N/A	N/A
rs2895811	HHIPL1	14	C/T	N/A	N/A	N/A	N/A	652.329	SHAPEIT (V2)	IMPUTE (V2)	0,99656
rs12936587	RAI1-PEMT-RASD1	17	G/A	N/A	N/A	N/A	N/A	652.329	SHAPEIT (V2)	IMPUTE (V2)	0,98541

Supplementary Table 1. Study-specific genotyping information **(continued)**

				SNP call rate (%)	p for HWE	Platform	Genotype calling algorithm	SNPs that met QC criteria (N)	Haplotyping software (version)	Imputation software (version)	Imputation nfo
rs1122608	LDLR	19	G/T	N/A	N/A	N/A	N/A	652.329	SHAPEIT (V2)	IMPUTE (V2)	0,98864
rs9982601	Gene desert (KCNE2)	21	T/C	N/A	N/A	N/A	N/A	652.329	SHAPEIT (V2)	IMPUTE (V2)	0,97750
rs17114036	PPAP2B	1	A/G	N/A	N/A	N/A	N/A	652.329	SHAPEIT (V2)	IMPUTE (V2)	0,98235
rs12205331	ANKS1A	6	C/T	N/A	N/A	N/A	N/A	652.329	SHAPEIT (V2)	IMPUTE (V2)	0,94352
rs9369640	PHACTR1	6	A/C	100	0,6987	Illumina Omni Express	Illuminus	N/A	N/A	N/A	N/A
rs2047009	CXCL12	10	C/A	100	0,7050	Illumina Omni Express	Illuminus	N/A	N/A	N/A	N/A
rs11203042	LIPA	11	T/C	100	0,3697	Illumina Omni Express	Illuminus	N/A	N/A	N/A	N/A
rs15563	UBE2Z	17	C/T	100	1,0000	Illumina Omni Express	Illuminus	N/A	N/A	N/A	N/A
rs2281727	SMG6	17	C/T	N/A	N/A	N/A	N/A	652.329	SHAPEIT (V2)	IMPUTE (V2)	0,99940
rs2075650	ApoE-ApoC1	19	G/A	100	0,2550	Illumina Omni Express	Illuminus	N/A	N/A	N/A	N/A
rs17464857	MIA3	1	T/G	100	1,0000	Illumina Omni Express	Illuminus	N/A	N/A	N/A	N/A
rs12539895	7q22	7	A/C	N/A	N/A	N/A	N/A	652.329	SHAPEIT (V2)	IMPUTE (V2)	0,99739
rs9326246	ZNF259-APOA5-APOA1	11	C/G	N/A	N/A	N/A	N/A	652.329	SHAPEIT (V2)	IMPUTE (V2)	0,99489
rs7173743	ADAMTS7	15	T/C	100	0,8014	Illumina Omni Express	Illuminus	N/A	N/A	N/A	N/A
rs4845625	IL6R	1	T/C	100	0,5703	Illumina Omni Express	Illuminus	N/A	N/A	N/A	N/A
rs515135	APOB	2	G/A	N/A	N/A	N/A	N/A	652.329	SHAPEIT (V2)	IMPUTE (V2)	0,99408
rs2252641	ZEB2-ACO74093.1	2	G/A	100	0,7043	Illumina Omni Express	Illuminus	N/A	N/A	N/A	N/A
rs1561198	VAMP5-VAMP8-GGCX	2	A/G	N/A	N/A	N/A	N/A	652.329	SHAPEIT (V2)	IMPUTE (V2)	0,99624
rs7692387	GUCY1A3	4	G/A	N/A	N/A	N/A	N/A	652.329	SHAPEIT (V2)	IMPUTE (V2)	0,99734
rs273909	SLC22A4-SLC22A5	5	C/T	N/A	N/A	N/A	N/A	652.329	SHAPEIT (V2)	IMPUTE (V2)	0,94627

Supplementary Table 1. Study-specific genotyping information **(continued)**											
				SNP call rate (%)	**p for HWE**	**Platform**	**Genotype calling algorithm**	**SNPs that met QC criteria (N)**	**Haplotyping software (version)**	**Imputation software (version)**	**Imputation nfo**
rs10947789	KCNK5	6	T/C	N/A	N/A	N/A	N/A	652.329	SHAPEIT (V2)	IMPUTE (V2)	0,99879
rs4252120	PLG	6	T/C	100	0,2767	Illumina Omni Express	Illuminus	N/A	N/A	N/A	N/A
rs264	LPL	8	G/A	N/A	N/A	N/A	N/A	652.329	SHAPEIT (V2)	IMPUTE (V2)	0,96875
rs9319428	FLT1	13	A/G	N/A	N/A	N/A	N/A	652.329	SHAPEIT (V2)	IMPUTE (V2)	0,99868
rs17514846	FURIN-FES	15	A/C	100	0,3101	Illumina Omni Express	Illuminus	N/A	N/A	N/A	N/A
rs2954029	TRIB1	8	A/T	N/A	N/A	N/A	N/A	652.329	SHAPEIT (V2)	IMPUTE (V2)	0,98480
rs6544713	ABCG5-ABCG8	2	T/C	N/A	N/A	N/A	N/A	652.329	SHAPEIT (V2)	IMPUTE (V2)	0,99591
rs1878406	EDNRA	4	T/C	N/A	N/A	N/A	N/A	652.329	SHAPEIT (V2)	IMPUTE (V2)	0,99456
rs2023938	HDAC9	7	G/A	100	0,4119	Illumina Omni Express	Illuminus	N/A	N/A	N/A	N/A
rs17087335	REST-NOA1	4	T/G	100	0,1419	Illumina Omni Express	Illuminus	N/A	N/A	N/A	N/A
rs3918226	NOS3	7	T/C	100	0,6884	Illumina Omni Express	Illuminus	N/A	N/A	N/A	N/A
rs10840293	SWAP70	11	A/G	N/A	N/A	N/A	N/A	652.329	SHAPEIT (V2)	IMPUTE (V2)	0,94733
rs56062135	SMAD3	15	C/T	N/A	N/A	N/A	N/A	652.329	SHAPEIT (V2)	IMPUTE (V2)	0,99414
rs8042271	MFGE8-ABHD2	15	G/A	N/A	N/A	N/A	N/A	652.329	SHAPEIT (V2)	IMPUTE (V2)	0,97432
rs7212798	BCAS3	17	C/T	N/A	N/A	N/A	N/A	652.329	SHAPEIT (V2)	IMPUTE (V2)	0,97600
rs663129	PMAIP1-MC4R	18	A/G	N/A	N/A	N/A	N/A	652.329	SHAPEIT (V2)	IMPUTE (V2)	0,99794
rs180803	POM121L9P-ADORA2A	22	G/T	N/A	N/A	N/A	N/A	652.329	SHAPEIT (V2)	IMPUTE (V2)	0,68771

Supplementary Table 2. Association summary statistics for known susceptibility CAD loci

CARDIoGRAMplusC4D association summary statistics							THISEAS association summary statistics			
SNP	**Candidate Gene**	**Chr**	**Effect allele/ Non-effect allele**	**EAF***	**OR ***	***p*-value***	**EAF**	**OR**	**95% CI**	***p*-value**
rs602633	SORT1	1	C/A	0.77	1.13	2.19×10^{-18}	0.80	1.27	1,02-1,59	1,140
rs11206510	PCSK9	1	T/C	0.84	1.04	5.09×10^{-3}	0.80	0.87	0.59-1.10	0.129
rs6725887	WDR12	2	C/T	0.11	1.10	5.29×10^{-8}	0.15	1.35	1.06-1.73	0.198
rs9818870	MRAS	3	T/C	0.14	1.05	1.83×10^{-3}	0.14	1.20	0.93-1.55	0.548
rs12190287	TCF21	6	C/G	0.59	1.04	6.48×10^{-4}	0.64	0.95	0.74-1.12	0.819
rs2048327	SLC22A3-LPAL2-LPA	6	G/A	0.35	1.05	1.09×10^{-5}	0.27	0.92	0.75-1.13	0.558
rs11556924	ZC3HC1	7	C/T	0.65	1.08	1.45×10^{-9}	0.67	0.95	0.73--1.13	0.702
rs1333049	CDKN2BAS1	9	C/G	0.47	1.21	1.08×10^{-34}	0.44	1.16	0.97-1.39	0.581
rs579459	ABO	9	C/T	0.21	1.04	2.13×10^{-2}	0.21	1.04	0.83-1.29	0.537
rs12413409	CYP17A1-CNNM2-NT5C2	10	G/A	0.89	1.08	4.12×10^{-3}	0.89	1.17	0.89-1.38	0.406
rs2505083	KIAA1462	10	C/T	0.42	1.06	2.82×10^{-7}	0.46	0.93	0.78-1.11	0.796
rs974819	PDGFD	11	A/G	0.29	1.08	2.03×10^{-9}	0.37	1.00	0.80-1.17	0.678
rs3184504	SH2B3	12	T/C	0.40	1.07	6.13×10^{-3}	0.48	0.99	0.80-1.16	0.882
rs9515203	COL4A1-COL4A2	13	T/C	0.74	1.08	1.13×10^{-8}	0.80	1.13	0.92-1.31	0.084
rs2895811	HHIPL1	14	C/T	0.43	1.04	1.18×10^{-4}	0.39	1.02	0.85-1.22	0.952
rs12936587	RAI1-PEMT-RASD1	17	G/A	0.59	1.04	2.06×10^{-4}	0.61	1.01	0.82-1.17	0.745
rs1122608	LDLR	19	G/T	0.76	1.06	3.72×10^{-6}	0.71	0.96	0.74-1.14	0.755
rs9982601	Gene desert (KCNE2)	21	T/C	0.13	1.10	8.69×10^{-9}	0.13	1.07	0.83-1.39	0.508
rs17114036	PPAP2B	1	A/G	0.91	1.09	2.68×10^{-5}	0.91	1.02	0.67-1.27	0.469
rs12205331	ANKS1A	6	C/T	0.81	1.01	4.36×10^{-1}	0.90	1.12	0.83-1.33	0.408
rs9369640	PHACTR1	6	A/C	0.65	1.09	1.11×10^{-12}	0.58	1.07	0.89-1.28	0.230
rs2047009	CXCL12	10	C/A	0.48	1.05	9.66×10^{-6}	0.47	1.11	0.93-1.33	0.291
rs11203042	LIPA	10	T/C	0.44	1.03	9.86×10^{-3}	0.43	0.99	0.80-1.16	0.983
rs15563	UBE2Z	17	C/T	0.52	1.01	2.44×10^{-1}	0.53	1.10	0.92-1.31	0.377
rs2281727	SMG6	17	C/T	0.36	1.04	8.46×10^{-4}	0.42	0.99	0.83-1.19	0.898
rs2075650	ApoE-ApoC1	19	G/A	0.14	1.11	5.86×10^{-11}	0.11	1.03	0.77-1.37	0.852
rs17464857	MIA3	1	T/G	0.87	1.02	1.55×10^{-1}	0.84	0.93	0.64-1.16	0,699
rs12539895	7q22	7	A/C	0.19	1.02	4.00×10^{-2}	0.18	0.80	0.64-1.01	0.121

Supplementary Table 2. Association summary statistics for known susceptibility CAD loci **(continued)**

				CARDIoGRAMplusC4D association summary statistics			**THISEAS association summary statistics**			
SNP	**Candidate Gene**	**Chr**	**Effect allele/ Non-effect allele**	**EAF***	**OR ***	***p*-value***	**EAF**	**OR**	**95% CI**	***p*-value**
rs9326246	ZNF259-APOA5-APOA1	11	C/G	0.10	1.04	2.90×10^{-2}	0.10	1.03	0.68-1.28	0.843
rs7173743	ADAMTS7	15	T/C	0.58	1.06	2.46×10^{-7}	0.50	1.10	0.93-1.24	0.419
rs4845625	IL6R	1	T/C	0.47	1.04	3.46×10^{-8}	0.47	0.96	0.76-1.13	0,374
rs515135	APOB	2	G/A	0.83	1.08	2.17×10^{-8}	0.81	1.06	0.85-1.33	0.347
rs2252641	ZEB2-AC074093.1	2	G/A	0.46	1.04	1.27×10^{-4}	0.46	0.99	0.83-1.19	0.896
rs1561198	VAMP5-VAMP8-GGCX	2	A/G	0.45	1.05	2.57×10^{-6}	0.45	1.05	0.88-1.25	0.315
rs7692387	GUCY1A3	4	G/A	0.81	1.06	1.85×10^{-5}	0.80	1.09	0,86-1.27	0.368
rs273909	SLC22A4-SLC22A5	5	C/T	0.14	1.09	2.00×10^{-7}	0.08	0.93	0.68-1.28	0.986
rs10947789	KCNK5	6	T/C	0.76	1.06	1.22×10^{-5}	0.79	0.96	0.69-1.15	0.783
rs4252120	PLG	6	T/C	0.73	1.06	1.82×10^{-5}	0.68	1.08	0.89-1.24	0.447
rs264	LPL	8	G/A	0.86	1.05	7.30×10^{-4}	0.84	1.15	0.93-1.33	0.248
rs9319428	FLT1	13	A/G	0.32	1.05	5.70×10^{-6}	0.29	1.05	0.86-1.28	0.961
rs17514846	FURIN-FES	15	A/C	0.44	1.05	7.35×10^{-7}	0.45	0.98	0.82-1.17	0.908
rs2954029	TRIB1	8	A/T	0.55	1.04	7.75×10^{-5}	0.59	1.07	0.88-1.22	0.280
rs6544713	ABCG5-ABCG8	2	T/C	0.30	1.06	1.57×10^{-7}	0.36	1.12	0.94-1,27	0.302
rs1878406	EDNRA	4	T/C	0.15	1.06	3.54×10^{-3}	0.17	1.10	0.87-1.39	0.310
rs2023938	HDAC9	7	G/A	0.10	1.07	5.25×10^{-5}	0.13	0.83	0.64-1.09	0.954
rs17087335	REST-NOA1	4	T/G	0.21	1.06	4.60×10^{-8}	0.18	0.99	0.79-1.25	0.942
rs3918226	**NOS3**	**7**	**T/C**	**0.06**	**1.14**	**1.70×10^{-9}**	**0.08**	**1.32**	**0.96-1.80**	**0.007**
rs10840293	SWAP70	11	A/G	0.55	1.06	1.30×10^{-8}	0.61	1.06	0.88-1.27	0.406
rs56062135	SMAD3	15	C/T	0.79	1.07	4.50×10^{-9}	0.79	1.08	0.85-1.25	0.938
rs8042271	**MFGE8-ABHD2**	**15**	**G/A**	**0.90**	**1.10**	**3.70×10^{-8}**	**0.93**	**1.48**	**1.25-1.64**	**0.0009**
rs7212798	BCAS3	17	C/T	0.15	1.08	1.90×10^{-8}	0.15	1.14	0.89-1.45	0.860
rs663129	PMAIP1-MC4R	18	A/G	0.26	1.06	3.20×10^{-8}	0.25	1.13	0.93-1.39	0.671
rs180803	POM121L9P-ADORA2A	22	G/T	0.97	1.20	1.60×10^{-10}	0.99	1.50	0.65-1.81	0.016

SNP= single nucleotide polymorphism, Chr=chromosome, EAF=effect allele frequency, OR=odds ratio, 95% CI= 95% Confidence Interval, HWE=Hardy Weinberg Equilibrium

Results were obtained using adjusted logistic regression for age and sex. Allelic test p-value, OR and 95% CI are shown for each single SNP.

*EAF, OR and allelic test p-value as published by previous literature. Bold high-lighted locus indicated nominal evidence for association with CAD risk.

Supplementary Table 3. Student's t test association of the unwGRS-53

	CAD patients N=422		Controls N=576		p-value
	Mean	SD	Mean	SD	
unwGRS-53	52.3 (±4,4)	4,7	51,5	4,3	0,006

CAD=coronary artery disease, unwGRS-53=unweighted genetic risk score

Data are expressed as mean and standard deviation (SD)

Supplementary Table 4. Results from the logistic regression models for the evaluation of wGRS-53 on the risk of developing CAD

	OR	95% CI	*p*-value
Model 1*	1,04	1,01-1,07	0,006
Model 2†	1,05	1,02-1,08	0,005
Model 3‡	1,06	1,02-1,10	0,001

wGRS-53= unweighted genetic risk score, CAD=coronary artery disease, OR=odds ratio, CI=confidence interval

* **Model 1** unadjusted

† **Model 2** adjusted for age, sex

‡ **Model 3** adjusted for age, sex and body mass index

Supplementary Table 5. Logistic regression for CAD incidence by risk factors in the THISEAS study

	N	ORs	95% CI	*p*- value
Hypertension (yes vs. no)	964			
Model 1[*]		6.51	4.35-9.72	0.000
Model 2[†]		6.39	4.27-9.57	0.000
Type 2 diabetes mellitus (yes vs. no)	983			
Model 1		2.79	1.95-3.99	0.000
Model 2		2.75	1.92-3.93	0.000
Smoking status I (current/former vs. never smokers)	851			
Model 1		3.30	2.25-4.85	0.000
Model 2		3.31	2.25-4.88	0.000
Smoking status II (current vs. never/former smokers)	851			
Model 1		2.07	1.47-2.92	0.000
Model 2		2.07	1.46-2.92	0.000
Physical activity adoption (no vs. yes)	650			
Model 1		1.95	1.17-3.26	0.010
Model 2		1.90	1.14-3.18	0.014
Energy intake (above or equal to vs below median)	538			
Model 1		1.94	1.29-2.91	0.001
Model 2		1.92	1.28-2.89	0.002
Waist-to-hip ratio (above or equal to vs below median)	693			
Model 1		2.18	1.48-3.19	0.000
Model 2		2.20	1.48-3.24	0.000

CAD=coronary artery disease, THISEAS= The Hellenic study of Interactions between Single nucleotide polymorphisms and Eating in Atherosclerosis Susceptibility

*****adjusted for age, sex and body mass index

†adjusted for age, sex, body mass index and unweighted genetic risk score

Chapter 7

Appendix

7 APPENDIX

REVIEW PAPER

This paper review is entitled "Gene-Diet Interactions in Cardiovascular disease. The purpose of the manuscript was to review current evidence regarding the effects of genetic variation on responsiveness to diet in CAD.

Key points

- Papers published from 2010 to 2012 were used for the current review.
- The evidence regarding gene-diet interaction in CAD was limited.
- Investigated gene-diet interactions could not give plausible explanations regarding the biology and molecular pathways of the disease.
- Diet is one of the most important environmental factor that can induce epigenome perturbations.

Curr Nutr Rep
DOI 10.1007/s13668-012-0020-4

GENETICS (GVZ DEDOUSSIS, SECTION EDITOR)

Gene–Diet Interactions in Cardiovascular Disease

Maria E. Dimitriou · George V. Z. Dedoussis

Abstract Cardiovascular disease (CVD), the leading cause of mortality worldwide, results from a complex interplay between genetic and environmental factors. Genetic studies identified genetic variants providing insights into the pathogenesis and treatment of the disease. However, the mechanisms linking the genotypic and phenotypic expression remain to be elucidated. Gene–diet interaction studies attempt to elucidate how a modifiable factor interacts with the genetic background. The knowledge gained thus far confers to small increments of CVD risk and cannot explain the molecular mechanisms of the disease. Epigenetic studies attempt to elucidate the molecular pathways affected by an environmental stimulus, such as dietary exposure. The epigenomic changes and their link to gene–diet interactions remain a challenging area for research. Understanding the complex interplay among the epigenome, genome, and dietary exposure should lead to accurate prediction, prevention, or treatment of the disease.

Keywords Gene–diet interaction · Polymorphisms · Diet · Genes · Genetics · Cardiovascular disease · Epigenetics · Methylation

Introduction

Cardiovascular disease (CVD) is a class of disorders of the heart and blood vessels and includes coronary artery disease (CAD), cerebrovascular disease (stroke), peripheral artery disease, rheumatic heart disease, congenital heart disease, and heart failure. CVD is the leading cause of mortality worldwide, with more than 80 % of deaths taking place in low- and middle-income countries. In a recent report, the World Health Organization (WHO, 2011) estimated that more than 17 million people died from CVD in the year 2008, thus accounting for 30 % of all global deaths. By the year 2030, it is estimated that CVD will account for 23.6 million deaths worldwide [1••].

The underlying cause of these manifestations is atherosclerosis. The atherosclerotic process accounts for many types of CVD, predominantly CAD and stroke, and is advanced by the time heart problems are detected. Genetic predisposition, behavioral, metabolic, and other risk factors are responsible for the onset of atherosclerosis, which finally leads to CVD (Table 1) [1••].

Among the established risk factors of CVD, there is strong scientific evidence that genetic background in response to dietary exposure plays a key role in the onset of the disease [2, 3]. Thus far, implemented guidelines and strategies in primary and secondary prevention of CVD mainly focus on a healthy lifestyle and endeavor to behavioral, social, and environmental changes. Nonetheless, further understanding of the pathogenesis and etiology of the disease requires research into the complex interaction between genetic and environmental factors.

Nutrigenetic and nutrigenomic research intends to explore the mechanisms of the interplay between diet and genetic variants in diet-related diseases (eg, CVD). This will lead to even more specialized guidelines according to individual genotypic architecture. The aim is to reduce the prevalence of chronic diseases, to prevent them, or even to cure them. For the purpose of the current review, outcomes regarding the effects of genetic variation on responsiveness to diet in CVD—specifically CAD—are examined.

M. E. Dimitriou (✉) · G. V. Z. Dedoussis
Department of Nutrition and Dietetics, Harokopio University,
70 El. Venizelou Street,
17671 Kallithea, Athens, Greece
e-mail: mdimi@hua.gr

G. V. Z. Dedoussis
e-mail: dedousi@hua.gr

Published online: 19 June 2012

Table 1 Risk factors associated with cardiovascular disease according to the World Health Organization (2011)

Behavioral risk factors	Unhealthy diet
	Physical inactivity
	Harmful use of alcohol
	Tobacco use
Metabolic risk factors	Hypertension
	Diabetes
	Elevated blood lipids (ie, hypercholesterolemia)
	Increased body weight (overweight, obesity)
Other risk factors	Genetic predisposition
	Gender
	Population ageing
	Elevated homocysteine levels
	Poverty and low educational status
	Psychological status (eg, stress, depression)

Genetic Determinants for Cardiovascular Disease

Individuals differ in their susceptibility to CVD, and genetic variants play a key role in this. Indicatively, in a twin study, it has been shown that there is a 57 % and 38 % heritability for CAD in men and women, respectively [4]. In younger individuals, it has been demonstrated that genetic variants may contribute to a 20 % to 60 % increase in CAD risk [5]. Therefore, research to identify these variants is extended through the past two decades [6, 7].

The candidate gene approach is based on an *a priori* hypothesis of the plausible involvement of the selected gene in the pathogenesis and process of the disease under investigation. Many studies examined single nucleotide polymorphisms (SNPs) in single candidate genes, while others focused on more SNPs in the same gene. Historically, the success of the first candidate gene study in identifying a genetic variant in the susceptibility of CVD was published in *Nature* in 1992. Specifically, the study explored a variant found in the gene encoding angiotensin-converting enzyme (ACE) and showed that homozygotes for a deletion in the *ACE* gene were at higher myocardial infarction (MI) risk [7]. Another example reflecting the success of this approach is the identification of variants in the apolipoprotein E (*APOE*) gene, which has a major role in cholesterol metabolism [8]. On the other hand, animal and human studies have provided strong evidence regarding the pro- and antiatherogenic roles of the cholesterol ester transfer protein (*CETP*) gene by playing a key role in reverse cholesterol transport [9]. The breakthrough of the candidate gene studies was the knowledge obtained in the pathogenesis and the treatment of CVD through the identification of variants in genes associated with rare and monogenic forms of CVD—that is, mendelian CVD. However, candidate gene studies appeared unsuccessful in identifying the genetic background of polygenic CVD, including CAD, mainly due to limited tested polymorphisms and small sample sizes. Furthermore, a significant number of SNPs involved in disease pathways were left aside.

Genome-wide association studies (GWAS) enabled the scanning of the entire genome in order to seek out associations between hundreds of thousands of SNPs and diseases. The first GWAS on CVD identified some SNPs on 9p21.3 loci associated with CAD [10–13]. Reviewing the data up to 2011, 35 variants associated with CAD are reported [14••]. A meta-analysis performed by the CARDIoGRAM Consortium found that rs1333049 on the 9p21 region confers a 29 % increase in MI risk per allele [15••]. Schunkert et al. [16••] identified 13 new loci associated with CAD. Specifically, the risk alleles for the new loci were associated with an increase in CAD risk ranging from 6 % to 17 % per allele. The C4D Consortium identified five novel loci, including the lipase (*LIPA*) gene [17]. The CARDIoGRAMplusC4D Consortium raised the number of variants associated with CAD to 46 in the Caucasian population. In total, all the genetic variants identified thus far explain approximately 10 % of the cardiovascular risk in humans (unpublished data). Although the genome-wide approach led to the discovery of many SNPs associated with CVD, mechanisms linking the genotypic and phenotypic expression remain to be elucidated. More insights will be gained by the investigation of gene–environment interactions that will help us further understand the molecular pathways of the disease and explain the interindividual variability. Diet is one of the environmental factors that plays a major role in CVD.

Dietary Risk Factors for Cardiovascular Disease

Many researchers have tried to explore the relationship between dietary factors and CVD for almost half a century. The effects of many nutrients, foods, and dietary patterns on CVD have been evaluated by numerous studies. The importance of diet as a risk factor for CVD was first demonstrated by the Seven Countries Study [18–21]. Specifically, the Mediterranean-type diet was inversely associated with the incidence of CAD in Southern Europe when compared with the United States and Northern Europe, after adjusting for confounding factors.

The Seven Countries Study paved the way for other large-scale cohorts to follow, which examined the effects of dietary food groups or dietary patterns on CVD risk. Only a few of the major findings of some studies are highlighted in the next lines, as this is not the purpose of the current review. The CARDIO2000 study, a multicenter case-control study, showed that exclusive olive oil consumption was associated with a 0.55 times lower likelihood of having

acute coronary syndrome (ACS) among hypercholesterolemic subjects [22]. Another case-control study, the INTERHEART study, demonstrated that a healthy dietary pattern characterized by increased consumption of fruits and vegetables has a protective role against MI [23]. The protective effect of vegetable intake in CAD has been also demonstrated by the Physicians' Health Study in male subjects. Specifically, in this study, men who consumed at least 2.5 servings/d of vegetables had a relative risk (RR) of 0.77 (95 % CI, 0.60–0.98) for CAD compared with men with vegetable consumption of less than 1 serving/d [24]. The WHO (2009) attributes about 11 % of ischemic heart disease deaths to insufficient intake of fruits and vegetables [25].

Furthermore, prospective studies have examined the role of dietary patterns in relationship to CAD risk. Specifically, the Multi-Ethnic Study (MESA) demonstrated a negative association between the prevalence of CVD and the consumption of fruits, vegetables, whole grains, and low-fat content dairy after a 4-year follow-up. On the contrary, high consumption of red meat, fried food, and sweets increased the prevalence of CAD [26]. Along the same lines, the ATTICA study demonstrated a protective role of fruits, vegetables, and exclusive olive oil use against CAD. On the other hand, high consumption of red meat, margarine, sweets, cheese, and alcohol was positively associated with the prevalence of CAD [27].

Regarding micronutrients, the consumption of foods rich in antioxidant compounds has been associated with decreased MI markers by exerting protective effects against low-density lipoprotein cholesterol (LDL-C) oxidation and endothelial vascular dysfunction. However, the benefit of their clinical use requires further investigation [28].

Gene–Diet Interactions in Cardiovascular Disease Risk

Gene–diet interaction studies are emerging and attempting to elucidate how a modifiable factor—that is, dietary exposure—interacts with genetic background. The new insights from gene–diet interaction studies promise to determine the individual risk of the disease, which will further lead to the implementation of personalized diets.

To date, the vast majority of studies that examined gene–diet interactions in CVD are referring mostly to cardiovascular risk factor outcomes. For the purpose of the current review, we searched PubMed, limited to studies on human subjects, with an emphasis on CAD and MI. A combination of keywords, namely genes, SNPs, polymorphisms, diet, dietary patterns, food groups, fat, cardiovascular disease, and coronary artery disease, was used. To provide updated knowledge, papers published during the past 2 years were used. During this period, the amount of information regarding gene–diet interactions in CVD in humans is rather limited. A summary of these studies is provided in Table 2.

A recent study on 9p21 loci, a chromosome region, extensively explored identified interactions between dietary patterns and rs2383206 on CVD risk. The results of the study showed that homozygous individuals (from the INTERHEART study) for the risk allele with a less prudent diet have a 1.6- to 2.0-fold increased risk of MI [29]. A prudent diet was a dietary pattern that consisted of raw vegetables, fruits, leafy green vegetables, nuts, desserts, and dairy products (food items that had factor loadings more than 0.25) [30]. A proxy SNP for rs2383206—that is, rs4977574—was also investigated for genetic association with CVD in a different sample (from the FINRISK study). It was found that the risk allele had an effect on the incidence of CVD among subjects with low and medium consumption of vegetables, fruits, and berries [29]. This result is similar to the one from the INTERHEART analysis. Furthermore, decreased CAD incidence has also been demonstrated in individuals with the *BHMT* minor allele in a case-control study [31•].

Regarding gene–diet interactions in clinical manifestations of the atherosclerotic process, there is only one study in humans examining the effect of the interaction between the 5-lipoxygenase promoter (*ALOX5*) genotype and dietary intake on carotid artery intima-media thickness. Although this study was not published within the last two years, we are referring to it as hallmark. The results of the study demonstrated that increased dietary intake of n-6 fatty acids, namely arachidonic and linoleic acid, was associated with increased intima-media thickness only among individuals with two variant alleles of the *ALOX5* gene. In contrast, dietary intake of n-3 fatty acids was inversely associated with intima-media thickness only among individuals carrying two variant alleles. These results suggest for the first time that interactions between genes in inflammatory pathways and diet may lead to the onset of the atherosclerotic process in humans. However, these results need to be replicated [32].

Significant genotype–diet interactions have been also reported for the *APOE* gene. Frequent intake of olive oil and polyunsaturated fatty acids (PUFAs) was associated with an increase in triglyceride and LDL-C levels in E4 allele carriers. On the other hand, when analysis was performed only in male subjects, it was demonstrated that high olive oil intake was associated with low LDL-C levels in E2 carriers. Moreover, high intake of PUFAs was associated with decreased triglyceride levels in E2 male and female carriers [33]. Regarding saturated fatty acids (SFAs), individuals carrying the E4 allele had a greater CAD risk for SFA intake of more than 10 % of energy compared with individuals carrying the E2 allele [34]. In response to a high-fat diet, individuals carrying an allele for rs4246444 on the

Table 2 Gene–diet interactions studies on CVD risk factors

Chromosome or gene (SNP)	Dietary factor	Population	Sample size, *n*	Major findings	Reference
9p21 (rs10757274, rs2383206, rs10757278, rs1333049)	3 dietary patterns (oriental, Western, prudent)	INTERHEART study	8,114	In response to a low prudent diet score, homozygous individuals for the risk allele of rs2383206 have a 1.6- to 2.0-fold increase in the risk for MI	[29]
9p21 (rs 4977574, proxy for rs2383206)	3 diet groups (low, medium, high consumption of vegetables, fruits, and berries)	FINRISK study	19,129 (including 1,014 cases of CVD)	Proxy rs4977574 showed an effect of the risk allele on incident CVD in the "low" and "medium" consumption group No effect was demonstrated in the high consumption group	[29]
BHMT (742 G→A SNP)	Micronutrient intake	Indian	504 (252 cases with CAD, 252 controls)	In the presence of the BHMT minor allele and decreased folate, consumption may increase the risk of CAD	[31•]
ALOX5 promoter	Diet	Los Angeles Atherosclerosis study	470 (healthy subjects)	Increased dietary intake of arachidonic and linoleic acid was associated with increased intima-media thickness among individuals with 2 variant alleles of the ALOX5 gene On the contrary, dietary intake of marine n-3 fatty acids was inversely associated intima-media thickness. Gene–diet interactions were not observed for MUFA or SFA	[32]
APOE (E2, E3, E4); APOAI (g-75a); APOAV (S19W); SRBI (Gly2Ser, c1050t); ABCA1(R219K); HL (g-250a); CETP (TaqIB1/2)	Dietary habits (9 food groups)	Brazilian	567	Frequent intake of olive oil was associated with increased triglyceride levels in individuals carrying the E4 allele Frequent PUFA intake was associated with increased HDL-C levels in individuals who are not carriers of the E4 allele Frequent intake of foods rich in PUFA was associated with increased LDL levels in individuals carrying the E4 allele High intake of PUFA was associated with decrease in TG levels in individuals with the E2 allele Frequent intake of olive oil was associated with low LDL-C levels in male individuals carrying the E2 allele	[33]
APOE (E2, E3, E4)	SFA, alcohol intake	The Spanish EPIC cohort	1657 (1,123 controls, 534 cases with CAD)	A greater CAD risk was determined when SFA intake was greater than 10 % of energy in individuals carrying the E4 allele when compared with E2 allele carriers Alcohol consumption increases HDL-C levels in E2 carriers; among drinkers, E4 carriers have the lower HDL-C levels	[34]
FASN (5 SNPs)	Fat	Quebec Family study	592	Individuals carrying an allele for rs4246444 had smaller LDL-PPD only when consuming a high-fat diet	[35•]
CETP (rs708272, rs183130)	Alcohol consumption, fat intake(total fat, SFA, MUFA), adherence to the Mediterranean diet	The PREDIMED study	4,210 (1,840 women and 2,370 men)	Alcohol, dietary fat, and adherence to the Mediterranean diet did not interact with the CETP polymorphism in determining HDL-C levels	[36]

Table 2 (continued)

Chromosome or gene (SNP)	Dietary factor	Population	Sample size, *n*	Major findings	Reference
LPL (Ser447Stop, Hind III)	Dietary intervention (control diet for 7 d, followed by a HC/LF diet for 6 d)	Chinese	56 (students)	In response to an HC/LF diet, male H- carriers of the Hind III polymorphism had increased HDL-C and APO AI levels; females with genotype S447S had increased TG levels after the dietary intervention	[37•]
PRAPα (Leu162Val); PRAPγ (Pro12Ala)	Dietary fatty acids	The RISCK study	466 (at increased cardiometabolic risk)	In habitual diet, TC was high in individuals with Pro12Ala genotype when dietary PUFA:SFA ratio was low; a reduction of 10 % in TC was observed as PUFA:SFA ratio was increased. PUFA:SFA ratio x PPAR-γ Pro12Ala genotype also influenced LDL-C and TG. These results referred to white subjects.	[41]

CAD coronary artery disease; *CVD* cardiovascular disease; *HC* high carbohydrate; *HDL-C* high-density lipoprotein cholesterol; *LDL-C* low-density lipoprotein cholesterol; *LF* low fat; *MI* myocardial infarction; *MUFA* monounsaturated fatty acids; *PPD* peak particle diameter; *PUFA* polyunsaturated fatty acids; *SFA* saturated fatty acids; *SNP* single nucleotide polymorphism; *TC* total cholesterol; *TG* triglycerides

FASN *gene* had LDL-C with smaller peak particle diameter (LDL-PPD) compared with non-carriers [35•].

Regarding the *CETP* gene, a recent study demonstrated no interaction between *CETP* polymorphisms and dietary factors, namely alcohol, SFA, and adherence to the Mediterranean diet, in determining high-density lipoprotein cholesterol (HDL-C) levels [36]. A study on young Chinese subjects examined the interaction between lipoprotein lipase (*LPL*) gene variants and high-carbohydrate (HC)/low-fat diet [37•]. The polymorphisms examined modified the lipid traits of the young population in various ways (Table 1). What makes the study noteworthy is that the study population was made up of young individuals, in contrast to middle-aged or older individuals. Furthermore, the Chinese population is documented to have a diet with more carbohydrates and less fat, better lipid profiles, and lower risk of CAD incidence [37•, 38, 39]. Most studies tend to examine the effect of an HC diet on lipids, especially triglycerides, among middle-aged or older subjects.

In mammals, peroxisomal proliferator-activated receptor-α (PRAPα) and peroxisomal proliferator-activated receptor-γ (PRAPγ) are transcription factor that regulate the cardiac energy production from fatty streaks. PRAPα promotes fatty acid β-oxidation and PRAPγ promotes triglycerol synthesis [40]. A study on individuals at high cardiometabolic risk investigated the effect of interaction of *PRAPα* and *PRAPγ* genes with dietary fatty acids on lipid traits. It was shown that in habitual diet, the PUFA:SFA ratio influenced total cholesterol, LDL-C, and triglyceride levels in white individuals carrying the 12Ala allele. As the ratio increased, plasma lipids decreased [41].

GWAS have identified many genetic variants associated with CVD that provide insights into genetic background. However, the variants identified confer small increments of CVD risk and cannot explain interindividual susceptibility. Thus far, investigated gene–diet interactions cannot give plausible explanations regarding the molecular mechanisms of the disease, and we are far beyond CVD occurrence prediction or personalized diet implementation. Scientific evidence in identical twins demonstrated that the different susceptibility to polygenic disease, such as CVD, is attributable to alterations in the epigenome, followed by different environmental exposure [42, 43]. The question that arises is how environmental exposures, such as diet, could result in different individual susceptibility.

Epigenetics

Epigenetics are heritable changes in the genome that may lead to altered gene expression, but with no alterations in the nucleotide sequence. Histone modification, DNA methylation, and noncoding RNA functions are examples of the epigenetic process. DNA methylation is mostly explored in epigenetic studies. It refers to a biochemical procedure during which a methyl group is added to CpG dinucleotides, and it normally confers silence of gene expression. Epigenome is established during the development of the embryo and changes to regulation mechanisms of gene expression by an environmental stimulus conferring phenotype variability and different susceptibility to chronic diseases. Diet is one of the most important environmental factors that can induce epigenome perturbations. Nutrients such as folic acid, B vitamins, methionine, and choline are methyl-donating components. Figure 1 shows a model presenting the genetic variants, dietary exposure, and epigenome interactions.

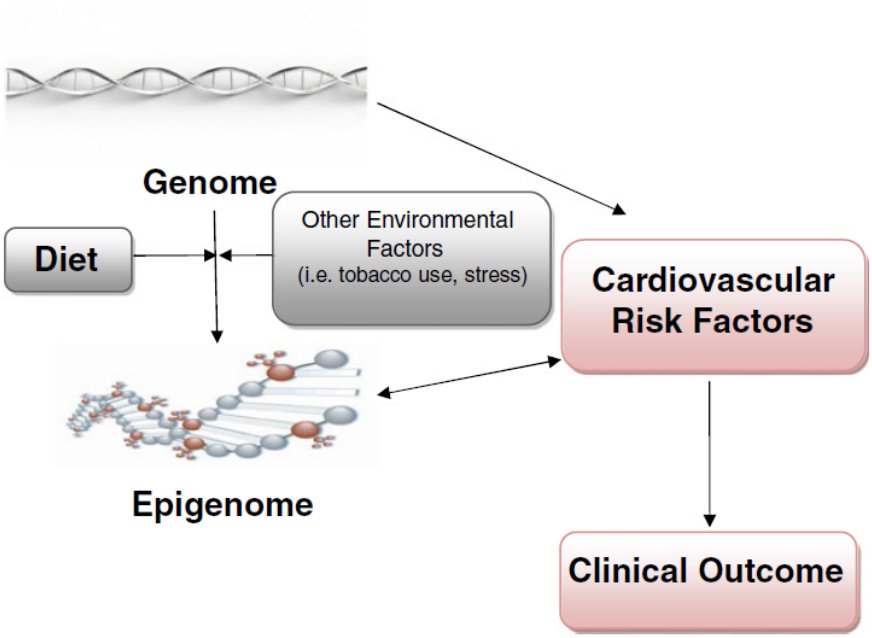


Fig 1 Model representing the interactions of genetic variants, dietary exposure, and epigenome modifications in cardiovascular disease manifestations

Influence of Diet on Epigenome in Cardiovascular Disease

Studies on animal models and humans showed that exposure to imbalanced nutrient supply in fetal life, deficient growth in gestational periods, or during infancy is associated with higher CVD mortality [43, 44•]. Nutrient excess or deficiency results in epigenetic alterations that are carried throughout the life span, are inherited, and may play a major role in phenotype diversity. The Dutch famine (1944–1945) demonstrated strong evidence that maternal nutrient deficiency was associated with epigenetic alterations and increased risk of chronic diseases, including CVD, later in the life of the fetus [45, 46].

A recent study in mice revealed that female offsprings ($APOE^{+/-}$) were more susceptible to neointima formation when exposed to maternal hypercholesterolemia ($APOE^{-/-}$) compared with identical offsprings from normocholesterolemic (wild-type) mothers. Furthermore, diet-induced hypercholesterolemia during the postnatal period further enhanced the susceptibility to atherosclerosis in offspring from $APOE^{-/-}$ mothers compared with those from wild-type mothers. The results of the study demonstrate that fetal programming and dietary exposure influence histone methylation and affect the epigenome [47]. A previous study has similarly shown that $APOE^{-/-}$ mice with atherosusceptibility present DNA methylation changes that promote atherosclerosis [48].

Other studies in mice have investigated whether diet can have epigenetic effects on *PRAPα* and *PRAPγ* regulation. *PRAPα*-altered regulation is associated with diabetic cardiomyopathy in animal models [49]. Specifically, in mice, *PRAPα* methylation was lower in offspring exposed to protein-restricted maternal diet compared with offspring exposed to protein-sufficient maternal diet. Changes in *PRAPα* methylation exert their effect in mRNA expression only later in life. The results of the study suggest that maternal deficiency leads to alterations in epigenomic background, resulting in health consequences later in life [50].

The exact molecular pathways that are affected by epigenetic changes need further investigation. More studies need to be employed regarding the influence of dietary exposure to the epigenome in the context of CVD. Therefore, understanding epigenetic disruptions and the relationship between diet and epigenome is undoubtedly one of the most important challenges in recent research.

Conclusions

CVD is a multifactorial, complex disease that involves interaction of genetic and environmental factors. The International Hap Map project enabled GWAS to identify novel genetic variants that contribute to the disease manifestations [51]. Although GWAS offer a powerful contribution to identifying common alleles in disease genes, the SNPs identified thus far provide small increments of disease risk and cannot provide predictive accuracy. In further elucidating the pathways resulting in the disease, the genetics of CVD is explored in relation to environmental factors. Gene–diet interaction studies remain limited to the literature and are still in progress, yet they promise to unravel some of the underlying mechanisms of the disease and determine the individual risk. Studies on epigenome modifications seem to play a key role in understanding diversities in individual

susceptibility. DNA methylation and histone modification in response to dietary exposure may help us elucidate the interplay between and diet and disease and understand how this interaction leads to increased CVD risk. Epigenetic disruptions are potentially reversible; thus, the development of epigenomic drugs and functional foods or supplements that can regulate the epigenome will be an upcoming challenge for the drug and food research industry.

Disclosure No potential conflicts of interest relevant to this article were reported.

References

Papers of particular interest, published recently, have been highlighted as:

- • Of importance
- •• Of major importance

1. •• World Health Organisation: Global Atlas on cardiovascular disease prevention and control. Available at: http://www.who.int/cardiovascular_diseases/publications/atlas_cvd/en/index/html. Accessed September 2011. *This report presents the most current trends in CVD.*
2. Lovegrove JA, Gitau R. Nutrigenetics and CVD: what does the future hold? Proc Nutr Soc. 2008;67:206–13.
3. Ordovas JM. Nutrigenetics, plasma lipids, and cardiovascular risk. J Am Diet Assoc. 2006;106:1074–81.
4. Zdravkovic S, Wienke A, Pedersen NL, et al. Heritability of death from coronary heart disease: a 36-year follow-up of 20 966 Swedish twins. J Intern Med. 2002;252:247–54.
5. Chaer RA, Billeh R, Massad MG. Genetics and gene manipulation therapy of premature coronary artery disease. Cardiology. 2004;101:122–30.
6. Lee DS, Pencina MJ, Benjamin EJ, et al. Association of parental heart failure with risk of heart failure in offspring. N Engl J Med. 2006;355:138–47.
7. Cambien F, Poirier O, Lecerf L, et al. Deletion polymorphism in the gene for angiotensin-converting enzyme is a potent risk factor for myocardial infarction. Nature. 1992;359:641–4.
8. Song Y, Stampfer MJ, Liu S. Meta-analysis: apolipoprotein E genotypes and risk for coronary heart disease. Ann Intern Med. 2004;141:137–47.
9. Vourvouhaki E, Dedoussis GV. Cholesterol ester transfer protein: a therapeutic target in atherosclerosis? Expert Opin Ther Targets. 2008;12:937–48.
10. McPherson R, Pertsemlidis A, Kavaslar N, et al. A common allele on chromosome 9 associated with coronary heart disease. Science. 2007;316:1488–91.
11. Helgadottir A, Thorleifsson G, Manolescu A, et al. A common variant on chromosome 9p21 affects the risk of myocardial infarction. Science. 2007;316:1491–3.
12. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 2007;447:661–78.
13. Samani NJ, Erdmann J, Hall AS, Hengstenberg C, et al. WTCCC and the Cardiogenics Consortium. Genomewide association analysis of coronary artery disease. N Engl J Med. 2007;357:443–53.
14. •• Peden JF, Farrall M. Thirty-five common variants for coronary artery disease: the fruits of much collaborative labour. Hum Mol Genet. 2011;20:198–205. *This article gives a summary of GWAS and large-scale genecentric genotyping studies that have reported genetic variants associated with CAD.*
15. •• Preuss M, König IR, Thompson JR, et al. CARDIoGRAM Consortium. Design of the Coronary ARtery DIsease Genome-Wide Replication And Meta-Analysis (CARDIoGRAM) Study: A Genome-wide association meta-analysis involving more than 22 000 cases and 60 000 controls. Circ Cardiovasc Genet. 2010;3:475–83. *This meta-analysis provides strong evidence regarding the role of 9p21 chromosome in CVD risk.*
16. •• Schunkert H, Konig IR, Kathiresan S, et al. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. Nat Genet. 2011;43:333–8. *This analysis is important mainly to the number of new loci identified in relation to CAD.*
17. Kathiresan S, Voight BF, Purcell S, et al. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. Nat Genet. 2009;41:334–41.
18. Keys A, Aravanis C, Blackburn HW, et al. Epidemiological studies related to coronary heart disease: characteristics of men aged 40-59 in seven countries. Acta Med Scand Suppl. 1966;460:1–392.
19. Keys A (Ed). Coronary heart disease in seven countries. Circulation. 1970, 41:1–200.
20. Keys A, Menotti A, Aravanis C, et al. The seven countries study: 2,289 deaths in 15 years. Prev Med. 1984;13:141–54.
21. Menotti A, Keys A, Aravanis C, et al. Seven Countries Study. First 20-year mortality data in 12 cohorts of six countries. Ann Med. 1989;21:175–9.
22. Kontogianni MD, Panagiotakos DB, Chrysohoou C, et al. The impact of olive oil consumption pattern on the risk of acute coronary syndromes: The CARDIO2000 case-control study. Clin Cardiol. 2007;30:125–9.
23. Iqbal R, Anand S, Ounpuu S. Islam Set al. INTERHEART Study Investigators. Dietary patterns and the risk of acute myocardial infarction in 52 countries: results of the INTERHEART study. Circulation. 2008;118:1929–37.
24. Liu S, Lee IM, Ajani U, et al. Intake of vegetables rich in carotenoids and risk of coronary heart disease in men: The Physicians' Health Study. Int J Epidemiol. 2001;30:130–5.
25. World Health Organisation: Global health risks 2009. Available at http://www.who.int/healthinfo/global_burden_disease/global_health_risks/en/index.html, Accessed December 2009.
26. Nettleton JA, Polak JF, Tracy R, et al. Dietary patterns and incident cardiovascular disease in the Multi-Ethnic Study of Atherosclerosis. Am J Clin Nutr. 2009;90:647–54.
27. Panagiotakos DB, Pitsavos C, Chrysohoou C, et al. ATTICA Study. Five-year incidence of cardiovascular disease and its predictors in Greece: the ATTICA study. Vasc Med. 2008;13:113–21.
28. Kaliora AC, Dedoussis GV. Natural antioxidant compounds in risk factors for CVD. Pharmacol Res. 2007;56:99–109.
29. Do R, Xie C, Zhang X, et al. The effect of chromosome 9p21 variants on cardiovascular disease may be modified by dietary intake: evidence from a case/control and a prospective study. PLoS Med. 2011;8:e1001106.
30. Iqbal R, Anand S, Ounpuu S, et al. Dietary patterns and the risk of acute myocardial infarction in 52 countries: results of the INTERHEART study. Circulation. 2008;118:1929–37.
31. • Singh PR, Lele SS, Mukherjee MS. Gene polymorphisms and low dietary intake of micronutrients in coronary artery disease. J Nutrigenet Nutrigenomics. 2011;4:203–9. *This is a recent gene–diet interaction study regarding micronutrients and SNP interactions.*
32. Dwyer JH, Allayee H, Dwyer KM, Fan J, Wu H, Mar R, Lusis AJ, Mehrabian M. Arachidonate 5-lipoxygenase promoter genotype,

dietary arachidonic acid, and atherosclerosis. N Engl J Med. 2004;350:29–37.
33. de Andrade FM, Bulhões AC, Maluf SW, et al. The influence of nutrigenetics on the lipid profile: interaction between genes and dietary habits. Biochem Genet. 2010;48:342–55.
34. Corella D, Portolés O, Arriola L, Chirlaque MD, et al. Saturated fat intake and alcohol consumption modulate the association between the APOE polymorphism and risk of future coronary heart disease: a nested case-control study in the Spanish EPIC cohort. J Nutr Biochem. 2011;22:487–94.
35. • Dolley G, Boisclair ME, Lamarche B, Després JP, Bouchard C, Pérusse L, Vohl MC. Interactions between dietary fat intake and FASN genetic variation influence LDL peak particle diameter. J Nutrigenet Nutrigenomics. 2011;4:137–45. *Little is known regarding the effect of diet on LDL-PPD. Most studies evaluate the influence on LDL values.*
36. Corella D, Carrasco P, Fitó M, et al. Gene-environment interactions of CETP gene variation in a high cardiovascular risk Mediterranean population. J Lipid Res. 2010;51:2798–807.
37. • Huang X, Gong R, Lin J, et al. Effects of lipoprotein lipase gene variations, a high-carbohydrate low-fat diet, and gender on serum lipid profiles in healthy Chinese Han youth. Biosci Trends. 2011;5 (5):198–204. *Not many studies have been conducted on young populations with a low CVD incidence.*
38. Saha N, Heng CK, Mozoomdar BP, et al. Racial variation of factor VII activity and antigen levels and their correlates in healthy Chinese and Indians at low and high risk for coronary artery disease. Atherosclerosis. 1995;117:33–42.
39. Lee MM, Wu-Williams A, Whittemore AS, et al. Comparison of dietary habits, physical activity and body size among Chinese in North America and China. Int J Epidemiol. 1994;23:984–90.
40. Bishop-Bailey D. Peroxisome proliferator-activated receptors in the cardiovascular system. Br J Pharmacol. 2000;129:823–34.
41. AlSaleh A, Sanders TA, O'Dell SD. Effect of interaction between PPARG, PPARA and ADIPOQ gene variants and dietary fatty acids on plasma lipid profile and adiponectin concentration in a large intervention study. Proc Nutr Soc. 2012;71(1):141–53.
42. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. Nature. 2009;461:747–53.
43. Barker DJ, Martyn CN. The maternal and fetal origins of cardiovascular disease. J Epidemiol Community Health. 1992;46:8–11.
44. • Jackson AA, Burdge GC, Lillicrop KA. Diet, nutrition and modulation of genomic expression in fetal origins of adult disease. J Nutrigenet Nutrigenomics. 2010;3:192–208. *This is an interesting review discussing how dietary exposure in early development affects later adult life through epigenome alterations.*
45. Painter RC, de Rooij SR, Bossuyt PM, et al. Blood pressure response to psychological stressors in adults after prenatal exposure to the Dutch famine. J Hypertens. 2006;24:1771–8.
46. Heijmans BT, Tobi EW, Stein AD, et al. Persistent epigenetic differences associated with prenatal exposure to famine in humans. Proc Natl Acad Sci USA. 2008;105:17046–9.
47. Alkemade FE, van Vliet P, Henneman P, et al. Prenatal exposure to apoE deficiency and postnatal hypercholesterolemia are associated with altered cell-specific lysine methyltransferase and histone methylation patterns in the vasculature. Am J Pathol. 2010;176:542–8.
48. Lund G, Andersson L, Lauria M, et al. DNA methylation polymorphisms precede any histological sign of atherosclerosis in mice lacking apolipoprotein E. J Biol Chem. 2004;279:29147–54.
49. Finck BN, Han X, Courtois M, et al. A critical role for PPARalpha-mediated lipotoxicity in the pathogenesis of diabetic cardiomyopathy: modulation by dietary fat content. Proc Natl Acad Sci U S A. 2003;100:1226–31.
50. Slater-Jefferies JL, Lillycrop KA, Townsend PA, et al. Feeding a protein-restricted diet during pregnancy induces altered epigenetic regulation of peroxisomal proliferator-activated receptor-α in the heart of the offspring. J Dev Orig Health Dis. 2011;2:250–5.
51. International HapMap Consortium. The International HapMap Project. Nature. 2003;18(426):789–96.