Oxidants and antioxidants as targets for cardiovascular disease prevention: evidence from observational and causal inference studies

Martens, L.G.

OXIDANTS AND ANTIOXIDANTS AS TARGETS FOR CARDIOVASCULAR DISEASE PREVENTION:

evidence from observational and causal inference studies

Leon G. Martens

OXIDANTS AND ANTIOXIDANTS AS TARGETS FOR CARDIOVASCULAR DISEASE PREVENTION:

evidence from observational and causal inference studies

Leon G. Martens

Oxidants and antioxidants as targets for cardiovascular disease prevention: evidence from observational and causal inference studies

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr.ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op woensdag 29 maart 2023
klokke 13.45 uur

door

Leon Gilbert Martens
geboren te Haarlem
in 1994

Cover design: Max Roelofsen
Layout: Ilse Modder (www.ilsemodder.nl)
Printed by: Gildeprint Drukkerijen, Enschede (www.gildeprint.nl)
ISBN: 978-94-6419-743-3

Financial support by the Dutch Heart Foundation for the publication of this thesis is gratefully acknowledged.

Promotor
Prof. dr. K. Willems van Dijk

Co-promotores
Dr. ir. D. van Heemst
Dr. R. Noordam

Leden promotiecommissie
Prof. dr. S. le Cessie
Prof. dr. P.C.N. Rensen
Dr. S. Trompet
Dr. J. van Setten (University Medical Center Utrecht, Utrecht, the Netherlands)
Prof. dr. J.W.J. Beulens (Amsterdam University Medical Centers, Amsterdam, the Netherlands)

TABLE OF CONTENTS

1

GENERAL INTRODUCTION

INTRODUCTION

Developments in science, medicine and society, particularly throughout the last few centuries, have led to an increasing ability to treat and cure many previously deadly diseases. Although life expectancy has steadily increased from the 1800s onwards, in part due to better treatment of infectious diseases, different major causes of death emerged. The prevalence of non-communicable, age-related diseases such as cardiovascular disease (CVD) have greatly increased and this increase is expected to continue as our society further ages [1]. Although CVD has been studied for decades, and successful interventions have been identified and implemented, CVD remains a leading cause of death worldwide. CVD accounted for an estimated 18.6 million deaths worldwide in 2019 [2], and simultaneously accounted for a significant reduction in quality of life in those individuals that suffered a non-fatal event. For this reason, it is crucial to further expand our understanding of the mechanisms leading to CVD, to identify novel strategies that further reduce CVD disease risk.

Advancing age is a primary risk factor for many major chronic diseases [3]. Ageing is generally defined as broad, time-dependent functional decline. The cellular and molecular ageing processes have been separated into the 9 hallmarks of ageing: Genomic instability, telomere attrition, epigenetic alterations, loss of proteostasis, deregulated nutrient-sensing, mitochondrial dysfunction, cellular senescence, stem cell exhaustion, and altered intercellular communication [3]. These 9 hallmarks are thus processes generally thought to contribute to the ageing process and together determine the ageing phenotype.

Mitochondrial dysfunction is one of the hallmarks that may underly various adverse effects observed in ageing. Mitochondria are an important source of Reactive Oxygen Species (ROS), as an inevitable byproduct of their essential role in energy production [4]. ROS are generated during the process of oxidative phosphorylation and under physiological conditions efficiently scavenged. Although normal cellular function requires some degree of ROS (e.g., for intracellular communication purposes), pathological processes such as ageing may result in a disbalance between ROS production and scavenging (**Figure 1**). This disbalance is characterized by insufficient scavenging capacity to effectively eliminate ROS. As ROS are a highly reactive chemical, they could damage surrounding cellular components by reacting with them, ultimately causing cell damage or cell death [5]. To prevent ROS from inflicting cellular damage, different compounds collectively named antioxidants act as scavengers of these free radicals aiming to maintain appropriate ROS levels. Therefore, both optimal ROS production and sufficient antioxidant activity is pivotal to prevent oxidative damage [6].

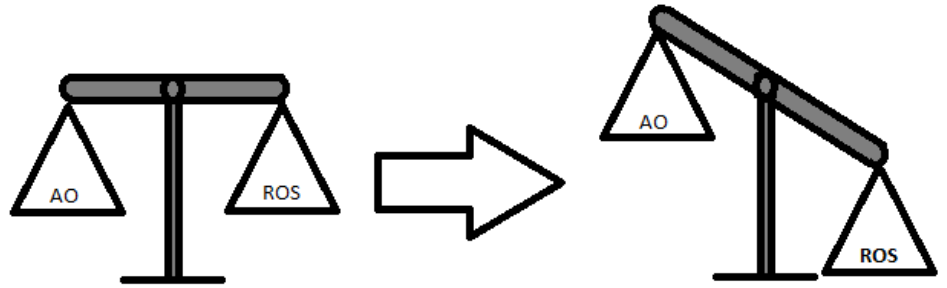

Figure 1. When either ROS production is increased or antioxidant activity is decreased, this may result in a disbalance. This disruption could lead to cellular damage or even cell death.

Oxidative stress is shown to play an important role in the pathophysiology of stroke, one of the main types of CVD [7, 8]. Increased ROS levels could lead to direct or indirect damage to the vascular wall via altered platelet aggregation, endothelial dysfunction, dysfunctional vasodilation, and/or disturbed vascular permeability [8]. These changes could lead to local lesions and might therefore increase the risk of clinical manifestations such as (ischemic) stroke. Lipid peroxidation by ROS is one of the oxidative processes that might contribute to the development of CVD. Peroxidized lipids can subsequently induce proinflammatory responses in the vascular system which play a crucial role in the initiation and progression of CVD [9-11].

Several studies have shown that decreased blood antioxidant levels are associated with increased CVD incidence [7, 12-15]. However, intervention studies exploring antioxidant supplementation have been unable to show reduction in the risk of developing clinical outcomes related to CVD [16, 17]. This lack of effectiveness of the studied antioxidant compounds thus far does not provide definitive proof for the role of ROS in CVD risk.

Although Randomized Controlled Trials (RCTs) are generally considered the gold standard of evidence, the effects of an altered oxidative stress balance might not be visible in the relatively short time period of an RCT. Additionally, the pathogenesis of atherosclerosis, a process where lesions are formed in the arteries that could lead to the buildup of plaque and eventually could result in CAD and stroke, often takes decades [18]. In other words, it might be improbable to demonstrate the effect of an altered oxidative stress balance on the development of CVD using RCTs. Thus, whether oxidative stress should ultimately be seen only as a marker of CVD risk might still require further investigation.

Epidemiology

Epidemiological research, while simple in principle, can be prone to oversimplification. To present a famous example: When studying lung cancer, one would probably establish that lung cancer patients have an increased chance of suffering from yellow stained

fingers. However, common sense would tell us that yellow fingers do not cause lung cancer. Simultaneously, lung cancer does not result in yellow fingers. The crux is that both phenotypes are caused by a third factor, smoking. Smoking in this case can be considered a *confounding* factor. Confounding is a common problem affecting most, if not all, epidemiological studies. The concept of confounding is visualized in **Figure 2**. A variable is considered a confounder if it can affect both the exposure and the outcome of the study. Knowing what potential confounders could be, and how to correct for them, is what makes epidemiological research complex.

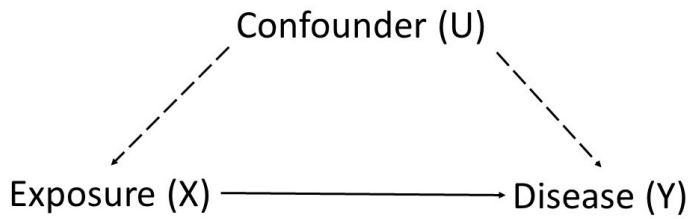

Figure 2. The concept of confounding. Both the exposure (X) and the outcome (Y) are influenced by the confounder (U). Therefore, the assumed association could be caused by the confounder.

This thesis aims to build on the basic premise of epidemiology, while simultaneously taking a more sub-group specific approach. Where risk factors and their effect are generally determined at a population level, the focus shifts towards the investigation of the difference in individual effect that these risk factors have on CVD, dependent on additional factors. This could potentially lead to a better understanding of the multidimensionality of CVD, and how risk factors affect the disease risk of other factors.

Reverse causation

Another problem within epidemiology is the concept of causality. Although observational research can link two occurrences together, the certainty that one conditions leads to another is not guaranteed. On top of that, the directionality of the relation cannot be assured. In theory, instead of an increased risk factor causing the observed outcome, the outcome could cause an increase in risk factor. This is what we call reverse causation. It is important to understand direction in a relationship, as identifying a causal variable is crucial in the eventual prevention of the outcome.

Mendelian Randomization

RCTs are the gold standard to assess whether an observational association is also a causal relation. However, RCTs are often costly, may be impossible for lack of intervention tools or may even be unethical. [19] An alternative approach one can use to assess causal relationships is the Mendelian Randomization (MR) method [20, 21]. MR uses independent genetic variants associated with the exposure as instrumental variable (**Figure 3)** [22]. Since the genetic make-up of an individual is unlikely to be affected by outside factors, when performing an analysis using genetic variants as a proxy for the determinants, the potential association is not susceptible to confounding or reverse causation and therefore allows us to approximate the causal association between the exposure and the outcome.

The genetic variants associated with the exposure are often SNPs obtained from Genome-Wide Association Studies (GWAS) on the exposure. Since the size and availability of genetic data has increased substantially over the last decade, more genetic instruments have become obtainable for use and thus more MR studies can be done on many exposures in large sample sizes that are implausible for RCTs. However, it is important to note that in order to be able to perform a Mendelian Randomization study and interpret causality, several assumptions are made. First, the genetic variants used as instrument variables for the exposure are associated with the exposure. By performing a GWAS on the exposure you ensure that this assumption is not violated. Second, there is no confounding of genetic variants with the outcome. Third, the genetic variants are only associated with the outcome through the exposure. When genetic variants are directly associated with the outcome, you are no longer able to discern the lone effect of the exposure [23]. These assumptions are often tested during a study, and although not all three of them can definitively be proven to be fulfilled, it is possible to assess the likelihood of an assumption being violated. Finally, it should be noted that MR assesses the assumed lifelong exposure of the determinant, compared with the limited exposure of an RCT. However, as MR attempts to assess the lifelong exposure, it is important to realize that MR assumes the effect of the genetic variants do not change over time, which might not always be true.

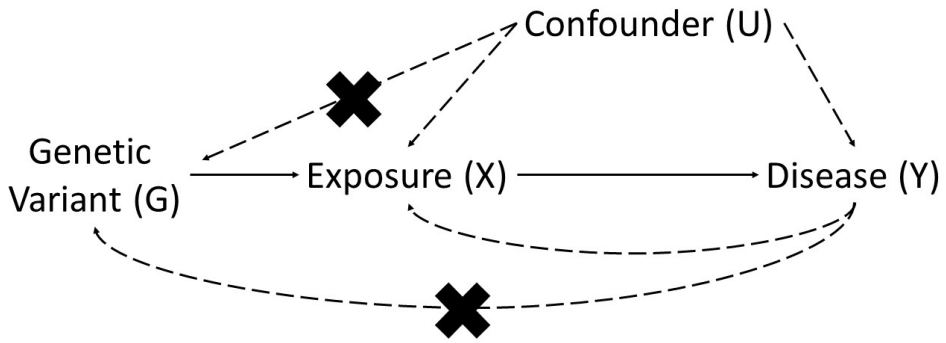

Figure 3. Mendelian Randomization: The genetic variants are associated with the exposure and used as a proxy. As the genotype is independent of confounders and unaffected by reverse causation, a found association would be causal.

Interplay of risk factors

For certain risk factors for CVD, causality has already been established. It is known, among others, that obesity, high blood pressure and high LDL cholesterol increase the risk on the development of CVD, and that interventions targeting these risk factors lower the risk substantially [2, 24-26]. However, CVD is a multifactorial, dynamic disease construct. Consequently, research should take into account the interplay of risk factors when studying CVD. Aside from the sum of effects of relevant risk factors, their potential interactive effect should also be taken into account. In other words, the presence of one risk factor might influence another risk factor in such a way it alters their association with CVD.

OUTLINE OF THIS THESIS

The general aim of this thesis is to address the phenotypical causes of mitochondrial dysfunction, resulting in oxidative imbalance, and the cerebrovascular consequences in the general population.

Possible risk factors will be studied for causality and established risk factors will be studied for causality differences in subgroups, with the ambition to increase the understanding of the impact of these known risk factors on individual cardiovascular disease risk. Therefore, a proof-of-concept study will be performed to investigate how causal relations between cardiovascular risk factors and coronary artery disease can be altered by socioeconomic status. The studies of this thesis will be performed in several unique cohorts. The Netherlands Epidemiology of Obesity (NEO) study is a population-based, prospective cohort study designed to study obesity-related pathways and diseases [27]. Although the main cohort is oversampled with individuals with overweight or obesity, this study will use a subsample from Leiderdorp, with no BMI selection for participation. The UK Biobank is a UK-based prospective cohort study recruited from the general population [28].

In **Chapter 2**, I will measure serum α-TOH and vitamin E metabolites, as a reflection of an individual's vitamin E status, and study their association with the lifestyle factors smoking, sleep, physical activity, and food and alcohol intake. This study will be performed in the NEO study.

Chapter 3 will focus on the causal relation between several antioxidants and cerebrovascular disease risk. Here, I look for a possible causal relationship between genetically-influenced diet-derived antioxidants with ischemic stroke using Mendelian Randomization (MR). This study will be done using summary level data from three large databanks: the UK Biobank, MEGASTROKE, and FinnGen, resulting in a total of 1.1 million participants.

Chapter 4 will study the association between mitochondrial DNA copy numbers (mtDNA-cn), as a marker for mitochondrial dysfunction which drives oxidative stress levels, and stroke using both an MR as well as a survival-analysis approach using the UK Biobank. The triangulation of causal inference as described by Lawlor et al. [29] will be adopted to increase the credibility of our results.

Chapter 5 is centered around the effect of individual characteristics on cardiovascular disease (CVD) risk factors. This study will be performed in the UK Biobank and is specifically focused on how socio-demographic characteristics could modify the causal association between classical CVD risk factors and coronary artery disease (CAD). In this study, I hypothesize that atherogenic cardiovascular disease is a multidimensional disease where different combinations of risk factors could alter the individual impact of said risk factors on CAD incidence.

In the final part of this thesis (**Chapter 6**), the main study findings and the future perspective of this research field will be discussed.

REFERENCES

1. Atella, V., A. Piano Mortari, J. Kopinska, F. Belotti, F. Lapi, C. Cricelli, and L. Fontana, *Trends in age-related disease burden and healthcare utilization.* Aging Cell, 2019. **18**: p. e12861.
2. Roth, G.A., G.A. Mensah, C.O. Johnson, G. Addolorato, E. Ammirati, L.M. Baddour, N.C. Barengo, A.Z. Beaton, E.J. Benjamin, C.P. Benziger, et al., *Global Burden of Cardiovascular Diseases and Risk Factors, 1990-2019: Update From the GBD 2019 Study.* J Am Coll Cardiol, 2020. **76**: p. 2982-3021.
3. Lopez-Otin, C., M.A. Blasco, L. Partridge, M. Serrano, and G. Kroemer, *The hallmarks of aging.* Cell, 2013. **153**: p. 1194-217.
4. Andreyev, A.Y., Y.E. Kushnareva, and A.A. Starkov, *Mitochondrial metabolism of reactive oxygen species.* Biochemistry (Mosc), 2005. **70**: p. 200-14.
5. Luo, J., K. Mills, S. le Cessie, R. Noordam, and D. van Heemst, *Ageing, age-related diseases and oxidative stress: What to do next?* Ageing Res Rev, 2020. **57**: p. 100982.
6. Parohan, M., J. Anjom-Shoae, M. Nasiri, M. Khodadost, S.R. Khatibi, and O. Sadeghi, *Dietary total antioxidant capacity and mortality from all causes, cardiovascular disease and cancer: a systematic review and dose-response meta-analysis of prospective cohort studies.* Eur J Nutr, 2019. **58**: p. 2175-2189.
7. Rodrigo, R., R. Fernandez-Gajardo, R. Gutierrez, J.M. Matamala, R. Carrasco, A. Miranda-Merchak, and W. Feuerhake, *Oxidative stress and pathophysiology of ischemic stroke: novel therapeutic opportunities.* CNS Neurol Disord Drug Targets, 2013. **12**: p. 698-714.
8. Allen, C.L. and U. Bayraktutan, *Oxidative stress and its role in the pathogenesis of ischaemic stroke.* Int J Stroke, 2009. **4**: p. 461-70.
9. Ridker, P.M., C.H. Hennekens, J.E. Buring, and N. Rifai, *C-reactive protein and other markers of inflammation in the prediction of cardiovascular disease in women.* N Engl J Med, 2000. **342**: p. 836-43.
10. Libby, P., P.M. Ridker, and A. Maseri, *Inflammation and atherosclerosis.* Circulation, 2002. **105**: p. 1135-43.
11. Ayala, A., M.F. Munoz, and S. Arguelles, *Lipid peroxidation: production, metabolism, and signaling mechanisms of malondialdehyde and 4-hydroxy-2-nonenal.* Oxid Med Cell Longev, 2014. **2014**: p. 360438.
12. Shirley, R., E.N. Ord, and L.M. Work, *Oxidative Stress and the Use of Antioxidants in Stroke.* Antioxidants (Basel), 2014. **3**: p. 472-501.
13. Sanchez-Moreno, C., A. Jimenez-Escrig, and A. Martin, *Stroke: roles of B vitamins, homocysteine and antioxidants.* Nutr Res Rev, 2009. **22**: p. 49-67.
14. Gillman, M.W., L.A. Cupples, D. Gagnon, B.M. Posner, R.C. Ellison, W.P. Castelli, and P.A. Wolf, *Protective effect of fruits and vegetables on development of stroke in men.* JAMA, 1995. **273**: p. 1113-7.
15. Yochum, L.A., A.R. Folsom, and L.H. Kushi, *Intake of antioxidant vitamins and risk of death from stroke in postmenopausal women.* Am J Clin Nutr, 2000. **72**: p. 476-83.
16. Leppala, J.M., J. Virtamo, R. Fogelholm, D. Albanes, P.R. Taylor, and O.P. Heinonen, *Vitamin E and beta carotene supplementation in high risk for stroke: a subgroup analysis of the Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study.* Arch Neurol, 2000. **57**: p. 1503-9.
17. Yusuf, S., G. Dagenais, J. Pogue, J. Bosch, and P. Sleight, *Vitamin E supplementation and cardiovascular events in high-risk patients.* N Engl J Med, 2000. **342**: p. 154-60.
18. Ahmadi, A., E. Argulian, J. Leipsic, D.E. Newby, and J. Narula, *From Subclinical Atherosclerosis to Plaque Progression and Acute Coronary Events: JACC State-of-the-Art Review.* J Am Coll Cardiol, 2019. **74**: p. 1608-1617.
19. Black, N., *Why we need observational studies to evaluate the effectiveness of health care.* BMJ, 1996. **312**: p. 1215-8.
20. Burgess, S., A. Butterworth, and S.G. Thompson, *Mendelian randomization analysis with multiple genetic variants using summarized data.* Genet Epidemiol, 2013. **37**: p. 658-65.
21. Burgess, S., R.A. Scott, N.J. Timpson, G. Davey Smith, S.G. Thompson, and E.-I. Consortium, *Using published data in Mendelian randomization: a blueprint for efficient identification of causal risk factors.* Eur J Epidemiol, 2015. **30**: p. 543-52.
22. Lawlor, D.A., R.M. Harbord, J.A. Sterne, N. Timpson, and G. Davey Smith, *Mendelian randomization: using genes as instruments for making causal inferences in epidemiology.* Stat Med, 2008. **27**: p. 1133-63.
23. de Leeuw, C., J. Savage, I.G. Bucur, T. Heskes, and D. Posthuma, *Understanding the assumptions underlying Mendelian randomization.* Eur J Hum Genet, 2022.
24. Koliaki, C., S. Liatis, and A. Kokkinos, *Obesity and cardiovascular disease: revisiting an old relationship.* Metabolism, 2019. **92**: p. 98-107.
25. O'Donnell, C.J. and R. Elosua, *[Cardiovascular risk factors. Insights from Framingham Heart Study].* Rev Esp Cardiol, 2008. **61**: p. 299-310.
26. Pucci, G., R. Alcidi, L. Tap, F. Battista, F. Mattace-Raso, and G. Schillaci, *Sex- and gender-related prevalence, cardiovascular risk and therapeutic approach in metabolic syndrome: A review of the literature.* Pharmacol Res, 2017. **120**: p. 34-42.
27. de Mutsert, R., M. den Heijer, T.J. Rabelink, J.W. Smit, J.A. Romijn, J.W. Jukema, A. de Roos, C.M. Cobbaert, M. Kloppenburg, S. le Cessie, et al., *The Netherlands Epidemiology of Obesity (NEO) study: study design and data collection.* Eur J Epidemiol, 2013. **28**: p. 513-23.
28. Sudlow, C., J. Gallacher, N. Allen, V. Beral, P. Burton, J. Danesh, P. Downey, P. Elliott, J. Green, M. Landray, et al., *UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age.* PLoS Med, 2015. **12**: p. e1001779.
29. Lawlor, D.A., K. Tilling, and G. Davey Smith, *Triangulation in aetiological epidemiology.* Int J Epidemiol, 2016. **45**: p. 1866-1886.

2

ASSOCIATIONS BETWEEN LIFESTYLE FACTORS AND VITAMIN E METABOLITES IN THE GENERAL POPULATION

Leon G Martens
Jiao Luo
Fleur L Meulmeester
Nadia Ashrafi
Esther W van Eekelen
Renée de Mutsert
Dennis O Mook-Kanamori
Frits R Rosendaal
Ko Willems van Dijk
Kevin Mills
Raymond Noordam
Diana van Heemst

Antioxidants (Basel), 2020 Dec 15;9(12):1280

ABSTRACT

Background and aims

The antioxidant vitamin E (α-tocopherol, α-TOH) protects lipids from oxidation by reactive oxygen species. We hypothesized that lifestyle factors associate with vitamin E metabolism marked by urinary α-tocopheronolactone hydroquinone (α-TLHQ) and α-carboxymethyl-hydroxychroman (α-CEHC levels), as a potential reflection of lipid oxidation.

Methods

We conducted a cross-sectional study in the Netherlands Epidemiology of Obesity Study. Serum α-TOH, and urinary α-TLHQ and α-CEHC were quantified by liquid chromatography coupled with tandem mass spectrometry. Information on the lifestyle factors (sleep, physical activity, smoking and alcohol) were collected through questionnaires. Multivariable linear regression analyses were performed to assess the associations between the lifestyle factors and α-TOH measures.

Results

530 participants (46% men) were included with mean (SD) age of 56(6) years. Of the examined lifestyle factors, only poor sleep was associated with a higher serum α-TOH (mean difference: 4% [95%CI: 1, 7%]). Current smoking was associated with higher urinary α-CEHC (32%: [14%, 53%]), with evidence of a dose-response relationship with smoking intensity (low packyears, 24% [2, 52%]; high packyears, 55% [25, 93%]). Moderate physical activity was associated with a lower α-TLHQ relative to α-CEHC (-17%: [-26,-6%], compared with low PA).

Conclusions

Only specific lifestyle factors associate with vitamin E metabolism. Examining serum α-TOH does not provide complete insight in vitamin E anti-oxidant capacity.

INTRODUCTION

The lipid-soluble antioxidant vitamin E is a defensive compound neutralizing reactive oxygen species (ROS), specifically when formed during lipid peroxidation [1]. The most common form of vitamin E in the blood is α-tocopherol (α-TOH) [2]. During the process of neutralizing lipid peroxidation reactions, α-TOH can be metabolized through two different mechanisms, notably enzymatic degradation and oxidation. When α-TOH is engaged in an oxidative reaction, the chromanol ring is opened which results in the formation of the oxidation product α-tocopheronolactone hydroquinone (α-TLHQ) [1, 3, 4]. The remaining is degraded through enzymatic conversion into α-carboxymethyl-hydroxychroman (α-CEHC), and is a measure of α-TOH status. Once formed, both metabolites are processed in the liver and excreted via the urine [4].

In most published studies, an individual's vitamin E status is reflected by the measured α-TOH level in the blood serum. In these studies, increased levels of α-TOH have been associated with reduced risks of age-associated diseases [5]. A meta-analysis done by Feng-Jiao Li and colleagues showed that dietary intake of vitamin E, vitamin C and β-carotene, all anti-oxidants, lowered the risk on Alzheimer's disease [6]. Furthermore, a study done in 39.910 middle-aged men observed a lower risk of coronary heart disease among men with a higher vitamin E intake [7]. However, recent studies have suggested that serum α-TOH does not correlate with vitamin E anti-oxidant activity nor with vitamin E metabolism [8]. Steven Steinhubl mentions in his review that although based on observational data there seems to an association between high vitamin E intake and a decreased risk of CVD, clinical trials involving vitamin E supplementation have failed to show an effect on the prevention of CVD [9]. These observations may explain why clinical trials have shown no beneficial effects of vitamin E supplementation on the risk of age-related disease [10-15]. Instead, there is increasing evidence that measuring vitamin E metabolites (notably CEHC and TLHQ) in urine provides a more accurate and reliable estimate of the vitamin E antioxidant status in the body [16]. However, it remains to be elucidated what factors influence vitamin E metabolite concentration.

A particular candidate for factors that affect vitamin E metabolism is lifestyle, which is generally accepted to have a serious impact on health [17-21]. For example, large studies have shown that smoking is associated with an increased the risk of diabetes mellitus and cardiovascular events [22, 23]. Furthermore, it is well-established that an increase in physical activity results in clinically relevant health benefits [24]. One potential mechanism linking lifestyle to health may be ROS production or antioxidant activity [1]. As lifestyle factors such as smoking and physical activity are strongly linked to lipid peroxidation, and vitamin E is the main antioxidant that confers protection against lipid peroxidation-induced damage [25, 26], we hypothesized that lifestyle factors are associated with vitamin E metabolite levels. Indeed, multiple of the main lifestyle factors have been studied previously in relation to serum α-TOH, but evidence is lacking for

the vitamin E metabolite levels [27, 28]. Previously, such hypothesis has already been postulated for nutrition [29].
In this study, we measured serum α-TOH levels in fasting serum samples and vitamin E metabolites in 24-hour urine and aimed to examine associations of the main lifestyle factors smoking, sleep, physical activity, and habitual food and alcohol intake with these levels in a cross-sectional cohort of middle-aged individuals.

MATERIALS AND METHODS

Setting and Study Design

The NEO study is a population-based, prospective cohort study designed to investigate pathways that lead to obesity-related diseases. The NEO study started in 2008 and includes 6,671 individuals aged 45–65 years, with an oversampling of individuals with overweight or obesity. The study design and population has been described in detail elsewhere [30]. The Medical Ethical Committee of the Leiden University Medical Center (LUMC) approved the design of the study. All participants gave their written informed consent.

In short, men and women living in the greater area of Leiden (in the West of the Netherlands) were invited by letters sent by general practitioners, municipalities and by local advertisements. They were invited to respond if they were aged between 45 and 65 years and had a self-reported Body Mass Index (BMI) of 27 kg/m2 or higher. In addition, all inhabitants aged between 45 and 65 years from one municipality (Leiderdorp) were invited to participate irrespective of their BMI. This resulted in a population of 1,671 participants with a BMI distribution similar to that of the general population [30].

Participants were invited to a baseline visit at the NEO study centre of the LUMC after an overnight fast. Prior to this study visit, participants collected their urine over 24 h and completed a general questionnaire at home to report demographic, lifestyle and clinical information, habitual food intake and physical activity. The participants were asked to bring all medication they were using in the month preceding the study visit and names and dosages of all medication were recorded. At the baseline visit several measurements were performed including anthropometry and blood pressure, and blood sampling after an overnight fast.

For the present analysis, we selected participants from the Leiderdorp municipality comprising of a Caucasian Dutch population with no BMI requirements for participation. At the baseline visit participants completed a screening form to identify contra-indications for undergoing magnetic resonance imaging (MRI) (most notably metallic devices, claustrophobia or a body circumference of more than 1.70m). Of the participants who were eligible, approximately 40% were randomly selected to undergo MRI. All individuals with available urine collected for more than 20 hours were included for vitamin E metabolites measurements (n=539). Individuals that had unrealistic levels of vitamin E metabolites were excluded (n=9).

Lifestyle factors

All lifestyle exposures were collected through self-reported questionnaires. Smoking was defined as current, former and never smoking. Long-term tobacco exposure was expressed in pack years of smoking, calculated by multiplying the number of packs of cigarettes smoked per day by the number of years the person smoked. One pack year is defined as twenty cigarettes smoked every day for one year. Additionally, to limit potential measurement error and have an objective measure for smoking exposure, we performed analyses on the relation between the metabolite cotinine and urinary vitamin E metabolites. Cotinine, a xenobiotic metabolite for nicotine, was measured using untargeted metabolomics measurements at Metabolon Inc. (Durham, North Carolina, USA) using their Metabolon™ Discovery HD4 platform. In brief, this process involves four independent ultra-high-performance liquid chromatography mass spectrometry (UHPLC-MS/MS) platforms [31, 32]. Two platforms used positive ionization reverse phase chromatography, one used negative ionization reverse phase chromatography, and one used hydrophilic interaction liquid chromatography (HILIC) negative ionization [32]. Diet quality was measured with the Dutch Healthy Diet Index (DHD-index), which measures the adherence to the Dutch dietary guidelines using a semiquantitative self-administered 125-item Food Frequency Questionnaire (FFQ) [33]. For the present study, we used an adapted version of the DHD-index with thirteen components instead of the original fifteen because we were not able to estimate the two components consumption of unfiltered coffee and of sodium on the basis of the FFQ used in our study. As a result, the DHD-index in our study ranges between 0 and 130, where a higher score indicates a better diet quality. Sleep quality was measured with the Pittsburgh Sleep Quality Index (PSQI) questionnaire [34]. This questionnaire consists of 19 items with a total score of 21, where a higher score indicates a worse sleep quality. Alcohol consumption was estimated using the FFQ and restructured to the unit of gram per day. Participants reported the frequency and duration of their usual physical activity during leisure time in the Short Questionnaire to assess health-enhancing physical activity (SQUASH), which was expressed in hours per week of metabolic equivalents [35].

Other variables

Body weight was measured and percent body fat was estimated by the Tanita bio impedance balance (TBF-310, Tanita International Division, Manchester, UK) without shoes and 1 kg was subtracted to correct for weight of clothing. BMI at baseline was calculated by dividing the weight in kg by the height in meters squared. LDL-cholesterol concentration was estimated using the Friedewald formula [36].

Serum and urinary vitamin E metabolites

Fasting blood samples were obtained from the antecubital vein. Standard clinical laboratory parameters including glucose and insulin concentrations and lipid profile were obtained in the clinical chemistry laboratory of the LUMC. The remaining blood samples were aliquoted and stored at-80° C for future research. Both measures were performed in 2019 making the storage period similar. Serum α-tocopherol levels were obtained using untargeted LC-MS/MS (Metabolon, Inc, North Carolina, USA), as highlighted in more detail above [37].

Prior to the study visit, participants collected urine over 24 hours and recorded the time of the first and last void. In addition, participants collected the first morning spot on the day of their study visit. 24-h Urinary oxidized α-TOH metabolites (α-TLHQ) and enzymatic metabolites (α-CEHC), presented as their sulfate or glucuronide conjugates (α-TLHQ-SO3, α-TLHQ-GLU, α-CEHC-SO3, α-CEHC-GLU), were measured by LC-MS/MS at University College London, UK.

Urine samples were thawed, and 100µl fresh urine was then centrifuged in Eppendorf tubes for 10 min at 13,000 rpm at room temperature and spiked with 10µl of the internal standards (100µmol/L), lithocholic acid sulfate (LA) and androsterone D4-glucuronide (AD4). Subsequently, samples were vortexed and transferred into screw-cap glass vials. 10µl was injected into the LC-MS/MS for detection.

The metabolites were separated using a Waters ACQUITY UPLC BEH C8 column (1.7µm particles, 50mm x 2.1mm; Waters Corp, Manchester, UK) plus a guard column containing an identical stationary phase. The mobile phase was a gradient elution of solvent A (99.98% water; 0.01% (v/v) formic acid) and solvent B (99.98% acetonitrile/MeCN; 0.01% (v/v) formic acid), which were LC-MS grade or equivalent (Sigma-Aldrich Co. Ltd). The flow rate was set to 0.8mL/min and the LC gradient was established by coordinating the solvents as follows: 95% solvent A plus 5% solvent B for 0 to 0.40 min; 80% solvent A plus 20% solvent B for 2 min; 0.1% solvent A plus 99.9% solvent B for 3.01 to 4 min; 95% solvent A plus 5% solvent B for 4.01 to 5 min. In order to minimize system contamination and carry over, the MS diverter valve was set up to discard the UPLC eluent before and after the sample elution, at 0 to 0.40 min and 4.01 to 5 min, respectively, as well as an additional run of blank sample (H2O: MeCN) between each run of urine samples. Two peaks were observed for α-TLHQ and α-CEHC glucuronide conjugates, corresponding to major and minor isoforms.

After separation, the metabolites were then analyzed by MS using a Waters ACQUITY UPLC coupled to a triple-quadrupole Xevo TQ-S fitted with an electrospray ionization in negative ion mode. The gas temperatures persisted 600°C for desolvation. In addition, nitrogen was used as the nebulizing gas with 7.0 Bar. The cone voltages were set at 56 V and 54 V, and the collision voltages at 28 eV and 30 eV for sulfate conjugates and glucuronide conjugates, respectively. Running time for each sample is 5 minutes with a 20µL injection volume together with a partial loop with needle overfill mode. Using multiple reaction monitoring (MRM) mode, specific parent and daughter ions were determined in scan mode and following collision activated dissociation (CAD) with argon. These ions were then used to quantify each α-TOH metabolite from transitions (α-TLHQ-GLU, 453.3>113.0; α-TLHQ-SO3, 357.1>79.9; α-CHE-GLU, 453.2>113.0; α-CHE-SO3, 357.1>79.9) that corresponded to their theoretical molecular masses.

Urinary creatinine concentrations (mmol/L) were also measured to correct for dilution differences for each metabolite, by triple-quadrupole Micro Quattro mass spectrometry (MicroMass, Waters, UK) using deuterated creatinine as the internal standard. Therefore, the concentrations of α-TOH metabolites are expressed as nmol per mmol of creatinine. A quality control (QC) assessment was performed throughout the quantification both in creatinine and α-TOH metabolite assays to deal with the variations in sample quality and UPLC-MS/MS performance over time. Four QC samples were systematically interleaved every 50 urine samples to limit the amount of sample loss. The whole measurement protocol was developed and further modified by the detection group in London [37, 38].

The final concentration of glucuronide conjugates for α-TLHQ and α-CEHC were the sum of their corresponding major and minor isoforms. In addition to the measured single metabolite, total, glucuronide and sulfate conjugates ratios were further determined to reflect the α-TOH antioxidative capacity as well as lipid peroxidation levels taking α-TOH status into consideration, namely α-TLHQ-to-α-CEHC ratio, α-TLHQ-GLU-to-α-CEHC-GLU ratio, and α-TLHQ-SO3-to-α-CEHC-SO3 ratio.

Statistical Analysis

We examined the characteristics of the total study population and stratified by sex. Characteristics are presented as mean (standard deviations), medians (with interquartile ranges; skewed variables only) and proportions. Serum α-TOH, and urinary α-TLHQ and α-CEHC had a skewed distribution and were natural log-transformed. To assess the correlation between α-TOH, α-TLHQ, and α-CEHC, we performed a Pearson correlation test.

Participants with a PSQI score of 5 or higher, were assigned to the poor sleep quality group. The group of participants with a score lower than 5 were assigned to the good sleep quality group and considered the reference. For smoking, participants were distributed into three groups: (i) non-smokers (reference), (ii) former smokers and (iii) current smokers. For both diet quality, alcohol use and physical activity, the group was divided into quartiles, with the lowest quartile as the reference.

For the analyses, we examined the associations between each individual lifestyle factor and vitamin E metabolite levels using multivariable linear regression analyses adjusted for age and sex. Additionally, we adjusted for BMI and we also performed a sensitivity analysis where we substituted BMI for total body fat. As results were not different between adjusting for BMI and total body fat, we only reported results adjusted for BMI. To study the possible association between smoking cessation on α-TOH and its metabolites among former smokers, we stratified the group of former smokers based on the duration of smoking cessation (less or more than 5 years of smoking cessation). As we

did not observe a difference in levels of α-TOH and its metabolites, we treated the group of former smokers as one in all analyses. As lifestyle factors are likely to be considerably different between men and women, we repeated all analyses after stratifying by sex. After the analyses, all beta estimates and 95% confidence limits were backtransformed to be able to present these as a percentage difference. Consequently, all results can be interpreted as the percentage difference in outcome compared with the reference category, with the 95% confidence interval.

RESULTS

Characteristics of the study population

After excluding the nine outliers, a total of 530 participants (46% men) were analyzed. The mean (SD) age was 55.9 (6.0), and BMI 25.9 (4.0). The baseline characteristics are presented in **Table 1**. Among current smokers, men had smoked more in pack years than women (median: 26.6, IQR: 14.0, 36.4 versus median: 13.3, IQR: 4.7, 22.7). Men also consumed more alcohol (gram/day) (median: 16.7, IQR: 5.1, 28.4 versus median: 7.2, IQR: 1.0, 14.4). α-TLHQ and α-CEHC were correlated with a Pearson correlation of 0.70. α-TLHQ and α-TOH had a Pearson correlation of 0.24. α-CEHC and α-TOH had a Pearson correlation of 0.21.

Table 1. Characteristics of the NEO participants stratified by sex

	All(N=530)	Men(N=246)	Women(N=284)
Demography	55.9(6.0)		
Sex, % men	46.4		
Age [years], mean (sd)	55.9(6.0)	56.2(6.2)	55.6(5.8)
Lifestyle factors			
BMI [kg/m^2], mean (sd)	25.9(4.0)	26.6(3.3)	25.4(4.4)
Smoking, N (%)			
Never	216(40.8)	92(37.4)	124(43.7)
Former	255(48.1)	124(50.4)	131(46.1)
Current	59(11.1)	30(12.2)	29(10.2)
Packyears[1], median (IQR)	18.1(8.4, 29.8)	26.6(14.0, 36.4)	13.3(4.7, 22.7)
Alcohol use [g/day], median (IQR)	9.2(2.5, 21.5)	16.7(5.1, 28.4)	7.2(1.0, 14.4)
Leisure activity [MET-hour], median (IQR)	30.0(16.5, 49.6)	30.0(16.0, 50.0)	29.9(16.5, 47.8)
Poor Sleep Quality, N (%)	205(38.7)	76(30.9)	129(45.4)
Diet Quality [0-130], mean (sd)	72.0(14.6)	68.4(13.4)	75.2(14.9)
Lipid lowering drugs, N (%)	38(7.2)	25(10.2)	13(4.6)
Vitamin E metabolites			
Measurements, median (IQR)			
α-THLQ (nmol/mmol creatinine)	1832(1337, 2745)	1519(1194, 2256)	2091.6(1470, 3074)
α-CEHC (nmol/mmol creatinine)	265(180, 439)	223(137, 352)	307(213, 508)
α-TLHQ/α-CEHC	2.0(1.6, 2.3)	2.0(1.7, 2.4)	2.0(1.6, 2.3)
α-TOH	$3.5*10^8$($3.2*10^8$, $3.9*10^8$)	$3.5*10^8$($3.2*10^8$, $3.8*10^8$)	$3.5*10^8$($3.1*10^8$, $3.9*10^8$)

Data are presented as mean (standard deviation) or median (interquartile range) for numerical variables, and number (proportions) for categorical variables. [1]Packyears only measured in current smokers.

Association between smoking (intensity) and vitamin E (metabolite) levels in serum and urine

The association between smoking and vitamin E serum and urinary metabolites is displayed in **Table 2**. After adjusting for age, sex, and potential confounding factors such as BMI and alcohol consumption, there was no evidence for a difference in mean α-TOH (-3% Beta [95% CI: -8, 2]) level between current smokers and never-smokers. For the urinary α-TOH metabolites, current smoking was associated with a 32% lower α-CEHC (-32% Beta [95% CI: -44, -18]) and a 32% higher TLHQ relative to CEHC ratio (32% Beta [95% CI: 14, 53]), after adjustments for possible confounding. In addition, we observed no difference in mean levels of α-TOH or its metabolite levels between never smokers and past smokers. An increased smoking intensity (measured in pack years, and restricted to never and current smokers) was associated with lower α-TOH levels, with current smokers grouped in the high packyears group having a 10% (95% CI: -16, -3]) lower α-TOH level than never smokers (**Figure 1D**). For urinary vitamin E metabolites, both low and high packyears were associated with reduced α-CEHC (**Figure 1A**) and an increased TLHQ relative to CEHC ratio (**Figure 1C**), in a clear dose-response relationship. Notably, α-CEHC was 25% (95% CI: -43, -2) lower in current smokers with low packyears and 46% (95% CI: -59, -28) lower in current smokers with high packyears, than in never smokers. The TLHQ relative to CEHC ratio was 24% (95% CI: 2, 52) higher in current smokers with low packyears and 55% (95% CI: 25, 93) higher in current smokers with high packyears, than in never smokers.

As an objective measure of smoking status, the presence of cotinine in serum was associated with lower α-CEHC (difference: -20%, 95%CI: -30, -8). Furthermore, after splitting the cotinine present group in two halves, the group with the highest cotinine level had a significantly lower α-CEHC (difference:-28%, 95%CI:-40,-14) when compared with the reference group (**Table 3**).

Table 2. Associations between lifestyle factors and different measurements of vitamin E activity.

Lifestyle factors		α -TOH		Total TLHQ		Total CEHC		Ratio TLHQ/CEHC	
	N =	% change	95% CI	% change	95% CI	% change	95% CI	% change	95% CI
Smoking									
Nonsmokers	216	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
Former smokers	255	-1	-4 to 2	4	-6 to 15	2	-10 to 15	2	-7 to 12
Current smokers	59	-4	-8 to 0	-11	-24 to 5	-32	-44 to-18	32	14 to 53
DHDI									
25.4- 61.5	132	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
61.9- 72.5	133	-2	-3 to 4	0	-13 to 15	6	-10 to 25	-7	-18 to 4
72.5- 81.5	133	1	-2 to 5	13	-2 to 29	6	-10 to 25	5	-7 to 18
81.6- 117.8	132	3	-1 to 7	12	-3 to 29	10	-7 to 31	-1	-13 to 12
Sleep Quality									
Good Quality	288	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
Poor Quality	205	3	1 to 6	0	-10 to 10	-0	-12 to 13	1	-9 to 9
PA in MET-hours/wk									
0.0- 16.3	129	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
16.5- 29.8	130	1	-2 to 5	-7	-18 to 7	11	-5 to 31	-17	-26 to-6
30.0- 49.5	132	-1	-4 to 3	-11	-22 to 2	0	-14 to 18	-12	-22 to-1
49.8- 242.5	130	-2	-6 to 1	-5	-17 to 14	4	-11 to 23	-9	-19 to 3
Alcohol use in g/day									
0- 2.49	132	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
2.5- 9.2	133	-2	-5 to 2	-8	-19 to 6	-1	-16 to 16	-8	-18 to 4
9.2- 21.5	133	-0	-3 to 4	-9	-21 to 4	2	-14 to 20	-12	-22 to 0
21.5- 171.6	132	3	-1 to 7	-17	-28 to 4	-4	-20 to 14	-13	-24 to-1

Results are derived from linear regression model adjusted for age, sex and BMI with 95% confidence interval (CI) and presented as percentage difference compared with the reference group. Analysis on smoking were additionally adjusted for alcohol use. Analysis on diet quality, sleep quality and alcohol use were additionally adjusted for smoking. Analysis on PA was additionally adjusted for smoking and alcohol.
DHDI, Dutch Healthy Diet Index; PA, Physical Activity; TOH, tocopherol; TLHQ, tocopheronolactone hydroquinone; CEHC, carboxymethyl-hydroxychroman.

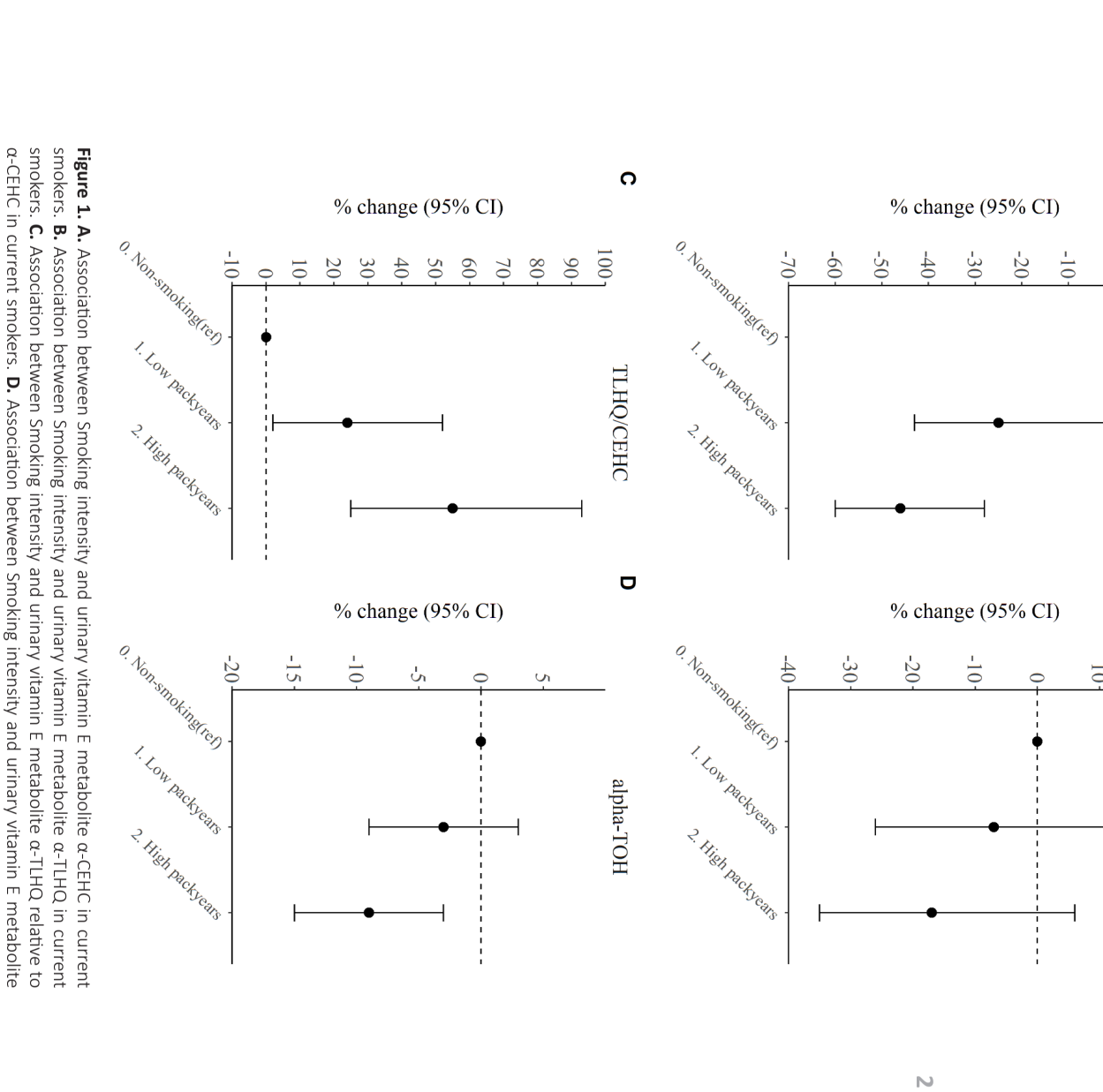

Figure 1. A. Association between Smoking intensity and urinary vitamin E metabolite α-CEHC in current smokers. **B.** Association between Smoking intensity and urinary vitamin E metabolite α-TLHQ in current smokers. **C.** Association between Smoking intensity and urinary vitamin E metabolite α-TLHQ relative to α-CEHC in current smokers. **D.** Association between Smoking intensity and urinary vitamin E metabolite α-TOH in current smokers. Results are derived from linear regression model adjusted for age, sex, BMI and alcohol with 95% confidence interval (CI) and presented as percentage difference compared with the reference group. Of the current smokers, the 50% with the lowest packyears is considered the low packyears group, whereas the highest 50% is in the high packyears group. CEHC, carboxymethyl-hydroxychroman.

Table 3. Associations between Cotinine and different measurements of vitamin E activity

Cotinine Metabolite		α -TOH		Total TLHQ		Total CEHC		Ratio TLHQ/CEHC	
	N =	% change	95% CI	% change	95% CI	% change	95% CI	% change	95% CI
Cotinine									
Absent	412	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
Present	118	2	-2 to 5	-11	-21 to-1	-20	-30 to-8	11	0 to 23
Cotinine levels									
Absent	412	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
Low	59	5	0 to 10	-6	-19 to 10	-11	-26 to 7	6	-7 to 21
High	59	-2	-6 to 3	-16	-29 to-3	-28	-40 to-14	16	1 to 33

Results are derived from linear regression model adjusted for age, sex, BMI or total body fat and alcohol use with 95% confidence interval (CI) and presented as percentage difference compared with the reference group. Of the group with cotinine levels present, the 50% with the lowest levels is considered the low group, whereas the highest 50% is in the high group.
α-TOH, α-tocopherol ; TLHQ, tocopheronolactone hydroquinone; CEHC, carboxymethyl-hydroxychroman

Diet quality and vitamin E serum and urinary levels

There was no association between the DHD-index and levels of vitamin E or its metabolites, (**Table 2**). For the fourth quartile, the level of α-TOH was not different (2% [95% CI:-2, 7]) from what was observed in those grouped having the worst diet quality. A similar result was observed for the vitamin E metabolites α-TLHQ (12% [95% CI: -3, 29]), α-CEHC (10% [95% CI:-7, 31]) and the α-TLHQ relative to α-CEHC ratio (-1% [95% CI:-13, 12]).

Sleep quality and vitamin E serum and urinary levels

Poor sleep quality was associated with increased serum α-TOH levels (4% Beta [95% CI:1, 7]). No association was found between sleep quality and any urinary metabolite levels (α-TLHQ: 0% Beta [95% CI:-10, 10], α-CEHC: 0% Beta [95% CI:-12, 13], α-TLHQ relative to α-CEHC ratio: 1% Beta [95% CI:-9, 9]) (**Table 2**).

Physical activity and vitamin E serum and urinary levels

Compared with the lowest quartile of physical activity during leisure time, medium and high physical activity showed no different mean level of α-TOH (medium PA: -1% Beta [95% CI:-5, 3], high PA: -3% Beta [95% CI:-7, 9]). For urinary vitamin E metabolites, we observed a u-shaped association with physical activity intensity. Notably, the TLHQ relative to CEHC ratio was lower in the second quartile (-17% [95% CI:-26,-6]) and third quartile (-12% [95% CI:-22,-1]) compared to the reference (**Table 2**). However, no there was no difference in TLHQ relative to CEHC ratio and any urinary vitamin E metabolite level between the highest quartile of physical activity and the lowest quartile.

Alcohol consumption and vitamin E serum and urinary levels

We did not observe an association between alcohol intake and α-TOH level in serum. However, after adjusting for age, sex, BMI and smoking, alcohol consumption in the highest quartile was associated with a 17% lower α-TLHQ (-17% [95% CI:-28,-4]), as well as a lower TLHQ relative to CEHC ratio (-13% [95% CI: -24, -1]) compare to the lowest quartile group (**Table 2**).

DISCUSSION

For this study, we aimed to investigate the associations between lifestyle factors with serum α-TOH and urinary α-TOH metabolite levels in a cross-sectional analysis in a middle-aged population. After adjusting for possible confounding factors such as BMI, we observed associations between smoking, alcohol use and physical activity, and urinary α-TOH metabolite levels, whereas no associations were observed for serum α-TOH levels. Additionally, the Pearson correlation showed that serum α-TOH and

urinary metabolites do not seem correlated to each other, suggesting both might reflect different biochemical processes. Especially given that observations were independent of BMI or total body fat, this indicated that the relation between certain lifestyle factors and urinary vitamin E metabolites are independent on an adverse adiposity profile. For sleep quality, we found an association with α-TOH, but not with urinary vitamin E metabolite levels. We did not find any association between diet quality and vitamin E levels in serum or urine.

Regarding the association between smoking behaviour and vitamin E levels, a strong association was observed between current smoking and α-CEHC. This is in accordance with multiple other studies, all linking smoking to various measurements of increased oxidant activity [25, 39-41]. Although our data showed that both the oxidative as well as enzymatic conversion of α-TOH were lower in current smokers, our data does suggest that there is relatively more oxidative conversion than enzymatic conversion, which is reflected by higher TLHQ relative to CEHC levels in current smokers. This suggests that there is a higher scavenging effect of oxidized lipids by α-TOH in current smokers. Additionally, we found a dose–response relationship in which current smokers with a higher number of pack years had high α-TLHQ relative to α-CECH and similar results were observed using the objective measure of current smoking cotinine. This supports the possibility of a direct effect of smoking on vitamin E metabolism. Furthermore, this may point to a higher amount of peroxidation, as well as lower antioxidant response capability in smokers than in non-smokers. The absence of this association in former-smokers suggests that smoking cessation reverses this observed difference. However, we acknowledge that our observations are cross-sectional, and longitudinal studies on smoking cessation are required to better understand the directional effects of smoking on the metabolism of vitamin E.

There was no association between diet quality and vitamin E levels. The DHD-index is a validated index and lower scores have been associated with an increased risk of chronic diseases [33]. The DHDI includes intake of the amount of total fat and saturated fat that theoretically can be associated with lipid levels or lipid peroxidation. A possible explanation for the lack of an association with vitamin E (metabolites) is that the DHDI is a reflection of multiple different food components, which might not all be relevant to lipid peroxidation. Consequently, two individuals may have the same DHDI score, but a very different habitual intake of the food components that are relevant for our study. In addition, we also acknowledge the potential measurement error in the questionnaire-derived data.

We did observe an association between poor sleep quality, indicated as a PSQI above 5, and higher α-TOH levels. As sleep deprivation is known to be associated with higher oxidative stress levels, we hypothesized an association between poor habitual sleep quality and α-TOH metabolite levels [42]. Although we found an association between poor sleep quality and higher α-TOH levels, poor sleep quality was not associated with urinary vitamin E metabolite levels. One study has shown that increased vitamin E levels can be associated with a reduction in sleep-deprivation induced problems [43]. However, additional studies are required to fully understand the relation between sleep disturbances and vitamin E (metabolism).

Our finding that alcohol consumption is associated with a lower α-TLHQ is not in line with earlier studies. Lipid peroxidation can be a consequence of ethanol-induced oxidative stress, especially in the brain [44]. A reason for this discrepancy could be that there are other confounding factors that we were not able to control for that are linked to both alcohol consumption and a lower α-TLHQ. Additional studies would be warranted to study this possible relation in more detail.

The observed U-shaped association between physical activity and α-TOH metabolite levels is in accordance with previous literature, as physical activity has shown to be effective at increasing oxidant resistance [26]. When comparing the highest physical activity quartile with the lowest physical activity quartile, the ratio of α-TLHQ relative to α-CHEC level was not different. A potential explanation for this finding could be that during heavy physical activity, the body actually starts activating fat cells, which leads to an increase in lipid peroxidation. In turn, the high physical activity may cancel the beneficial effect of physical activity on the process [26].

One of the strengths of our study is that where most studies only investigated health outcomes in relation to serum α-TOH levels, or studied vitamin E supplementation as an intervention, we also assessed urinary α-TOH metabolites. These metabolites are thought to give a better representation of the enzymatic and non-enzymatic activity as they are the breakdown products of the ROS scavenging process. Additionally, we were able to do these measurements in an adequately sized population of 530 participants.

There are also several limitations of this study that need to be considered. Firstly, although this research was able to provide some novel insights on how α-TOH metabolism is associated with certain lifestyle factors, the cross-sectional design and observational nature of the data do not allow us to determine whether the lifestyle factors preceded the changes in α-TOH metabolite levels. Secondly, α-TOH is only one of many antioxidants present in the human body. The full antioxidant defense system might also change with lifestyle, thus only assessing vitamin E metabolism might not be a complete representation of the functional antioxidant defense system. However, we did try to link the investigated lifestyle factors to lipid peroxidation, of which vitamin E is the main antioxidant response [2]. Lastly, all of the data on lifestyle factors was collected through questionnaires. With the exception of smoking, which we also studied using the metabolite cotinine, the other observations could have been influenced by recall bias and/or measurement error. Additionally, it is generally accepted that some lifestyle factors are frequently underreported, which specifically includes dietary intake and alcohol consumption. However, as we expect the underreporting is relatively independent from the vitamin E levels in both serum and urine, we do not expect this would majorly impact our observations.

Conclusions

In conclusion, our study suggests that, in our study population of relatively healthy middle-aged Dutch participants, the lifestyle factors smoking behavior, physical activity and alcohol consumption were associated with urinary vitamin E metabolites. Moreover, sleep quality was associated with serum α-TOH. These findings specifically highlight that measuring only vitamin E in serum does not provide sufficient insight in the antioxidant vitamin E capacity and activity. Alternatively, this might also explain why targeting antioxidant capacity only by increasing vitamin E levels in serum does not yield significant reduction in disease risk; targeting the conversion of vitamin E might provide additional health benefit beyond serum vitamin E concentrations. However, additional studies are required to provide further evidence of this hypothesis.

Supplementary Materials

The following are available online at www.mdpi.com/xxx/s1, Table S1: Associations between alcohol use and different measurements of vitamin E activity, Table S2: Associations between sleep quality, physical activity and different measurements of vitamin E activity, Table S3: Associations between smoking, diet quality and different measurements of vitamin E activity, Table S4: Associations between smoking intensity in current smokers and different measurements of vitamin E activity.

Author Contributions

Conceptualization, L.G.M., J.L, R.N. and D.v.H.; methodology, L.G.M., J.L., R.N. and D.v.H.; software, L.G.M., J.L. and R.N.; validation, R.N. and J.L.; formal analysis, L.G.M.; investigation, L.G.M., J.L., R.N., F.L.M., E.v.E., R.d.M., K.W.v.D., D.O.M.K., F.R.R. and K.M.; resources, F.L.M., N.A., E.v.E., R.d.M., F.R.R., K.W.v.D., K.M., R.N. and D.v.H.; data curation, J.L., F.L.M., R.d.M., K.W.v.D., D.v.H. and K.M.; writing—original draft preparation, L.G.M., J.L., F.L.M. and R.N.; writing—review and editing, D.O.M.K., E.v.E., R.d.M., F.R.R., K.W.v.D. and K.M.; visualization, L.G.M. and R.N.; supervision, R.N. and D.v.H.; project administration, R.d.M., F.R.R., K.W.v.D., K.M., R.N. and D.v.H.; funding acquisition, J.L., K.M., R.d.M., R.N. and D.v.H. All authors have read and agreed to the published version of the manuscript.

Funding

This work was supported by the VELUX Stiftung [grant number 1156] to D.v.H and R.N., and J.L. was supported by the China Scholarship Counsel [No. 201808500155]. We would like to thank the BRC at Great Ormond Street Hospital, the Alexandra Peto Foundation for their kind donations and the European Union for their funding via the HUMAN consortium. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health, also the kind donations from the Szeban Peto Foundation. The NEO study is supported by the participating Departments, the Division and the Board of Directors of the Leiden University Medical Centre, and by the Leiden University, Research Profile Area 'Vascular and Regenerative Medicine'.

Acknowledgments

We greatly appreciate all participants of the Netherlands Epidemiology of Obesity study, and all participating general practitioners for inviting eligible individuals. We furthermore thank P.R. van Beelen and all research nurses for collecting the data, P.J. Noordijk and her team for sample handling and storage and I. de Jonge, MSc for data management.

Conflicts of Interest

Dennis O Mook-Kanamori is a part-time research consultant at Metabolon, Inc. All other authors declare no conflict of interest.

REFERENCES

1. Luo, J., et al., Ageing, age-related diseases and oxidative stress: What to do next? Ageing Res Rev, 2020. 57: p. 100982.
2. Brigelius-Flohe, R. and M.G. Traber, Vitamin E: function and metabolism. FASEB J, 1999. 13(10): p. 1145-55.
3. Schmolz, L., et al., Complexity of vitamin E metabolism. World J Biol Chem, 2016. 7(1): p. 14-43.
4. Torquato, P., et al., Increased plasma levels of the lipoperoxyl radical-derived vitamin E metabolite alpha-tocopheryl quinone are an early indicator of lipotoxicity in fatty liver subjects. Free Radic Biol Med, 2019. 131: p. 115-125.
5. Wright, M.E., et al., Higher baseline serum concentrations of vitamin E are associated with lower total and cause-specific mortality in the Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study. Am J Clin Nutr, 2006. 84(5): p. 1200-7.
6. Li, F.J., L. Shen, and H.F. Ji, Dietary intakes of vitamin E, vitamin C, and beta-carotene and risk of Alzheimer's disease: a meta-analysis. J Alzheimers Dis, 2012. 31(2): p. 253-8.
7. Rimm, E.B., et al., Vitamin E consumption and the risk of coronary heart disease in men. N Engl J Med, 1993. 328(20): p. 1450-6.
8. Brewer, G.J., Why vitamin E therapy fails for treatment of Alzheimer's disease. J Alzheimers Dis, 2010. 19(1): p. 27-30.
9. Steinhubl, S.R., Why have antioxidants failed in clinical trials? Am J Cardiol, 2008. 101(10A): p. 14D-19D.
10. Wang, L., et al., Vitamin E and C supplementation and risk of cancer in men: posttrial follow-up in the Physicians' Health Study II randomized trial. Am J Clin Nutr, 2014. 100(3): p. 915-23.
11. Klein, E.A., et al., Vitamin E and the risk of prostate cancer: the Selenium and Vitamin E Cancer Prevention Trial (SELECT). JAMA, 2011. 306(14): p. 1549-56.
12. Farina, N., et al., Vitamin E for Alzheimer's dementia and mild cognitive impairment. Cochrane Database Syst Rev, 2017. 4: p. CD002854.
13. Miller, E.R., 3rd, et al., Meta-analysis: high-dosage vitamin E supplementation may increase all-cause mortality. Ann Intern Med, 2005. 142(1): p. 37-46.
14. Bjelakovic, G., et al., Mortality in randomized trials of antioxidant supplements for primary and secondary prevention: systematic review and meta-analysis. JAMA, 2007. 297(8): p. 842-57.
15. Myung, S.K., et al., Efficacy of vitamin and antioxidant supplements in prevention of cardiovascular disease: systematic review and meta-analysis of randomised controlled trials. BMJ, 2013. 346: p. f10.
16. Lebold, K.M., et al., Urinary alpha-carboxyethyl hydroxychroman can be used as a predictor of alpha-tocopherol adequacy, as demonstrated in the Energetics Study. Am J Clin Nutr, 2012. 96(4): p. 801-9.
17. Pate, R.R., et al., Physical activity and public health. A recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. JAMA, 1995. 273(5): p. 402-7.
18. Sofi, F., et al., Adherence to Mediterranean diet and health status: meta-analysis. BMJ, 2008. 337: p. a1344.
19. Ruitenberg, A., et al., Alcohol consumption and risk of dementia: the Rotterdam Study. Lancet, 2002. 359(9303): p. 281-6.
20. Inoue-Choi, M., et al., Association of Long-term, Low-Intensity Smoking With All-Cause and Cause-Specific Mortality in the National Institutes of Health-AARP Diet and Health Study. JAMA Intern Med, 2017. 177(1): p. 87-95.
21. Gottlieb, D.J., et al., Association of usual sleep duration with hypertension: the Sleep Heart Health Study. Sleep, 2006. 29(8): p. 1009-14.
22. Will, J.C., et al., Cigarette smoking and diabetes mellitus: evidence of a positive association from a large prospective cohort study. Int J Epidemiol, 2001. 30(3): p. 540-6.
23. Lee, W., et al., The association between smoking or passive smoking and cardiovascular diseases using a Bayesian hierarchical model: based on the 2008-2013 Korea Community Health Survey. Epidemiol Health, 2017. 39: p. e2017026.
24. Warburton, D.E.R. and S.S.D. Bredin, Health benefits of physical activity: a systematic review of current systematic reviews. Curr Opin Cardiol, 2017. 32(5): p. 541-556.
25. Basyigit, I., et al., Effects of erdosteine on smoking-induced lipid peroxidation in healthy smokers. Drugs R D, 2005. 6(2): p. 83-9.
26. Barranco-Ruiz, Y., et al., A lifelong competitive training practice attenuates age-related lipid peroxidation. J Physiol Biochem, 2017. 73(1): p. 37-48.
27. Beydoun M.A., Gamaldo A.A., Canas J.A., Beydoun H.A., Shah M.T., McNeely J.M., Zonderman A.B. Serum Nutritional Biomarkers and Their Associations with Sleep among US Adults in Recent National Surveys. PLoS ONE. 2014;9:e103490.
28. Evans W.J. Vitamin E, vitamin C, and exercise. Am. J. Clin. Nutr. 2000;72:647–652.
29. Bjorklund, G. and S. Chirumbolo, Role of oxidative stress and antioxidants in daily nutrition and human health. Nutrition, 2017. 33: p. 311-321.
30. de Mutsert, R., et al., The Netherlands Epidemiology of Obesity (NEO) study: study design and data collection. Eur J Epidemiol, 2013. 28(6): p. 513-23.
31. Evans, A., et al., High resolution mass spectrometry improves data quantity and quality as compared to unit mass resolution mass spectrometry in high-throughput profiling metabolomics. 2014. 4(2): p. 1.
32. Rhee, E.P., et al., Variability of Two Metabolomic Platforms in CKD. Clinical Journal of the American Society of Nephrology, 2019. 14(1): p. 40.
33. Looman, M., et al., Development and evaluation of the Dutch Healthy Diet index 2015. Public Health Nutr, 2017. 20(13): p. 2289-2299.
34. Buysse, D.J., et al., The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res, 1989. 28(2): p. 193-213.
35. Wendel-Vos, G.C., et al., Reproducibility and relative validity of the short questionnaire to assess health-enhancing physical activity. J Clin Epidemiol, 2003. 56(12): p. 1163-9.
36. Friedewald W.T., Levy R.I., Fredrickson D.S. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. Clin. Chem. 1972;18:499–502.
37. Sharma, G., et al., Urinary conjugated alpha-tocopheronolactone--a biomarker of oxidative stress in children with type 1 diabetes. Free Radic Biol Med, 2013. 55: p. 54-62.
38. Sharma, G., et al., A novel method for the direct measurement of urinary conjugated metabolites of alpha-tocopherol and its use in diabetes. Mol Nutr Food Res, 2010. 54(5): p. 599-600.
39. Karademirci, M., R. Kutlu, and I. Kilinc, Relationship between smoking and total antioxidant status, total oxidant status, oxidative stress index, vit C, vit E. Clin Respir J, 2018. 12(6): p. 2006-2012.
40. Kosecik, M., et al., Increased oxidative stress in children exposed to passive smoking. Int J Cardiol, 2005. 100(1): p. 61-4.
41. Lourenco, M.A.M., et al., Lipid damage is the best marker of oxidative injury during the cardiac remodeling process induced by tobacco smoke. BMC Pharmacol Toxicol, 2018. 19(1): p. 74.
42. Villafuerte, G., et al., Sleep deprivation and oxidative stress in animal models: a systematic review. Oxid Med Cell Longev, 2015. 2015: p. 234952.
43. Alzoubi, K.H., et al., The neuroprotective effect of vitamin E on chronic sleep deprivation-induced memory impairment: the role of oxidative stress. Behav Brain Res, 2012. 226(1): p. 205-10.
44. Mayas, M.D., et al., Ethanol modifies differently aspartyl- and glutamyl-aminopeptidase activities in mouse frontal cortex synaptosomes. Brain Res Bull, 2002. 57(2): p. 195-203.

3

DIET-DERIVED ANTIOXIDANTS DO NOT DECREASE RISK OF ISCHEMIC STROKE: A MENDELIAN RANDOMIZATION STUDY IN 1 MILLION PEOPLE

Leon G Martens
Jiao Luo
Ko Willems van Dijk
J Wouter Jukema
Raymond Noordam
Diana van Heemst

Journal of the American Heart Association, 2021 Dec 7;20(23):e022567

ABSTRACT

Background and aims

Dietary intake and blood concentrations of vitamin E, C, lycopene and carotenoids have been associated with a lower risk of incident (ischemic) stroke. However, causality cannot not be inferred from these associations. Here, we investigated causality by analyzing the associations between genetically-influenced antioxidant levels in blood and ischemic stroke using Mendelian Randomization (MR).

Methods

For each circulating antioxidant (vitamin E, C, lycopene, β-carotene and retinol), which were assessed as either absolute blood levels and/or high-throughput metabolite levels, independent genetic instrumental variables were selected from earlier genome wide association studies ($p < 5x10^{-8}$). We used summary statistics for Single-Nucleotide Polymorphisms (SNP)-stroke associations from three European-ancestry cohorts (cases/controls): MEGASTROKE (60,341/454,450), UK Biobank (2,404/368,771) and FinnGen study (8,046/164,286). MR analyses were performed on each exposure per outcome cohort using inverse-variance weighted analyses, and subsequently meta-analyzed.

Results

In a combined sample of 1,058,298 individuals (70,791 cases), none of the genetically-influenced absolute antioxidants or antioxidant metabolite concentrations were causally associated with a lower risk of ischemic stroke. For absolute antioxidants levels, the odds ratios (95% CI) ranged between 0.94 (95% CI: 0.85 to 1.05) for vitamin C and 1.04 (95% CI: 0.99 to 1.08) for lycopene. For metabolites, odds ratios ranged between 1.01 (95% CI: 0.98 to 1.03) for retinol and 1.12 (95% CI: 0.88 to 1.42) for vitamin E.

Conclusions

This study did not provide evidence for a causal association between dietary-derived antioxidant levels and ischemic stroke. Therefore, antioxidant supplements to increase circulating levels are unlikely to be of clinical benefit to prevent ischemic stroke.

INTRODUCTION

Stroke is the second leading cause of death and loss of disability-adjusted life years worldwide [1]. Ischemic stroke is caused by a disruption of cerebral blood flow causing a lack of oxygen in the affected area [2]. Several classical risk factors have been described as important in the pathogenesis of ischemic stroke, including smoking, obesity, diabetes mellitus, hypertension and dyslipidemia [3-7]. In addition to the traditional risk factors, oxidative stress has been hypothesized to be a vital trigger in the occurrence of stroke via excessive production of Reactive Oxygen Species (ROS) [8, 9]. ROS-induced damage can cause significant changes in the vascular system, ultimately influencing cerebral blood flow [9]. These detrimental effects, including increasing vasodilation, platelet aggregation, increased endothelial permeability and the formation of local lesions [9], could consequently lead to an increased risk of stroke [10].

Therefore, antioxidants, which are scavengers of free radicals and thereby diminish oxidative damage, can be hypothesized to decrease the risk of disease occurrence. Antioxidants, such as vitamin E, C and carotenoids, are of specific interest given that they are accessible and their intake is easily modifiable. Several studies have already been conducted examining the association between antioxidants and the occurrence of stroke [8, 10-14]. In these studies, dietary intake, either as dietary components or supplements, or blood concentration of vitamin E, C and carotenoids were associated with lower risk of first occurrence of ischemic stroke [8, 11-14]. Similarly, adherence to a diet rich in antioxidants, irrespective of the type of antioxidants, has been associated with a lower risk of incident stroke [15].

However, associations in these observational studies inevitably suffer from the possibility of reverse causation and residual confounding, and should be interpreted with caution. Therefore, the causality between dietary-derived antioxidants and stroke is still unclear. A randomized controlled trial on vitamin E and β-carotene supplementation and stroke risk found inconsistent results [16]. Whereas a similar study on vitamin E supplementation and cardiovascular events found no effect [17]. In addition to randomized clinical trials, Mendelian Randomization (MR), in which genetic variants of a certain exposure are used as instrumental variables, is an alternative approach to infer causality of life-long risk factors (exposure) on diseases (outcome) [18-20]. In the present study, we used two-sample MR to assess the causal associations between genetically-influenced dietary-derived circulating antioxidants and their metabolites with ischemic stroke.

METHODS

The data on which the results were based were primarily summary-level data. The data from the UK Biobank is available upon acceptance of a research proposal by UK Biobank Recourses and payment of an access fee. The data from MEGASTROKE and FinnGen are freely available from their respective websites (MEGASTROKE: https://www. https://www.megastroke.org/, FinnGen: https://www.finngen.fi/en/).

Study Design

For our study, we conducted a two-sample MR, which uses genetic instrumental variables as a proxy for the exposure [21]. Using this design, SNP-exposure and equivalent SNP-outcome associations were derived from genome-wide association studies (GWAS). Subsequently, the SNP-outcome effect is divided by the SNP-exposure effect, and these are subsequently meta-analyzed to approximate the average genetically-influenced effect on the outcome. MR is based on several assumptions. First, the genetic variants should be associated with the exposure. Secondly, the association between these genetic variants and the outcome should be solely through the exposure. Finally, the genetic variants should be independent of any measured and unmeasured confounders.

Selection of genetic instrumental variables

To study the association between genetically-influenced levels of antioxidants and stroke, the genetic instruments of five diet-derived antioxidants were used, which included vitamin E (α-tocopherol and γ-tocopherol), β-carotene, lycopene, vitamin C (L-ascorbic acid or ascorbate) and retinol. For these antioxidants, we investigated either absolute blood levels, circulating metabolites which were quantified as relative concentrations in plasma or serum using high-throughput commercial platforms, or both. The metabolites were quantified as relative concentrations [22]. For each antioxidant, a search for GWAS was performed to extract the leading SNPs as genetic instrumental variables. Whenever multiple GWAS were identified, only the largest study including replication was used [23-29]. All selected variants were associated with their corresponding exposure ($p < 5 \times 10^{-8}$). A more detailed description and a summary table of the SNPs for each antioxidant used as instrumental variables can be found in the supplements and in our previous publication [22]. Different from our earlier work [22], we now used an updated genome-wide association study on vitamin C (all genetic instruments presented in **Supplementary Table 1**) [29]. All selected variants passed standard quality control filters with an imputation score of 0.90 or higher, except for retinol-associated SNP rs117468033, which passed with a score of 0.65.

Outcome datasets

Summary statistics on the associations of the exposure-related SNPs with ischemic stroke were extracted from three large cohorts; the MEGASTROKE consortium, the UK Biobank, and the FinnGen study. Both UK Biobank and FinnGen were not part of the main analyses of the MEGASTROKE consortium preventing inclusion of overlapping samples in the analyses. Available summary statistics can be found in **Table 1**.

Table 1. Summary characteristics MEGASTROKE, UK Biobank & FinnGen

	MEGASTROKE	UK Biobank	FinnGen
Cases	67 162	2 404	8 046
Age (SD), years	69.1 (10.8)	61.5 (6.7)	-
Women, %	43%	34%	-
Controls	454 450	368 771	164 286
Age (SD), years	56.6 (15.6)	56.7 (8.0)	-
Women, %	51%	54.5%	-

The MEGASTROKE consortium consisted of 60,341 cases and 454,450 controls collected from 29 studies. Of these participants, 86% were of European ancestry, 9% of East Asian ancestry, and the remaining participants were from African, South Asian, mixed Asian, and Latin American ancestry [30].

The UK Biobank cohort (project application number 56340) is a prospective general population cohort with 502,628 participants between the age of 40 and 70 years recruited from the general population between 2006 and 2010 [31], and more information can be found online (https://www.ukbiobank.ac.uk). The analyses were performed with participants of European ancestry, who were in the full released imputed genomics databases (UK10K + HRC). Follow-up information, including stroke occurrence, was retrieved through routinely available national datasets. In our dataset, we had data available on 2,404 cases of ischemic stroke, and 368,771 controls. We performed new genome-wide association analyses using logistic regression to assess the associations between genetic instruments and ischemic stroke, adjusted for age, sex and 10 principal components, and corrected for familial relationships using BOLT_LMM (v2.3.2).

The FinnGen study is an ongoing cohort study launched in 2017, which integrated the genetic data generated from biobank samples and health-related data from social and healthcare registers. Detailed information such as participating biobanks/cohorts, genotyping and data analysis are available at their website (https://www.finngen.fi/en/). For our current study, the freeze 4 data was used. Within this data, there are 8,046 reported cases of ischemic stroke, and 164,286 controls.

Statistical analysis

All the analyses were done using R (v3.6.1) statistical software (The R Foundation for Statistical Computing, Vienna, Austria). MR analyses were performed using the R-based

package “TwoSampleMR” (https://mrcieu.github.io/TwoSampleMR/).

Prior to the analysis, the “TwoSampleMR” package was used for harmonization of the exposure and outcome SNPs, in order to verify that the effect estimates are aligned in the same direction. Additionally, harmonization removes palindromic SNPs from the instrument pool.

For our primary MR analysis, Inverse-Variance weighted (IVW) regression analyses were performed. This method assumes the absence of invalid genetic instruments such as SNPs affecting multiple exposures (pleiotropy) causing possible directional pleiotropy [19]. First, causal estimates were calculated per genetic instrument using the Wald ratio (SNP – outcome association divided by the SNP – exposure association) and subsequently meta-analyzed using the inverse-weighted meta-analyses weighted on the standard error of the SNP-outcome association (assuming no measurement error in the exposure) [32]. The calculated estimates were expressed as odds ratios (OR) on ischemic stroke per unit difference of the corresponding absolute circulating antioxidant levels (natural log-transformed levels for β-carotene and retinol, µg/dL for lycopene or µmol/L for ascorbate) or 10-fold change in antioxidant metabolite concentrations.

To ensure that the results obtained from the IVW analyses were not biased due to directional pleiotropy, we performed MR-Egger regression analysis and Weighted-Median Estimator when applicable [32]. In MR-Egger, the intercept depicts the estimated average pleiotropic effect across the genetic variants, and a value that differs from zero indicates that the IVW estimate is biased [33]. Although considered as a relatively inefficient approach (e.g., large confidence intervals), this method does not force the regression line to go through the intercept. The Weighted-Median estimator analysis can provide a consistent valid estimate if at least half of the instrumental variables are valid [20]. In addition, MR-PRESSO (MR Pleiotropy RESidual Sum and Outlier) was applied when possible to detect and correct for horizontal pleiotropy through removing outliers [34], as implemented in the R-based package MR-PRESSO (https://github.com/rondolab/MR-PRESSO)”. The Cochran’s Q statistic was performed in order to test the heterogeneity between the estimated Wald ratios from different genetic variants [35]. Additional sensitivity analyses were performed for the antioxidants β-carotene and lycopene.

The main MR analyses were performed in the individual datasets, and subsequently meta-analyzed to derive the pooled estimates for each exposure on the risk of ischemic stroke. Heterogeneity of the estimates across three datasets was performed by I^2, and corresponding p-value was obtained from the Cochran’s Q test. When no heterogeneity was found amongst the three cohorts, a fixed-effect model meta-analysis was used to pool the instrumental variable estimates for each exposure. All meta-analyses were performed in the R-based “meta” package (https://cran.r-project.org/web/packages/meta/index.html).

A power calculation was performed for each genetically derived antioxidant separately (https://shiny.cnsgenomics.com/mRnd/). With power = 0.80, minimal effect sizes (OR) differed from 0.98 to 0.93, which we deemed reasonable given the results of observational studies [11-14]. We therefore concluded our study had enough power given the current parameters. Separate power calculations were done for the different subtypes of stroke (cardioembolic, large artery atherosclerosis and small-vessel) in MEGASTROKE, as data was available. Minimal effect sizes needed ranged from 0.94 to 0.78. Although some of these were considered as an unreasonably achievable estimate, we decided to perform these sub-analyses as well.

RESULTS

By combining the three cohorts, a total sample of 1,058,298 individuals, of which 70,791 cases of ischemic stroke, and 987,507 controls, were analyzed to assess the association between diet-derived antioxidants and ischemic stroke. Variance explained (R^2) by the instruments for each trait were either derived from the original study or calculated based on the derived summary statistics and in line with the method described previously [36], and ranged from 1.7% to 30.1% for absolute antioxidant levels, and from 3.3% to 18.6% for metabolite antioxidants (**Table 2**). In order to minimize potential weak instrument bias, we considered an F-statistic of at least 10 as sufficient for performing an MR analysis, which is well-accepted in the field [37]. The in-between SNP heterogeneity was non-significant for all antioxidants in each cohort ($p > 0.05$). Additionally, we found no heterogeneity in the summary estimates from the MR analyses between the included datasets (**Figure 1 & Figure 2**). All analyses were performed with outcome ischemic stroke, as well as with outcomes cardioembolic, large artery atherosclerosis, and small-vessel stroke separately (**Supplementary Figure 1 & Figure 2**). As no consistent strong evidence favoring the hypothesis of higher levels resulting in a lower risk of stroke subtypes was found, only the main results are given here.

Table 2. GWAS on genetically determined diet derived antioxidants

	Absolute antioxidants				Metabolite antioxidants			
Exposure Datasets	**N =**	**Genetic Variants**	**Explained Variance**	**Unit**	**N=**	**Genetic Variants**	**Explained Variance**	**Units**
a-Tocopherol	-	-	-	-	7276	11	3.3%	log10-transformed metabolites concentration
γ-Tocopherol	-	-	-	-	5822	13	15.0%	log10-transformed metabolites concentration

Table 2. Continued

	Absolute antioxidants				Metabolite antioxidants			
Exposure Datasets	**N =**	**Genetic Variants**	**Explained Variance**	**Unit**	**N=**	**Genetic Variants**	**Explained Variance**	**Units**
β-Caroten	2344	3	9.0%	µg/L in log-transformed scale	-	-	-	-
Lycopene	441	5	30.1%	µg/dL	-	-	-	-
Ascorbate	52018	11	1.87%	µmol/L	2063	13	18.6%	log10-transformed metabolites concentration
Retinol	5006	2	2.3%	µg/L in log-transformed scale	1957	24	4.8%	log10-transformed metabolites concentration

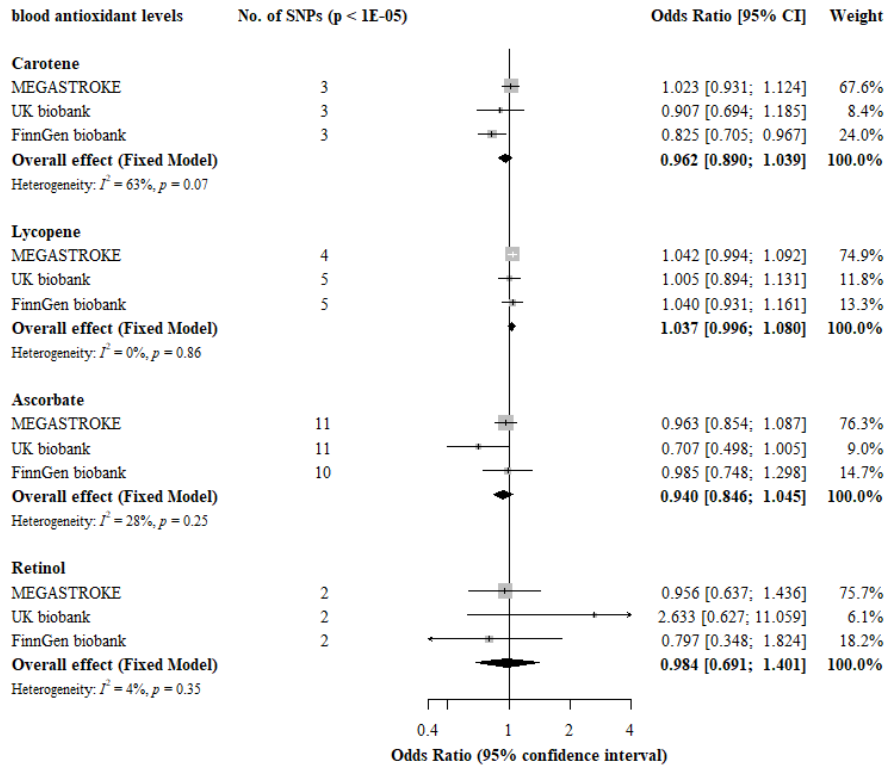

Figure 1. Causal association between absolute blood level antioxidants and ischemic stroke occurrence. Estimated ORs represent the effect per unit increase in ln-transformed β-Carotene and retinol, 1 µg/dL lycopene and 1 µmol/L ascorbate on ischemic stroke. Results were obtained from an IVW analysis per outcome database and combined over the three databases using fixed-effect meta-analyses.

CI = confidence interval; OR = odds ratio;

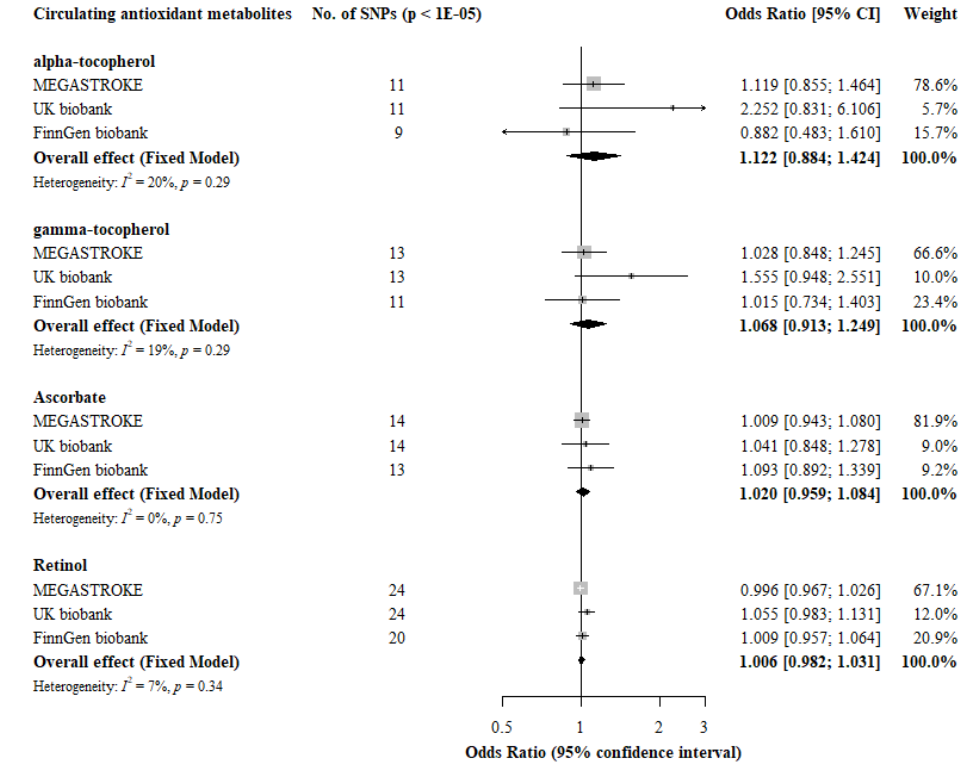

Figure 2. Causal association between circulating antioxidant metabolites and ischemic stroke occurrence. Estimated ORs represent the effect per 10-fold increase in antioxidant metabolites' concentration on ischemic stroke. Results were obtained from an IVW analysis per outcome database and combined over the three databases using fixed-effect meta-analyses.

CI = confidence interval; OR = odds ratio;

Absolute antioxidant levels

For absolute blood antioxidants levels (**Figure 1**), we observed no evidence for associations with ischemic stroke, except for β-carotene when analyzed only within the FinnGen cohort. However, the pooled estimate for all three cohorts was not significantly associated with ischemic stroke. Most notably, the pooled odds ratios (95% confidence intervals) were 0.94 (95% CI: 0.85, 1.05) per 1 µmol/L ascorbate, 1.04 (95% CI: 0.99, 1.08) per 1 µg/dL lycopene, 0.96 (95% CI: 0.89, 1.04) and 0.98 (95% CI: 0.69, 1.40) per natural log-transformed β-carotene and retinol, respectively.

MR-Egger and weighted-median estimator regression were performed for antioxidants with more than three genetic instruments (notably β-carotene and lycopene). The estimates of both MR-Egger and weighted-median estimator were comparable with the IVW analyses. Furthermore, the MR-Egger intercept did not

deviate from zero (p-values>0.05). Additionally, MR-PRESSO did not detect any outliers, and Cochran's Q statistics detected no heterogeneity for the analyses of β-carotene or lycopene on ischemic stroke.

Circulating antioxidant metabolite levels

For circulating antioxidant metabolites (**Figure 2**), the pooled OR (95%CI) for ischemic stroke per 10-fold increase in metabolite concentration were 1.12 (95% CI: 0.88, 1.42) for α-tocopherol, 1.07 (95% CI: 0.91, 1.25) for γ-tocopherol, 1.02 (95% CI: 0.96, 1.08) for ascorbate and 1.01 (95% CI: 0.98, 1.03) for retinol using IVW.

Estimates derived from the MR-Egger and weighted-median estimator analyses were of similar direction and magnitude as the IVW analyses. Furthermore, no pleiotropic effect was identified by the intercept from MR-Egger or MR PRESSO, and no potential outlier was found via MR PRESSO. Cochran's Q statistics only detected heterogeneity in γ-tocopherol in the MEGASTROKE cohort.

DISCUSSION

We investigated whether we could find evidence supporting causality of the association between diet-derived antioxidants and ischemic stroke risk using MR. Circulating antioxidants, irrespective of how the levels were determined, were proxied by using genetic variants as instrumental variables. In an extreme sample size of 1,058,298 participants, 70,791 cases and 987,507 controls, we did not find evidence that genetically-influenced diet-derived antioxidant levels were associated with a lower risk of developing ischemic stroke. While B-carotene was associated with ischemic stroke in the FinnGen population, this association was not observed in the other included study populations and was therefore considered as a likely chance finding. These findings suggest that the previously observed association between antioxidants (either by dietary intake and/or serum levels) and ischemic stroke is not causal.

Previously, a meta-analysis of observational studies identified a 17% risk reduction of stroke in participants with high dietary vitamin E intake compared with those with low intake. However, the authors were cautious with interpreting this result given that only 3,284 events of ischemic stroke were reported, despite the study comprising over 220,000 participants, as well as the large heterogeneity of the individual contributing studies in the meta-analysis [38]. To date, there are no MR studies that have assessed the causal association of circulating antioxidants and ischemic stroke risk. In our previous study, we demonstrated that the effect of genetic variants on circulating antioxidant levels are generally comparable with those which would be achieved by dietary supplementation [22]. Given the robust and generally consistent null results in the present study that both absolute blood antioxidant levels and metabolites measured with high throughput technology are not causally associated with ischemic stroke, it is not likely that dietary supplements that are to increase antioxidant concentrations in blood reduce the risk of ischemic stroke.

Despite the use of data from more than 1 million participants, our study population could be seen as relatively young for stroke occurrence (especially in the UK Biobank population who were below age 70 years at study inclusion). In a population-based cohort study, stroke incidence increased from 1.7 at age 55-59, to 19.9 at age 80-84 [39]. Thus, the current population might show different results when studied several years from now. The low number of cases observed in the UK Biobank population is likely explained by the lower mean age of the individuals included in the analysis, and the UK Biobank is considered a healthier population than the general UK population. Furthermore, MEGASTROKE contained cohorts with an oversampling of stroke cases. As our primary research hypothesis was focused on ischemic stroke, we did not investigate different stroke subtypes as primary outcomes. This decision was motivated by the up to 6 times lower number of cases per subtype. Therefore, analyses of these stroke subtypes were considered as underpowered to draw firm conclusions from, as could also be seen from our power calculation. However, importantly, our findings on ischemic stroke were robust across different analysis methods and were generally consistent across three independent cohorts with a large number of cases. As our null findings were generally consistent across our study cohorts, it is very unlikely that changes in antioxidant levels will yield any clinically relevant reduction in ischemic stroke risk. These findings were also in line with our earlier work in which we found no associations between these antioxidants in relation to coronary heart disease [22]. Together with our present findings, this would suggest that antioxidants do not affect the risk of developing atherogenic cardiovascular disease. However, this is contrary to the data from experimental settings in which oxidative stress does play an important role in the onset of atherosclerosis [40, 41]. This has given rise to the hypothesis that circulating antioxidant levels might not be representative of antioxidant capacity, and that increasing antioxidant levels in blood (either by nutritional intake or supplements) do not necessarily result in additional antioxidative effects. This hypothesis was supported by a number of our earlier studies in which we investigated the associations of vitamin E and its enzymatic and oxidative metabolites with lifestyle factors and subclinical disease outcomes [42, 43]. In brief, we showed in these studies that vitamin E concentrations were not correlated with the urinary enzymatic and oxidative vitamin E metabolite levels, and lifestyle factors and subclinical disease outcomes showed different associations with vitamin E concentrations than with their oxidized metabolites.

The present study has several strengths. First, a large sample size was studied, by combining three cohorts comprising a total of 70,791 cases and performing a meta-analysis. Individually, the results from the three cohorts are consistent with each other and with the final meta-analysis. We did not detect any heterogeneity between-SNPs

for every antioxidant in each cohort, or across each cohort. Additionally, by performing MR-egger and weighted-median estimator analyses, the final MR estimates should be seen as a reliable result despite the sometimes low explained variance of certain genetic instruments. Second, we used separate sets of instrumental variables by looking at both absolute blood levels as well as metabolite levels of antioxidants. The similar findings, especially in regards to ascorbate and retinol which are analyzed with both their absolute blood concentration and the relative metabolite levels, suggests general robustness of our findings. However, some limitations should also be considered with respect to the interpretation of the results. First, participants included in our study are predominantly of European descent, which will limit the extrapolation to other populations. Second, sensitivity analyses for some instrumental variables with limited number of genetic variants as instrumental variables could not be performed. Third, since we were only able to study a selected number of antioxidants, we cannot fully exclude the hypothesis that other antioxidants could have protective effects on ischemic stroke.

In summary, our study did not provide evidence supporting a causal association between diet-derived levels of antioxidants vitamin E, C, lycopene, β-carotene and retinol, and ischemic stroke. Therefore, antioxidant supplementation is unlikely to be of clinical benefit to prevent ischemic stroke.

Acknowledgements

The authors are grateful to the METASTROKE consortium and the FinnGen Biobank for making their summary statistics available for further studies. The analyses done in UK Biobank were done under project number 56340.

LGM, JL, RN and DvH designed research; LGM and JL conducted research; LGM and JL performed statistical analysis; LGM and JL, RN wrote paper; LGM had primary responsibility for final content. DvH, KWvD and JWJ contributed to the data interpretation and commented on initial versions of the manuscript; All authors read and approved the final manuscript.

Funding

This work was supported by the VELUX Stiftung [grant number 1156] to DvH and RN, and JL was supported by the China Scholarship Counsel [No. 201808500155]. RN was supported by an innovation grant from the Dutch Heart Foundation [grant number 2019T103 to R.N.].

Disclosures

None

Conflict of interest statement

The authors declare to have no conflict of interest.

Supplemental Material

Data S1. Supplemental Methods, Table S1, Figures S1-S2.

REFERENCES

1. Wang, H., *Global age-sex-specific fertility, mortality, healthy life expectancy (HALE), and population estimates in 204 countries and territories, 1950-2019: a comprehensive demographic analysis for the Global Burden of Disease Study 2019.* Lancet, 2020. **396**: p. 1160-1203.
2. Boehme, A.K., C. Esenwa, and M.S. Elkind, *Stroke Risk Factors, Genetics, and Prevention.* Circ Res, 2017. **120**: p. 472-495.
3. Banerjee, C., Y.P. Moon, M.C. Paik, T. Rundek, C. Mora-McLaughlin, J.R. Vieira, R.L. Sacco, and M.S. Elkind, *Duration of diabetes and risk of ischemic stroke: the Northern Manhattan Study.* Stroke, 2012. **43**: p. 1212-7.
4. Bhat, V.M., J.W. Cole, J.D. Sorkin, M.A. Wozniak, A.M. Malarcher, W.H. Giles, B.J. Stern, and S.J. Kittner, *Dose-response relationship between cigarette smoking and risk of ischemic stroke in young women.* Stroke, 2008. **39**: p. 2439-43.
5. Horenstein, R.B., D.E. Smith, and L. Mosca, *Cholesterol predicts stroke mortality in the Women's Pooling Project.* Stroke, 2002. **33**: p. 1863-8.
6. Lewington, S., R. Clarke, N. Qizilbash, R. Peto, R. Collins, and C. Prospective Studies, *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**: p. 1903-13.
7. Suk, S.H., R.L. Sacco, B. Boden-Albala, J.F. Cheun, J.G. Pittman, M.S. Elkind, M.C. Paik, and S. Northern Manhattan Stroke, *Abdominal obesity and risk of ischemic stroke: the Northern Manhattan Stroke Study.* Stroke, 2003. **34**: p. 1586-92.
8. Rodrigo, R., R. Fernandez-Gajardo, R. Gutierrez, J.M. Matamala, R. Carrasco, A. Miranda-Merchak, and W. Feuerhake, *Oxidative stress and pathophysiology of ischemic stroke: novel therapeutic opportunities.* CNS Neurol Disord Drug Targets, 2013. **12**: p. 698-714.
9. Allen, C.L. and U. Bayraktutan, *Oxidative stress and its role in the pathogenesis of ischaemic stroke.* Int J Stroke, 2009. **4**: p. 461-70.
10. Shirley, R., E.N. Ord, and L.M. Work, *Oxidative Stress and the Use of Antioxidants in Stroke.* Antioxidants (Basel), 2014. **3**: p. 472-501.
11. Yochum, L.A., A.R. Folsom, and L.H. Kushi, *Intake of antioxidant vitamins and risk of death from stroke in postmenopausal women.* Am J Clin Nutr, 2000. **72**: p. 476-83.
12. Liu, S., J.E. Manson, M.J. Stampfer, K.M. Rexrode, F.B. Hu, E.B. Rimm, and W.C. Willett, *Whole grain consumption and risk of ischemic stroke in women: A prospective study.* JAMA, 2000. **284**: p. 1534-40.
13. Sanchez-Moreno, C., A. Jimenez-Escrig, and A. Martin, *Stroke: roles of B vitamins, homocysteine and antioxidants.* Nutr Res Rev, 2009. **22**: p. 49-67.
14. Gillman, M.W., L.A. Cupples, D. Gagnon, B.M. Posner, R.C. Ellison, W.P. Castelli, and P.A. Wolf, *Protective effect of fruits and vegetables on development of stroke in men.* JAMA, 1995. **273**: p. 1113-7.
15. Parohan, M., J. Anjom-Shoae, M. Nasiri, M. Khodadost, S.R. Khatibi, and O. Sadeghi, *Dietary total antioxidant capacity and mortality from all causes, cardiovascular disease and cancer: a systematic review and dose-response meta-analysis of prospective cohort studies.* Eur J Nutr, 2019. **58**: p. 2175-2189.
16. Leppala, J.M., J. Virtamo, R. Fogelholm, D. Albanes, P.R. Taylor, and O.P. Heinonen, *Vitamin E and beta carotene supplementation in high risk for stroke: a subgroup analysis of the Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study.* Arch Neurol, 2000. **57**: p. 1503-9.
17. Yusuf, S., G. Dagenais, J. Pogue, J. Bosch, and P. Sleight, *Vitamin E supplementation and cardiovascular events in high-risk patients.* N Engl J Med, 2000. **342**: p. 154-60.
18. Lawlor, D.A., R.M. Harbord, J.A. Sterne, N. Timpson, and G. Davey Smith, *Mendelian randomization: using genes as instruments for making causal inferences in epidemiology.* Stat Med, 2008. **27**: p. 1133-63.
19. Burgess, S., A. Butterworth, and S.G. Thompson, *Mendelian randomization analysis with multiple genetic variants using summarized data.* Genet Epidemiol, 2013. **37**: p. 658-65.
20. Bowden, J., G. Davey Smith, P.C. Haycock, and S. Burgess, *Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator.* Genet Epidemiol, 2016. **40**: p. 304-14.
21. Burgess, S., R.A. Scott, N.J. Timpson, G. Davey Smith, S.G. Thompson, and E.-I. Consortium, *Using published data in Mendelian randomization: a blueprint for efficient identification of causal risk factors.* Eur J Epidemiol, 2015. **30**: p. 543-52.
22. Luo, J., S. le Cessie, D. van Heemst, and R. Noordam, *Diet-Derived Circulating Antioxidants and Risk of Coronary Heart Disease: A Mendelian Randomization Study.* J Am Coll Cardiol, 2021. **77**: p. 45-54.
23. Long, T., M. Hicks, H.C. Yu, W.H. Biggs, E.F. Kirkness, C. Menni, J. Zierer, K.S. Small, M. Mangino, H. Messier, et al., *Whole-genome sequencing identifies common-to-rare variants associated with human blood metabolites.* Nat Genet, 2017. **49**: p. 568-578.
24. D'Adamo, C.R., A. D'Urso, K.A. Ryan, L.M. Yerges-Armstrong, R.D. Semba, N.I. Steinle, B.D. Mitchell, A.R. Shuldiner, and P.F. McArdle, *A Common Variant in the SETD7 Gene Predicts Serum Lycopene Concentrations.* Nutrients, 2016. **8**: p. 82.
25. Shin, S.Y., E.B. Fauman, A.K. Petersen, J. Krumsiek, R. Santos, J. Huang, M. Arnold, I. Erte, V. Forgetta, T.P. Yang, et al., *An atlas of genetic influences on human blood metabolites.* Nat Genet, 2014. **46**: p. 543-550.
26. Hendrickson, S.J., A. Hazra, C. Chen, A.H. Eliassen, P. Kraft, B.A. Rosner, and W.C. Willett, *beta-Carotene 15,15'-monooxygenase 1 single nucleotide polymorphisms in relation to plasma carotenoid and retinol concentrations in women of European descent.* Am J Clin Nutr, 2012. **96**: p. 1379-89.
27. Mondul, A.M., K. Yu, W. Wheeler, H. Zhang, S.J. Weinstein, J.M. Major, M.C. Cornelis, S. Mannisto, A. Hazra, A.W. Hsing, et al., *Genome-wide association study of circulating retinol levels.* Hum Mol Genet, 2011. **20**: p. 4724-31.
28. Major, J.M., K. Yu, W. Wheeler, H. Zhang, M.C. Cornelis, M.E. Wright, M. Yeager, K. Snyder, S.J. Weinstein, A. Mondul, et al., *Genome-wide association study identifies common variants associated with circulating vitamin E levels.* Hum Mol Genet, 2011. **20**: p. 3876-83.
29. Zheng, J.S., J. Luan, E. Sofianopoulou, F. Imamura, I.D. Stewart, F.R. Day, M. Pietzner, E. Wheeler, L.A. Lotta, T.E. Gundersen, et al., *Plasma Vitamin C and Type 2 Diabetes: Genome-Wide Association Study and Mendelian Randomization Analysis in European Populations.* Diabetes Care, 2021. **44**: p. 98-106.
30. Malik, R., G. Chauhan, M. Traylor, M. Sargurupremraj, Y. Okada, A. Mishra, L. Rutten-Jacobs, A.K. Giese, S.W. van der Laan, S. Gretarsdottir, et al., *Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes.* Nat Genet, 2018. **50**: p. 524-537.
31. Sudlow, C., J. Gallacher, N. Allen, V. Beral, P. Burton, J. Danesh, P. Downey, P. Elliott, J. Green, M. Landray, et al., *UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age.* PLoS Med, 2015. **12**: p. e1001779.
32. Burgess, S. and S.G. Thompson, *Interpreting findings from Mendelian randomization using the MR-Egger method.* Eur J Epidemiol, 2017. **32**: p. 377-389.
33. Bowden, J., G. Davey Smith, and S. Burgess, *Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression.* Int J Epidemiol, 2015. **44**: p. 512-25.
34. Verbanck, M., C.Y. Chen, B. Neale, and R. Do, *Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases.* Nat Genet, 2018. **50**: p. 693-698.
35. Greco, M.F., C. Minelli, N.A. Sheehan, and J.R. Thompson, *Detecting pleiotropy in Mendelian randomisation studies with summary data and a continuous outcome.* Stat Med, 2015. **34**: p. 2926-40.
36. Teslovich, T.M., K. Musunuru, A.V. Smith, A.C. Edmondson, I.M. Stylianou, M. Koseki, J.P. Pirruccello, S. Ripatti, D.I. Chasman, C.J. Willer, et al., *Biological, clinical and population relevance of 95 loci for blood lipids.* Nature, 2010. **466**: p. 707-13.
37. Pierce, B.L., H. Ahsan, and T.J. Vanderweele, *Power and instrument strength requirements for Mendelian randomization studies using multiple genetic variants.* Int J Epidemiol, 2011. **40**: p. 740-52.
38. Cheng, P., L. Wang, S. Ning, Z. Liu, H. Lin, S. Chen, and J. Zhu, *Vitamin E intake and risk of stroke: a meta-analysis.* Br J Nutr, 2018. **120**: p. 1181-1188.
39. Hollander, M., P.J. Koudstaal, M.L. Bots, D.E. Grobbee, A. Hofman, and M.M. Breteler, *Incidence, risk, and case fatality of first ever stroke in the elderly population. The Rotterdam Study.* J Neurol Neurosurg Psychiatry, 2003. **74**: p. 317-21.
40. Kattoor, A.J., N.V.K. Pothineni, D. Palagiri, and J.L. Mehta, *Oxidative Stress in Atherosclerosis.* Curr Atheroscler Rep, 2017. **19**: p. 42.
41. Forstermann, U., N. Xia, and H. Li, *Roles of Vascular Oxidative Stress and Nitric Oxide in the Pathogenesis of Atherosclerosis.* Circ Res, 2017. **120**: p. 713-735.
42. Martens, L.G., J. Luo, F.L. Meulmeester, N. Ashrafi, E.W. van Eekelen, R. de Mutsert, D.O. Mook-Kanamori, F.R. Rosendaal, K.W. van Dijk, K. Mills, et al., *Associations between Lifestyle Factors and Vitamin E Metabolites in the General Population.* Antioxidants (Basel), 2020. **9**.

43. Luo, J., F.L. Meulmeester, L.G. Martens, N. Ashrafi, R. de Mutsert, D.O. Mook-Kanamori, F.R. Rosendaal, K. Willems van Dijk, S. le Cessie, K. Mills, et al., *Urinary oxidized, but not enzymatic vitamin E metabolites are inversely associated with measures of glucose homeostasis in middle-aged healthy individuals.* Clin Nutr, 2021. **40**: p. 4192-4200.

4

THE ASSOCIATION BETWEEN MITOCHONDRIAL DNA ABUNDANCE AND STROKE: A COMBINATION OF MULTIVARIABLE-ADJUSTED SURVIVAL AND MENDELIAN RANDOMIZATION ANALYSES.

Leon G Martens
Jiao Luo
Marieke J H Wermer
Ko Willems van Dijk
Sara Hägg
Felix Grassmann
Raymond Noordam
Diana van Heemst

Atherosclerosis, 2022 Jun 19;354:1-7

ABSTRACT

Background and aims

Mitochondrial dysfunction is associated with increased Reactive Oxygen Species that are thought to drive disease risk, including stroke. We investigated the association between mtDNA abundance, as a proxy measure of mitochondrial function, and incident stroke using multivariable-adjusted survival and Mendelian Randomization (MR) analyses.

Methods

Cox-proportional hazard model analyses were conducted to assess the association between mtDNA abundance, and incident ischemic and hemorrhagic stroke over a maximum of 14-years follow-up in European-ancestry participants from UK Biobank. MR was conducted using independent (R^2<0.001) lead variants for mtDNA abundance (p < 5x10-8) as instrumental variables. Single-Nucleotide Polymorphism (SNP)-ischemic stroke associations were derived from three published open source European-ancestry results databases (cases/controls): MEGASTROKE (60,341/454,450), UK Biobank (2,404/368,771) and FinnGen (10,551/202,223). MR was performed per study, and results were subsequently meta-analyzed.

Results

In total, 288,572 unrelated participants (46% men) with mean (SD) age of 57 (8) years were included in the Cox-proportional hazard analyses. After correction for considered confounders (BMI, hypertension, cholesterol, T2D), no association was found between low versus high mtDNA abundance and ischemic (HR: 1.06 [95% CI: 0.95, 1.18]) or hemorrhagic (HR: 0.97 [95% CI: 0.82, 1.15]) stroke. However, in the MR analyses after removal of platelet count-associated SNPs, we found evidence for an association between genetically-influenced mtDNA abundance and ischemic stroke (odds ratio, 1.17; confidence interval, 1.03, 1.32).

Conclusions

Although the results from both multivariable-adjusted prospective and basis MR analyses did not show an association between low mtDNA and increased risk of ischemic stroke, in-depth MR sensitivity analyses may suggest evidence for a causal relationship.

INTRODUCTION

Stroke is the second leading cause of death and loss of disability-adjusted life years worldwide [1]. Oxidative stress has been hypothesized to play an important role in the pathophysiology of stroke by aggravating secondary damage and increases reperfusion injury after ischemic stroke [2-4]. As a result of direct or indirect Reactive Oxygen Species (ROS)-induced damage to the (cerebral) vascular wall, multiple aspects of the vascular system are affected including platelet aggregation, endothelial function, vascular permeability and vasodilation [3]. These local vessel changes induced by oxidative stress can also gradually develop before stroke onset, and therefore may also lead to an increased risk of stroke incidence [5].

Mitochondria are a major source of ROS production [6]. Mitochondrial dysfunction leads to an increase in ROS production due to a change in redox homeostasis [7]. Additionally, impaired mitochondrial function, frequently proxied by the mitochondrial copy number (mtDNA-CN) [8, 9], has been associated with diseases such as diabetes, heart failure, and neurological defects [10]. mtDNA-CN can be assessed relatively easy in large populations by estimating the mtDNA abundance from the intensities of genotyping probes representing mitochondrial DNA on genotyping arrays [9, 11, 12]. Increased ROS production drives mitochondrial dysfunction causing increased defects in mitochondrial fusion, fission, and mitophagy activation [13], which subsequently lead to subsequent excessive ROS production [13].

Although a relatively small study was not able to provide evidence of an association between low mtDNA-CN and increased stroke risk [14], we hypothesized that leukocyte mtDNA might affect brain pathologies, given the available biological data. Based on the combination of the postulated detrimental biological effect of blood oxidative stress on the (cerebro)vascular endothelial system and its role in secondary damage after stroke occurrence, we investigated the prospective association between mtDNA abundance and incident ischemic and hemorrhagic stroke in a large cohort of European-ancestry participants from the UK Biobank. In addition, we applied Mendelian Randomization (MR) to provide evidence for possible causality [15, 16] as a way to triangulate the results from the prospective analyses by obtaining results from two analysis methods, both with different assumptions and limitations [15].

MATERIALS AND METHODS

Population description

The UK Biobank cohort is a prospective general population cohort with 502,628 participants between the age of 40 and 70 years recruited from the general population between 2006 and 2010 [17] (more information can be found online https://www.

ukbiobank.ac.uk). Blood samples were collected for genotyping. Access for information to invite participants was approved by the Patient Information Advisory Group (PIAG) from England and Wales. All participants in the UK Biobank provided a written informed consent and local research ethics committees and institutional review boards approved the study. The present study was accepted under project number 56340.

In the present study, genotyped European-ancestry participants were followed (N = 488 377). Exclusion criteria included: 1) non-European ancestry; 2) participants who failed genotyping quality control and/or with low call rate; 3) related individuals defined by principal components (PCs); 4) participants with high SD of autosomal probes; 5) history of any stroke at baseline; 6) missingness on covariates. After the exclusions, the final analyses were performed in 288,572 participants.

Mitochondrial DNA abundance

We used somatic mtDNA abundance as a proxy measure of mtDNA-CN, as the exposure, which is determined from the intensities of genotyping probes on the mitochondrial chromosome on the Affymetrix Array. The method for computing mtDNA abundance has been described previously [11]. In brief, the relative amount of mtDNA hybridized to the array at each probe was the log2 transformed ratio (L2R) of the observed genotyping probe intensity divided by the intensity at the same probe observed in a set of reference samples. We used the median L2R values across all 265 variants passing quality control on the MT chromosome as an initial raw measure of mtDNA abundance. To correct for confounding induced by poorly performing probes, we weighted L2R values of each probe by multiplying the weight of the probe that are generated from a multivariate linear regression model in which those intensities statistically significantly predicted normalized mitochondrial coverage from exome sequencing data, resulting in a single mtDNA abundance estimate for each individual. To eliminate the plate effect, we subsequently normalized the abundance to mean of zero and SD of one within each genotyping plate consisting of 96 wells [9].

Covariates

In addition to age and sex, we took into account data based on self-reported questionnaires (smoking, alcohol consumption, disease status, medication use), blood cell counts (white blood cell counts and platelet counts), body mass index (BMI) in kg/m^2, serum lipid levels (total and LDL cholesterol) in mmol/L, and systolic and diastolic blood pressure in mmHg.

Outcome

The outcome in the analysis was ischemic and hemorrhagic stroke separately, as well as combined, in the time period August 2006 up to January 2021. Stroke incidence was obtained via hospital admission data and national health register data and used to identify the date of the first stroke or stroke-related death after baseline assessment. The primary outcomes were any stroke incidence and further specified ischemic and hemorrhagic stroke incidence. Incident disease diagnoses are coded according to the International Classification of Diseases edition 10 (ICD-10); Ischemic stroke was defined as I63 and hemorrhagic stroke as I61. Any stroke was defined as the combination of I63 and I61. Follow-up time is computed from the baseline visit to the diagnosis of incident disease, loss-to-follow-up or death, or the end of the study period, whichever came first.

Data required for the Mendelian Randomization analyses

For the MR, genetic variants of mtDNA abundance were used as instrument variables. In a previous study, 129 independent Single-Nucleotide Polymorphisms (SNPs) as genetic variants were found to be independently associated with mtDNA abundance at a genome-wide significance threshold ($p < 5x10^{-8}$); SNPs were additionally pruned to an LD $R^2<0.0001$ [18]. The study was performed in a total of 465,809 individuals using a combined population of the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium and the UK Biobank.

Mendelian Randomization outcome datasets

For the extraction of summary statistics on the associations of the mtDNA abundance related SNPs with ischemic stroke, which was defined as any ischemic stroke (I63), three large studies were used: the MEGASTROKE consortium, the UK Biobank, and the FinnGen study [17, 19]. Both UK Biobank and FinnGen were not part of the main analyses of the MEGASTROKE consortium preventing inclusion of overlapping samples in the analyses. In the three studies insufficient data on hemorrhagic stroke was available.

The trans-ancestry meta-analysis from the MEGASTROKE consortium was used to retrieve the ischemic stroke SNP-outcome data and was based on 60,341 cases and 454,450 controls collected from 29 studies of predominantly the European ancestry (86%) [19].

For the MR analyses in UK Biobank, cases developed before and after enrolment were considered. Follow-up information that included ischemic stroke occurrence was retrieved through the routinely available NHS database. In the European-ancestry dataset with full genomics data available, we had data on 2,404 cases of ischemic stroke and 368,771 controls. We performed new genome-wide association analyses using linear mixed models to assess the associations between genetic instruments and ischemic stroke, adjusted for age, sex and 10 principal components, and corrected for familial relationships using BOLT_LMM (v2.3.2).

Data from FinnGen (Freeze 5; https://www.finngen.fi/en/), which is an ongoing cohort study launched in 2017, and analyses were based on 10,551 cases of ischemic stroke, and 202,223 controls.

Although with lower numbers, we additionally performed MR on subtypes of

ischemic stroke (Cardio Embolic Stroke: 7,193 cases, 406,111 controls, Large Artery Atherosclerosis: 4,373 cases, 406,111 controls, Small Vessel Stroke: 5,386 cases, 192,662 controls) using data from MEGASTROKE and hemorrhagic stroke (1,687cases, 201,146 controls) from FinnGen.

Statistical analysis

Multivariable-adjusted analyses

For the analyses, and for presentation purposes, we divided the study population in 5 equally-sized groups based on the mtDNA abundance, with the first quintile containing the group with the lowest levels of mtDNA abundance and the fifth quintile containing the highest levels (used as reference).

Baseline characteristics of the study population were presented separately per quintile of mtDNA abundance, as mean (SD) for continuous variables if they followed a normal distribution, or as median (interquartile range) otherwise, and frequency (proportion) for categoric variables.

The cumulative incidence for competing risk (CICR) was used to plot the cumulative incidence of ischemic and hemorrhagic stroke against follow-up time separately using a Kaplan-Meier survival curve by mtDNA abundance quintiles, where death was accounted for as a competing event. For any, ischemic, and hemorrhagic stroke, a Cox proportional hazards model was used to estimate the hazard ratio (HR) and 95% confidence interval (CI) presented as stroke incidence, comparing the lowest 20% mtDNA abundance with the highest 20%. Analyses were additionally done stratified by sex. Two multivariate regression models were fitted, where for model 2 covariates were first added individually:

- Model 1: age, sex, batch, the first 10 genetic principal components, white blood cell counts, platelet count
- Model 2: Model 1 + BMI, smoking, alcohol consumption, total cholesterol, hypertension, diabetes, cholesterol lowering medication, blood pressure lowering medication

Covariates were included in the regression models given their known relation with both exposure and outcome (age, sex, smoking, alcohol consumption, total cholesterol, disease status, medication status), or were included as a technical correction due to the measurement composition (batch, white blood cell count, platelet count). Participants were censored in the event of loss-to-follow-up or death. In order to check whether the proportional hazards assumption was fulfilled, a Cox proportional hazard assumption test ("*cox.zph*" from R package "Survival") was performed. Additionally, mtDNA-CN was assessed continuously as a one-SD lower mtDNA-CN on stroke incidence. Analyses were performed using the "Survival" (cran.r-project.org/web/packages/survival) package in R (v4.1.0)

Mendelian Randomization

All the analyses were done using R (v4.1.0) statistical software (The R Foundation for Statistical Computing, Vienna, Austria). MR analyses were performed using the R-based package "TwoSampleMR" (https://mrcieu.github.io/TwoSampleMR/) [20].

For our primary MR analysis, Inverse-Variance weighted (IVW) regression analyses were performed [16]. Estimates were calculated for each genetic instrument using the Wald ratio (SNP – outcome association divided by the SNP – exposure association) and subsequently meta-analyzed using the inverse-weighted meta-analyses weighted on the standard error of the SNP-outcome association (assuming no measurement error [NOME] in the exposure) [21]. The calculated estimates were expressed as odds ratios (OR) on ischemic stroke per SD (obtained from the exposure data) difference in mtDNA abundance.

To ensure that the results obtained from the IVW analyses were not biased due to directional pleiotropy, we performed MR-Egger regression analysis and Weighted-Median Estimator [21]. Although MR-Egger is considered as a relatively inefficient approach (e.g., large confidence intervals), this method does not force the regression line to go through the intercept. The intercept depicts the estimated average pleiotropic effect across the genetic variants, and a value that differs from zero indicates that the IVW estimate is biased. [22] The Weighted-Median estimator analysis can provide a consistent valid estimate if at least half of the instrumental variables are valid [23]. In addition, MR-PRESSO (MR Pleiotropy RESidual Sum and Outlier) was applied to detect and correct for horizontal pleiotropy through removing outlying causal estimates based on individual instruments [24], as implemented in the R-based package "MR-PRESSO" (https://github.com/rondolab/MR-PRESSO). The Cochran's Q statistic was performed in order to test the heterogeneity between the estimated Wald ratios from different genetic variants [25]. Additionally, a Steiger directionality test was performed to ensure consistent causal direction-of-effect. A power calculation was performed with mRnd (https://shiny.cnsgenomics.com/mRnd/) [26]. With power = 0.80, minimal effect size (OR) was 1.076.

Recent research has proven that two-sample MR methods can safely be used for one-sample MR in large databases [27]. This allows us to use the UK Biobank database in our sample set despite also being used as our exposure dataset. As a limitation to this method, results of MR-Egger analyses are to be interpreted with caution when used to check for pleiotropy [27].

The main MR analyses were performed in the individual datasets, and subsequently meta-analyzed to derive the pooled estimates for the exposure on the risk of ischemic stroke using a fixed-effect model. Heterogeneity testing of the estimates across three datasets was performed by I^2, and corresponding p-value was obtained from the Cochran's Q test. All meta-analyses were performed in the R-based "meta" package (https://cran.r-project.org/web/packages/meta/index.html).

Sensitivity analysis after stratification of genetic instruments

SNPs identified in relation to mtDNA-CN have been found in relation to platelet activation and megakaryocyte proliferation [18], which both could affect stroke risk and could potentially lead to biased results. We first examined the associations between the SNPs and platelet count in our study sample (adjusted for age, sex, and the first 10 genetic principal components); all SNPs with P<(0.05/123) in its association with platelet count were excluded from further MR sensitivity analyses.

RESULTS

Baseline characteristics of the study population

A total of 288,572 participants were included in the final study sample (see full procedure in **Supplementary Figure 1**) for multivariable-adjusted survival analyses. Participants excluded were due to unavailable genetic data (N = 14,251), not used to compute the genetic principal components (N = 81,623), or having unrealistic SD of autosomal probes (N = 9,440) were excluded according to standard UK Biobank quality control recommendations. Subsequently, we excluded related participants (N = 38,642), and with a non-white British ancestry (N = 65,498). Finally, 4,602 participants were excluded due to a history of stroke before study enrollment. Participants in the lower mtDNA abundance quintile (**Table 1**) had a mean age of 57.5 versus 56.1 year in the highest quintile, a mean BMI of 27.7 versus 27.0 kg/m^2, T2D prevalence of 2.8% versus 2.0%, and 11.3% were current smokers compared with 8.4% in the highest quintile.

Table 1. Baseline characteristics of the study participants stratified by quintiles of mtDNA abundance

	Q1	Q2	Q3	Q4	Q5
N	57 715	57 714	57 714	57 714	57 715
mtDNA abundance (normalized)	-1.4 (0.5)	-0.5 (0.2)	0.0 (0.1)	0.5 (0.2)	1.4 (0.5)
Age (years)	57.5 (8.0)	57.1 (8.0)	56.8 (8.0)	56.5 (8.0)	56.1 (8.0)
Sex (female %)	52.0	53.2	54.0	54.3	55.7
BMI (kg/m^2)	27.7 (5.0)	27.5 (4.8)	27.4 (4.7)	27.2 (4.6)	27.0 (4.5)
Diastolic blood pressure (mmHg)	82.6 (10.2)	82.4 (10.0)	82.3 (10.0)	82.1 (10.1)	81.7 (10.1)
Systolic blood pressure (mmHg)	139.5 (18.8)	138.6 (18.7)	138.2 (18.6)	137.6 (18.3)	136.7 (18.3)
White Blood Cell count (10^9 cells/L)	7.4 (1.8)	7.1 (1.7)	6.9 (1.7)	6.6 (1.7)	6.4 (2.7)
Platelet count (10^9 cells/L)	245.5 (58.0)	250.8 (57.7)	253.4 (58.4)	256.3 (59.2)	259.5 (63.8)
Blood pressure-lowering medication %					
Yes	19.8	18.4	17.5	16.6	15.6
No	80.2	81.6	(82.5	83.4	84.4
Cholesterol (mmol/L)	5.8 (1.2)	5.7 (1.1)	5.7 (1.1)	5.7 (1.1)	5.7 (1.1)
Cholesterol lowering medication %					
Yes	13.9	13.5	13.1	12.5	12.2
No	86.1	86.5	86.9	87.5	87.8

Table 1. Continued

	Q1	Q2	Q3	Q4	Q5
Alcohol consumption %					
Less than once per week	29.1	28.2	27.8	27.3	26.6
Once or twice per week	25.6	26.4	26.2	26.4	26.6
More than four times per week	45.3	45.2	45.9	46.3	46.7
Smoking %					
Never	53.3	54.1	55.0	55.6	56.5
Past	35.0	35.2	34.8	34.9	34.9
Current	11.3	10.4	9.9	9.4	8.4
Type 2 Diabetes %					
Yes	2.8	2.5	2.3	2.2	2.0
No	97.2	97.5	97.7	97.8	98.0

Data are mean (SD) for continuous variables or percentages for dichotomous variables. mtDNA abundance is presented as normalized in unit standard deviations. Abbreviations: BMI, Body Mass Index.

Multivariable-adjusted survival analyses mtDNA abundance and stroke

A total of 6,218 of the 288,572 participants (2.15%) had a stroke incidence, of which 3,994 (1.38%) were ischemic and 1,883 (0.65%) hemorrhagic over a median (IQR) follow-up of 11.8 (11.1 – 12.5) years. The incidence of ischemic stroke was higher in the lower mtDNA-CN quintiles than in the higher quintiles (**Figure 1A**), while hemorrhagic stroke incidence was similar in all mtDNA-CN quintiles (**Figure 1B**); in both cases the analyses fulfilled the proportional hazard assumption (*p-value:* 0.84 & 0.88).

After stratification based on mtDNA-CN (**Table 2**), in model 1, mtDNA abundance was associated with any stroke and ischemic stroke incidence, when comparing the first quintile with the highest 20% mtDNA abundance (any stroke: hazard ratio (HR), 1.11; 95% confidence interval (CI): 1.02 to 1.20; ischemic stroke: HR, 1.15; 95% CI: 1.04 to 1.27). Similarly, a one-SD increase in mtDNA abundance was associated with lower risk of incident ischemic stroke (HR, 0.96; 95% CI: 0.93 to 0.99). No association was found between mtDNA abundance and incident hemorrhagic stroke.

After correcting for other confounders, the associations with stroke and ischemic stroke attenuated (any stroke: HR, 1.06; 95% CI: 0.97 to 1.16; ischemic stroke: HR, 1.07; 95% CI: 0.95 to 1.19), as did the continuous model on ischemic stroke (HR, 0.98; 95% CI: 0.94 to 1.01).

Mendelian Randomization on mtDNA abundance and ischemic stroke

Main analyses

We did not observe evidence favoring an association between genetically-influenced lower mtDNA-CN and ischemic stroke (**Figure 2**). The odds ratios per 1 SD less mtDNA-CN were 1.07 (95%CI: 0.95, 1.20) in MEGASTROKE, 1.04 (95%CI: 0.79, 1.37) in the UK Biobank, and 0.99 (95%CI: 0.82, 1.20) in FinnGen. After meta-analysis, in a combined sample size of 1,098,740 (of which 73,296 cases), the pooled odds ratio was 1.04 (95%CI:

0.95 to 1.15) per 1-SD decrease in genetically-influenced mtDNA abundance.

The exact set of variants, their corresponding coefficients, standard errors, and p-values are presented in **Supplementary Table 1**. Variance explained (R^2) was 2.0% and calculated based on the derived summary statistics. The MR-Egger intercept indicated no pleiotropy (p > 0.05). Although several outliers were identified with MR-PRESSO in MEGASTROKE and FinnGen, results remained similar after removal of these outlying SNPs. The Steiger test of directionality showed a correct causal direction, indicating that there is no evidence for reverse causation, and no different results were observed with MR-sensitivity analyses MR-Egger and weighted median (**Supplementary Table 2.**).

Sub-analyses performed with separate outcomes cardioembolic, large artery atherosclerosis, small-vessel, and hemorrhagic stroke (**Supplementary Figure 2 & 3**) showed no evidence favoring a different result.

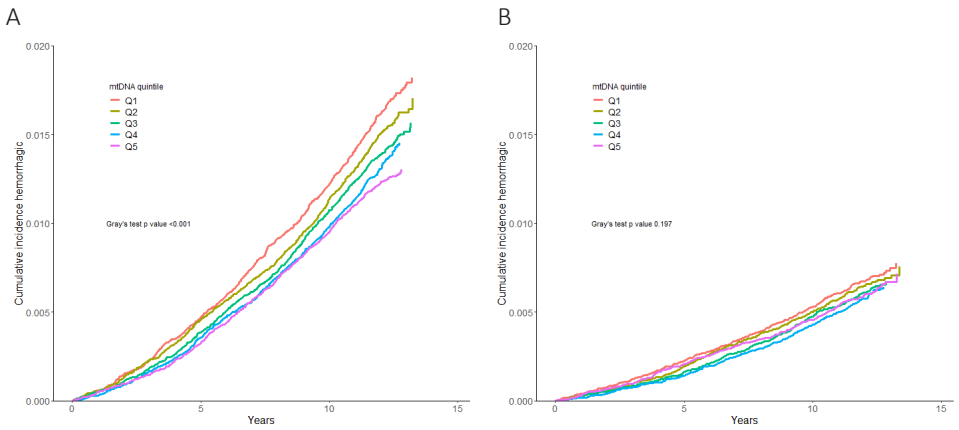

Figure 1. Cumulative incidence of ischemic (A) and hemorrhagic (B) stroke by quintiles of mtDNA abundance
We calculated the Cumulative incidence for ischemic and hemorrhagic stroke, accounting for death as a competing event. Differences in cumulative incidence between groups were assessed using Gray's test.

Additional sensitivity analyses

A total of 61 SNPS were associated with platelet count, which were subsequently excluded from additional sensitivity analyses. In the full sample, a 1-SD genetically-determined lower mtDNA abundance was associated with a higher risk of ischemic stroke (OR: 1.165; 95% CI: 1.026 to 1.323), although results from FinnGen did not align with those obtained in UK Biobank and MEGASTROKE (**Supplementary Figure 4**).

Table 2. The multivariable-adjusted association between mtDNA abundance and incident stroke in European-ancestry participants from UK Biobank.

		Continuous	Q1	Q2	Q3	Q4	Q5
Stroke incidence		**HR (95%CI)**	**HR (95%CI)**	**HR (95%CI)**	**HR (95%CI)**	**HR (95%CI)**	**HR (95%CI)**
Any	Model 1	0.97 (0.95, 1.00)	1.11 (1.02, 1.20)	1.07 (0.98, 1.16)	1.03 (0.95, 1.12)	1.03 (0.95, 1.12)	1.00 (ref)
	Model 2	0.99 (0.96, 1.02)	1.06 (0.97, 1.16)	1.05 (0.96, 1.15)	1.02 (0.94, 1.12)	1.01 (0.92, 1.10)	1.00 (ref)
Ischemic	Model 1	0.96 (0.93, 0.99)	1.15 (1.04, 1.27)	1.11 (1.00, 1.23)	1.07 (0.96, 1.18)	1.05 (0.95, 1.16)	1.00 (ref)
	Model 2	0.98 (0.94, 1.01)	1.07 (0.96, 1.19)	1.08 (0.97, 1.21)	1.03 (0.92, 1.15)	1.00 (0.89, 1.12)	1.00 (ref)
Hemorrhagic	Model 1	1.02 (0.97, 1.07)	0.97 (0.84, 1.13)	0.98 (0.84, 1.13)	0.93 (0.80, 1.09)	0.92 (0.79, 1.07)	1.00 (ref)
	Model 2	1.02 (0.97, 1.07)	0.98 (0.83, 1.15)	0.99 (0.84, 1.17)	0.97 (0.82, 1.14)	0.93 (0.79, 1.10)	1.00 (ref)

Estimated hazard ratios per-SD increase in mtDNA abundance (continuous), or for the 1st to the 4th quintile compared with the 5th (reference) quintile (categorical) on any, ischemic, and hemorrhagic stroke. Model 1 includes age, sex, batch, PCs, white blood cell count, platelet. Model 2 includes model 1, BMI, smoking, total cholesterol, hypertension, diabetes, cholesterol lowering medication, blood pressure lowering medication. Abbreviations: CI, confidence interval; HR, hazard ratio.

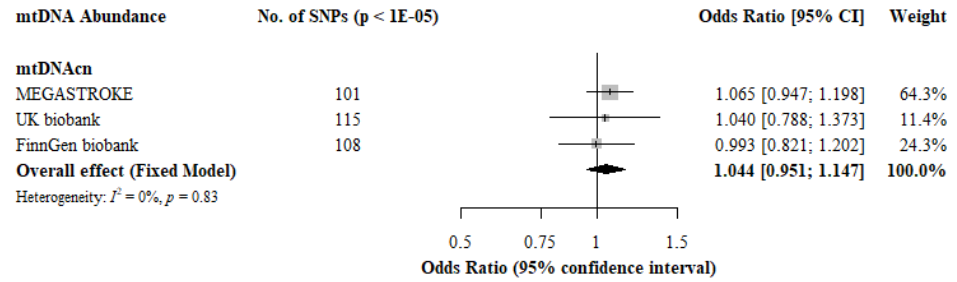

Figure 2. Causal association between mtDNA abundance and ischemic stroke occurrence.
Estimated ORs represent the effect per SD decrease in mtDNA copy number on ischemic stroke. Results were obtained using a Mendelian Randomization inverse-variance weighted method, analyzed per outcome database and combined over the three databases using fixed-effect meta-analyses.

DISCUSSION

In the UK Biobank cohort, consisting of 288,572 participants after exclusion, an initial association was found between mtDNA abundance and incident ischemic stroke, which attenuated after adjustment for confounders. Consistent with the prospective analyses, MR analyses, using a total sample size of 73,296 cases and 1,025,444 controls, showed no evidence for an association between genetically-predicted mtDNA abundance and ischemic stroke. However, some in-depth sensitivity analyses in which SNPs associated with platelet count were excluded, did provide some preliminary evident for low mtDNA-CN as possible causal driver for ischemic stroke.

Although of specific interest, caution of these results is warranted given that the results were mainly driven by results derived from MEGASTROKE, and to a lesser extent by UK Biobank. Furthermore, results from these additional MR sensitivity analyses deviated significantly from those observed in the prospective multivariable-adjusted analyses, and therefore do not meet the requirements for triangulation [15]. Collectively, our results indicate that there is only weak evidence for a causal association between mtDNA abundance and ischemic stroke, and more studies are required to elucidate the nature of the pleiotropy identified in our study, which goes beyond the current scope.

Previously, an association between low mtDNA-CN and increased risk of incident stroke in 20,162 participants who were followed over a 13.5-year period during which 1584 stroke events occurred [28], and therefore deviate from our study done in a larger sample of 288,752 participants with 6,218 stroke cases. Difference in baseline health characteristics are possible reasons explaining the observed differences in results.

Recent studies showed that mtDNA-CN could be a marker of stroke prognosis after hospitalization [29, 30]. By analyzing mtDNA-CN, and consequently oxidative stress, our findings did provide some, albeit circumstantial, evidence for a relationship between oxidative stress and stroke occurrence, although this association attenuated after adjustment for confounders. In the Mendelian Randomization analysis, after excluding SNPs associated with platelet count, we also found an association between genetically determined mtDNA abundance and ischemic stroke risk. In contrast to ischemic stroke, we did not find an association between mtDNA abundance and hemorrhagic stroke in the univariate or MR analyses. This difference might be explained because hemorrhagic stroke, in contrast to ischemic stroke, is also often caused by non-classic cardiovascular mechanisms such as vascular amyloid deposition in cerebral amyloid angiopathy [31].

Our data on mtDNA abundance was obtained from leukocytes. Although some of the leukocytes may be directly involved in the pathology of stroke, additional cell types such as endothelial and smooth muscle cells, that we did not query for mitochondrial abundance, are clearly more directly involved. This could potentially explain our overall null findings. Studies on the differences in mitochondrial function within an individual between cell groups are largely non-existent. However, mtDNA-CN measured in blood has been associated with gene expression in other tissues, which suggests mtDNA-CN derived from leukocytes can reflect metabolic health across multiple tissues [32]. Thus, the evidence so far indicates that mitochondrial dysfunction, as measured with leukocyte mtDNA-CN, is systemic. Of interest, using similar methodology as in our study, low genetically-influenced mtDNA has recently been associated with increased dementia risk [12]. This would further indicate that lower mtDNA-CN, although measured in leukocytes, can reflect processes of a systemic increase in disease risk.

A key strength of this study is the statistical power of the analyses of the association between stroke and mitochondrial abundance (288,572 participants for the multivariable survival analysis and 1,098,740 for the MR, respectively). Additionally, we adopted the triangulation of causal inference [15]. By using two different approaches in observational research to study the association between low mtDNA abundance and (ischemic) stroke risk, we increased the credibility of our results. Although results from both our used approaches were not exactly similar, they were directionally consistent.

Some limitations are to be considered. First, mtDNA abundance was determined from intensities of genotyping probes on the mitochondrial DNA, whereas the assessment with whole-exome sequencing is generally considered to result in more reliable mtDNA abundance estimates [33, 34]. Although Hägg et al showed a moderate correlation between mtDNA based on SNP array intensities and exome sequencing of 0.33 [11], analyses still indicated the measurements of SNP array intensities reflect underlying biology of mtDNA abundance. For this reason, the increased variance is most likely the result of nondifferential measurement error, and therefore considered to be mainly cause a reduction in statistical power. As a main consequence, the true associations, particularly those from the multivariable-adjusted prospective analyses are most likely larger than as observed. Second, our study population consists of predominantly Caucasian participants, limiting the generalizability of the results to other ancestry groups. Third, Mendelian Randomization functions on several assumptions. However,

by using several sensitivity analyses such as MR-Egger and MR-PRESSO, we can establish with some conviction that these are fulfilled. To add, although in a one-sample MR (as conducted in the UK Biobank) the assumption of independence does not hold up, previous studies have shown that two-sample MR methods can be used reliably with large enough biobanks [27]. Last, despite a large sample size in the multivariable adjusted analysis, stroke and especially hemorrhagic stroke incidences were relatively few. However, as an association was found before correction, we think our analyses had enough power to detect a difference between groups. In addition, we used one of the larger data sets available.

In conclusion, despite a large sample size our prospective study did not find evidence for an association between mtDNA abundance and ischemic or hemorrhagic stroke. After exclusion of pleiotropic SNPs associated with platelet count, we found some preliminary evidence for an association between genetically determined lower mtDNA-CN and ischemic stroke risk using MR. However, further studies are required for validation and to examine the nature of this type of pleiotropy.

Acknowledgements

The authors are grateful to the UK Biobank for allowing us the use of their data. The analyses done in UK Biobank were done under project number 56340. Furthermore, the authors acknowledge the participants and investigators of the MEGASTROKE consortium and the FinnGen Biobank who contributed to the summary statistics data which are made available for further studies.

LGM, JL, RN and DvH designed research; LGM and JL conducted research; LGM and JL performed statistical analysis; LGM, JL, RN, and MJHW wrote paper; LGM had primary responsibility for final content. DvH, KWvD, SH, FG and MJHW contributed to the data interpretation and commented on initial versions of the manuscript; All authors read and approved the final manuscript.

Funding

This work was supported by the VELUX Stiftung [grant number 1156] to DvH and RN, and JL was supported by the China Scholarship Counsel [No. 201808500155]. RN was supported by an innovation grant from the Dutch Heart Foundation [grant number 2019T103 to R.N.]. Parts of this work were funded by the Åke Wibergs Foundation (grant number M19-0294 to F.G).

Disclosures

None

Conflicts of interest statement

The authors declare to have no conflict of interest.

Supplemental Material

Table S1-2

Figure S1-4

REFERENCES

1. Wang, H., *Global age-sex-specific fertility, mortality, healthy life expectancy (HALE), and population estimates in 204 countries and territories, 1950-2019: a comprehensive demographic analysis for the Global Burden of Disease Study 2019.* Lancet, 2020. **396**: p. 1160-1203.
2. Rodrigo, R., R. Fernandez-Gajardo, R. Gutierrez, J.M. Matamala, R. Carrasco, A. Miranda-Merchak, and W. Feuerhake, *Oxidative stress and pathophysiology of ischemic stroke: novel therapeutic opportunities.* CNS Neurol Disord Drug Targets, 2013. **12**: p. 698-714.
3. Allen, C.L. and U. Bayraktutan, *Oxidative stress and its role in the pathogenesis of ischaemic stroke.* Int J Stroke, 2009. **4**: p. 461-70.
4. Duan, X., Z. Wen, H. Shen, M. Shen, and G. Chen, *Intracerebral Hemorrhage, Oxidative Stress, and Antioxidant Therapy.* Oxid Med Cell Longev, 2016. **2016**: p. 1203285.
5. Shirley, R., E.N. Ord, and L.M. Work, *Oxidative Stress and the Use of Antioxidants in Stroke.* Antioxidants (Basel), 2014. **3**: p. 472-501.
6. Lackner, L.L. and J.M. Nunnari, *The molecular mechanism and cellular functions of mitochondrial division.* Biochim Biophys Acta, 2009. **1792**: p. 1138-44.
7. Bandy, B. and A.J. Davison, *Mitochondrial mutations may increase oxidative stress: implications for carcinogenesis and aging?* Free Radic Biol Med, 1990. **8**: p. 523-39.
8. Chung, J.K., S.Y. Lee, M. Park, E.J. Joo, and S.A. Kim, *Investigation of mitochondrial DNA copy number in patients with major depressive disorder.* Psychiatry Res, 2019. **282**: p. 112616.
9. Longchamps, R.J., C.A. Castellani, S.Y. Yang, C.E. Newcomb, J.A. Sumpter, J. Lane, M.L. Grove, E. Guallar, N. Pankratz, K.D. Taylor, et al., *Evaluation of mitochondrial DNA copy number estimation techniques.* PLoS One, 2020. **15**: p. e0228166.
10. Yang, J.L., S. Mukda, and S.D. Chen, *Diverse roles of mitochondria in ischemic stroke.* Redox Biol, 2018. **16**: p. 263-275.
11. Hagg, S., J. Jylhava, Y. Wang, K. Czene, and F. Grassmann, *Deciphering the genetic and epidemiological landscape of mitochondrial DNA abundance.* Hum Genet, 2021. **140**: p. 849-861.
12. Chong, M., P. Mohammadi-Shemirani, N. Perrot, W. Nelson, R. Morton, S. Narula, R. Lali, I. Khan, M. Khan, C. Judge, et al., *GWAS and ExWAS of blood Mitochondrial DNA copy number identifies 71 loci and highlights a potential causal role in dementia.* Elife, 2022. **11**.
13. Liu, F., J. Lu, A. Manaenko, J. Tang, and Q. Hu, *Mitochondria in Ischemic Stroke: New Insight and Implications.* Aging Dis, 2018. **9**: p. 924-937.
14. Wachsmuth, M., A. Hubner, M. Li, B. Madea, and M. Stoneking, *Age-Related and Heteroplasmy-Related Variation in Human mtDNA Copy Number.* PLoS Genet, 2016. **12**: p. e1005939.
15. Lawlor, D.A., K. Tilling, and G. Davey Smith, *Triangulation in aetiological epidemiology.* Int J Epidemiol, 2016. **45**: p. 1866-1886.
16. Burgess, S., A. Butterworth, and S.G. Thompson, *Mendelian randomization analysis with multiple genetic variants using summarized data.* Genet Epidemiol, 2013. **37**: p. 658-65.
17. Sudlow, C., J. Gallacher, N. Allen, V. Beral, P. Burton, J. Danesh, P. Downey, P. Elliott, J. Green, M. Landray, et al., *UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age.* PLoS Med, 2015. **12**: p. e1001779.
18. Longchamps, R.J., S.Y. Yang, C.A. Castellani, W. Shi, J. Lane, M.L. Grove, T.M. Bartz, C. Sarnowski, C. Liu, K. Burrows, et al., *Genome-wide analysis of mitochondrial DNA copy number reveals loci implicated in nucleotide metabolism, platelet activation, and megakaryocyte proliferation.* Hum Genet, 2022. **141**: p. 127-146.
19. Malik, R., G. Chauhan, M. Traylor, M. Sargurupremraj, Y. Okada, A. Mishra, L. Rutten-Jacobs, A.K. Giese, S.W. van der Laan, S. Gretarsdottir, et al., *Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes.* Nat Genet, 2018. **50**: p. 524-537.
20. Hemani, G., J. Zheng, B. Elsworth, K.H. Wade, V. Haberland, D. Baird, C. Laurin, S. Burgess, J. Bowden, R. Langdon, et al., *The MR-Base platform supports systematic causal inference across the human phenome.* Elife, 2018. **7**.
21. Burgess, S. and S.G. Thompson, *Interpreting findings from Mendelian randomization using the MR-Egger method.* Eur J Epidemiol, 2017. **32**: p. 377-389.
22. Bowden, J., G. Davey Smith, and S. Burgess, *Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression.* Int J Epidemiol, 2015. **44**: p. 512-25.
23. Bowden, J., G. Davey Smith, P.C. Haycock, and S. Burgess, *Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator.* Genet Epidemiol, 2016. **40**: p. 304-14.
24. Verbanck, M., C.Y. Chen, B. Neale, and R. Do, *Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases.* Nat Genet, 2018. **50**: p. 693-698.
25. Greco, M.F., C. Minelli, N.A. Sheehan, and J.R. Thompson, *Detecting pleiotropy in Mendelian randomisation studies with summary data and a continuous outcome.* Stat Med, 2015. **34**: p. 2926-40.
26. Brion, M.J., K. Shakhbazov, and P.M. Visscher, *Calculating statistical power in Mendelian randomization studies.* Int J Epidemiol, 2013. **42**: p. 1497-501.
27. Minelli, C., M.F. Del Greco, D.A. van der Plaat, J. Bowden, N.A. Sheehan, and J. Thompson, *The use of two-sample methods for Mendelian randomization analyses on single large datasets.* Int J Epidemiol, 2021.
28. Ashar, F.N., Y. Zhang, R.J. Longchamps, J. Lane, A. Moes, M.L. Grove, J.C. Mychaleckyj, K.D. Taylor, J. Coresh, J.I. Rotter, et al., *Association of Mitochondrial DNA Copy Number With Cardiovascular Disease.* JAMA Cardiol, 2017. **2**: p. 1247-1255.
29. Song, L., T. Liu, Y. Song, Y. Sun, H. Li, N. Xiao, H. Xu, J. Ge, C. Bai, H. Wen, et al., *mtDNA Copy Number Contributes to All-Cause Mortality of Lacunar Infarct in a Chinese Prospective Stroke Population.* J Cardiovasc Transl Res, 2020. **13**: p. 783-789.
30. Chong, M.R., S. Narula, R. Morton, C. Judge, L. Akhabir, N. Cawte, N. Pathan, R. Lali, P. Mohammadi-Shemirani, A. Shoamanesh, et al., *Mitochondrial DNA Copy Number as a Marker and Mediator of Stroke Prognosis: Observational and Mendelian Randomization Analyses.* Neurology, 2022. **98**: p. e470-e482.
31. Grysiewicz, R.A., K. Thomas, and D.K. Pandey, *Epidemiology of ischemic and hemorrhagic stroke: incidence, prevalence, mortality, and risk factors.* Neurol Clin, 2008. **26**: p. 871-95, vii.
32. Yang, S.Y., C.A. Castellani, R.J. Longchamps, V.K. Pillalamarri, B. O'Rourke, E. Guallar, and D.E. Arking, *Blood-derived mitochondrial DNA copy number is associated with gene expression across multiple tissues and is predictive for incident neurodegenerative disease.* Genome Res, 2021. **31**: p. 349-358.
33. Zhang, P., B.D. Lehmann, D.C. Samuels, S. Zhao, Y.Y. Zhao, Y. Shyr, and Y. Guo, *Estimating relative mitochondrial DNA copy number using high throughput sequencing data.* Genomics, 2017. **109**: p. 457-462.
34. Picardi, E. and G. Pesole, *Mitochondrial genomes gleaned from human whole-exome sequencing.* Nat Methods, 2012. **9**: p. 523-4.

5

THE IMPACT OF SOCIODEMOGRAPHIC STATUS ON THE ASSOCIATION BETWEEN CLASSICAL CARDIOVASCULAR RISK FACTORS AND CORONARY ARTERY DISEASE: A STRATIFIED MENDELIAN RANDOMIZATION STUDY

Leon G Martens
Daan van Hamersveld
Saskia le Cessie
Ko Willems van Dijk
Diana van Heemst
Raymond Noordam

Submitted

ABSTRACT

Background and aims

Low socioeconomic status (SES) is associated with cardiovascular risk factors and increased coronary artery disease (CAD) risk. We tested whether SES is an effect modifier of the association between classical cardiovascular risk factors and CAD using SES-stratified Mendelian Randomization (MR) in European-ancestry participants from UK Biobank.

Methods

We calculated weighted genetic risk scores (GRS) for the risk factors body mass index (BMI), systolic blood pressure, LDL cholesterol, and triglycerides. Participants were stratified by Townsend deprivation index (TDI) score. Logistic regression models were used to investigate associations between GRSs and CAD occurrence using MR. Additionally, stratification based on GRS-adjusted TDI residuals was conducted to correct for possible collider-stratification bias.

Results

In a total sample size of N=446,485, with 52,946 cases, the risk for CAD per SD increase in genetically-influenced BMI was highest in the group with the lowest 25% SES (OR: 1.126, 95% CI: 1.106-1.145; OR: 1.081, 95% CI: 1.059-1.103 in high SES), remaining similar after controlling for possible collider-stratification bias. The effects of genetically-influenced systolic blood pressure, LDL cholesterol, and triglyceride on CAD were similar between SES groups.

Conclusion

CAD risk attributable to increased BMI is not homogenous and could be modified by SES. This emphasizes the need of tailor-made approaches for BMI-associated CAD risk reduction.

INTRODUCTION

Individuals with low socioeconomic status (SES) are at an increased risk of developing cardiovascular disease (CVD) [1-4]. Although it is known that a low SES is associated with adverse lifestyle factors such as smoking, alcohol consumption, unhealthy diet, and physical inactivity [5-8], the exact mechanisms underlying the link between SES and risk for CVD are still relatively unclear. Additionally, classic cardiovascular risk factors such as high body mass index (BMI), elevated blood lipid levels and systolic blood pressure (BP) play prominent roles in the pathogenesis of (atherosclerotic) CVD [9]. Simultaneously, studies indicate that these classic risk factors are also generally more prevalent in individuals with lower SES [10-13], emphasizing the complex interplay between SES, CVD risk factors and the development of disease.

In a previous study conducted in the UK biobank, it was shown that as much as 40% of the association between low education, as a reflection of low SES, and increased CVD risk was mediated by BMI, BP and smoking behavior [14]. Additionally, evidence is emerging that the risk for CVD associated with the classical CVD risk factors is not universal throughout subgroups of the general population, and differs for example already for different age groups and for the different sexes in observational studies [15, 16]. The heterogeneous risk factor-CVD associations in different groups of the general population emphasizes the need of a “tailor-made” approach for clinical decision making. These observations are in line with the general hypothesis that atherogenic cardiovascular diseases is not a single disease entity but a dynamic disease construct with changing pathogenesis depending on specific patient characteristics throughout life. For example, and in line with this concept, we previously showed, using Mendelian Randomization approaches, that the impact of genetically-influenced increased BMI on the risk of developing type 2 diabetes was dependent on the age of diagnosis where older people with higher BMI were less susceptible for developing T2D [17]. In addition, the impact of classical genetically-influenced CVD risk factors on coronary artery disease (CAD) attenuated for increasing age of diagnosis [18].

We hypothesized that SES is an important factor that can modify the impact of classical CVD risk factors on CVD, in addition to age and sex. If confirmed, this would mean that interventions tailored to specific SES groups may achieve a larger reduction in CVD risk not only due to low SES groups having a higher average BMI, but also due to SES acting as a catalyst for BMI attributable CAD risk. To omit potential reverse causation and/or most confounding in our analyses, we used a two-sample Mendelian Randomization (MR) approach. Here, genetic variants are used as instrumental variables for given exposures to approximate the effect of life-long exposure to risk factors on the development of disease outcomes [19-21]. In the present study [22], we assessed the associations between classical genetically-influenced CVD risk factors and CAD, stratified for SES, in a large cohort of European-ancestry participants from the UK Biobank.

Methods

Study setting and population

The UK Biobank is a prospective general population cohort with 502,628 participants between the age of 40 and 70 years at the moment of enrollment [23]. Recruitment took place between 2006 and 2010 (more information can be found online https://www.ukbiobank.ac.uk). Invitation letters were sent to eligible adults registered to the National Health Services (NHS) and living within a 25 miles distance from one of the assessment centers. Participants provided information on their lifestyle and medical history through touch-screen questionnaires and physical measurements. Blood samples were collected for biochemistry analyses and genotyping.

The UK biobank study was approved by the North-West Multi-center Research Ethics Committee (MREC). Access for information to invite participants was approved by the Patient Information Advisory Group (PIAG) for England and Wales. All participants in the UK Biobank provided a written informed consent. The present study was accepted under project number 56340.

We restricted all our analyses to participants of European origin (N=446,485), as to limit confounding by ethnic genetic variation. Townsend Deprivation Index (TDI) scores, a measure of SES, were collected at baseline for nearly all of the participants in the study (N=445,965).

Genotyping, genetic imputations and genetic risk scores

For our study, we conducted a two-sample, stratified mendelian randomization analyses, where weighted genetic risk scores (GRS) were used to represent the genetically-determined higher BMI, low-density lipoprotein cholesterol (LDL-C) levels, triglyceride (TG) levels and systolic blood pressure (SBP). These weighted GRS were calculated using independent lead genetic variants (p-value<$5x10^{-8}$) that have been previously identified in genome-wide association meta-analyses in which the UK Biobank population did not contribute. The GRS score for BMI was based on data from 339,224 individuals; 76 SNPs) [24], LDL cholesterol level on188,577 individuals; 15 SNPs [25], triglycerides on 188,577 individuals; 20 SNPs [25], and systolic blood pressure 200,000 individuals; 42 SNPs) [26]. The published beta estimates for the independent lead variants in these meta-analyses were subsequently used to calculate the weighted GRS for each CVD risk factor for each participant in the UK biobank study. Overlapping independent lead variants [25] between LDL cholesterol and triglyceride levels in the genetic risk scores were not taken into account in the calculation of the GRS with the intention to limit bias by (directional) pleiotropy.

The genotyping of the UK Biobank population was performed for roughly 50,000 participants by Affymetrix, using a BiLEVE Axium array. For the other UK Biobank participants, the genotyping was performed using the Affymetrix UK Biobank Axiom array. More information on the genotyping processes can be found online (https://www.ukbiobank.ac.uk). Based on the genotyped data from these arrays, the UK Biobank resources performed imputation on the autosomal SNPs using the UK10K haplotype [27], 1000 Genomes Phase 3 [28], and Haplotype Reference Consortium [29] as reference panels.

Socioeconomic status

To stratify the UK biobank population into different SES groups, we used the Townsend deprivation index (TDI) [30]. This calculated index score, defined at the moment of enrolment, is a composition of four different variables, all related to SES: unemployment, non-ownership of a home, non-ownership of a car, and household overcrowding [30]. Importantly, the TDI is not linked to a specific individual, but instead linked to the postal codes from the UK Biobank participants and is therefore a reflection of overall socioeconomic status of the neighborhood in which the participants are living.

The TDI scores recorded within the UK Biobank ranged between-6.26 and 11.00, and lower scores are reflective of a higher SES in the neighborhood. Using quartiles of these scores, the population was divided into four groups. Because some individuals had very high TDI values, we performed sensitivity analyses by dividing the highest TDI group into two subgroups based on the 87.5 percentile of TDI, and repeated the main analysis accordingly.

Coronary artery disease

Coronary artery disease occurrence (either before or after enrollment in UK Biobank) was the primary outcome for the analyses. Diagnoses were coded according to the International Classification of Diseases (ICD) [23]. Here, the study outcome was CAD which we defined as: angina pectoris (I20), myocardial infarction (I21 and I22), and acute and chronic ischemic heart disease (I24 and I25). Cases were ascertained through a UKB algorithm combining data from linked hospital admissions, death registries, reports from the general practitioner and self-report.

Statistical analysis

Mendelian Randomization

All the analyses were done using R (v4.1.0) statistical software (The R Foundation for Statistical Computing, Vienna, Austria) [31]. In our MR analyses, the associations between genetically-determined CVD risk factors and CAD were calculated using multivariable-adjusted logistic regression analyses, with CAD as dependent and the weighted GRS score as exposure and corrected for age, sex, and the first 10 principal components. In addition, the analyses were stratified by the TDI score categories to study the possible effect modification of the association between the genetically-influenced CVD risk factor and CAD by SES. The results derived from these models (with

accompanied 95% confidence intervals) can be interpreted as the change in odds on CAD for every increase in standard deviation (SD) in genetically-determined exposure. All analyses were adjusted for age, sex and the first 10 genetic principal components to correct for possible population admixture. Additional analyses were performed where we stratified the study population for men and women. We tested for evidence for an interaction on a multiplicative scale by adding multiplicative interaction terms between the GRS's and TDI (both as continuous variables) in the multivariable-adjusted logistic regression models for CAD. For these analyses, we reported the p-values for interaction, corrected for multiple testing using the Bonferroni adjustment method. Therefore, we required a P-value < 0.0125.

However, stratification by TDI can introduce collider bias when there is a conditional relationship between the genetic risk score and TDI (e.g., the mean GRS score is higher in either the low or high TDI group). As explained in detail previously in *Coscia et al*, when a variable (TDI) in a causal diagram is directly affected by two other variables, such as the risk factor (BMI, blood pressure, LDL cholesterol and triglyceride levels) and corresponding GRS, conditioning on TDI might introduce a collider (**Figure 1**) [32]. In line with this paper, in a sensitivity analysis, we controlled for the possible presence of such bias in the main analyses, by defining strata defined by quantiles of the residual TDI collider. We calculated the residual TDI, a variable that is free from any influences of the instrument variable (the GRS) [32], by calculating residuals using linear regression analyses with TDI as outcome and the genetic risk score as independent variable. Using the residuals, new subgroups were defined based on quartiles, and the main analyses were repeated accordingly.

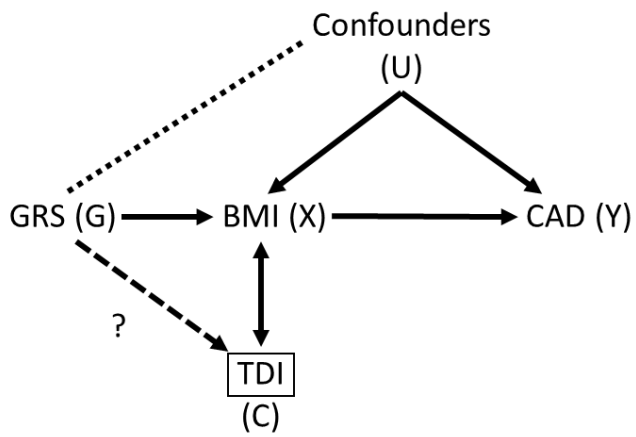

Figure 1. Directed Acyclic Graph (DAG) illustrating the relationship between the studied variables. When variable G (GRS) and variable C (TDI) are directly linked, TDI can be considered a collider variable, becoming a dependent variable when conditioned on.

RESULTS

Baseline characteristics of the UK biobank study population

When stratified in quartiles for TDI score, the study sample (**Table 1**) consisted of 446,495 individuals of which 52,946 (12%) had CAD. Participants in the group with the highest TDI scores had a higher mean measured BMI (28.0 versus 27.0 kg/m^2) and a lower mean age (55.9 versus 57.3 years) compared with the lowest TDI group.

Mendelian Randomization analyses

In the lowest TDI group, around 10% of the participants (N = 11,526) developed CAD prior or during follow-up. With a higher TDI, the percentage of participants with CAD increased to around 14% (N= 16,158) in the highest TDI group.

The logistic regression models in our MR analyses without stratification by TDI, showed that a one SD increase in genetically-determined BMI (OR: 1.107 [95% CI: 1.096, 1.117]), systolic blood pressure (OR: 1.068 [95% CI: 1.058, 1.078]), LDL cholesterol (OR: 1.086 [95% CI: 1.077, 1.097]) and triglycerides (OR: 1.053 [95% CI: 1.044, 1.063]) were all associated with a higher risk of CAD.

In the stratified analyses, we observed that the effect estimate of CAD by genetically-determined BMI increased as TDI increased (**Figure 2A**). The OR for CAD per SD increase in genetically-influenced BMI was 1.081 (95% CI: 1.059, 1.103) in the lowest TDI group versus 1.126 (95% CI: 1.106, 1.145) in the highest TDI group. Using a logistic model, interaction analyses showed that, after correcting for multiple testing, the OR of CAD per SD increase in BMI differed significantly between TDI groups (p-value for interaction = 0.0049). For systolic blood pressure, LDL cholesterol and triglyceride levels, the OR for CAD per SD increase was similar in the different TDI groups (**Figure 2B-D**). In addition, interaction analyses for these variables showed no significant difference between groups (p-value for interaction = 0.27, for systolic blood pressure, 0.44 for LDL and 0.073 for triglyceride respectively).

In subsequent analyses where we further stratified the highest TDI group because of the large range in TDI values in this group, the OR for CAD per SD increase in genetically-influenced BMI in the 75-87.5 percentile TDI group was 1.109 (95% CI: 1.080, 1.138), whereas in the group with TDI values above the 87.5 percentile the OR was 1.140 (95% CI: 1.113, 1.167).

Sensitivity analyses

In logistic regression analyses, only genetically-influenced BMI was associated with TDI. Therefore, the analysis that studied the association between genetically-influenced BMI and CAD, stratified for TDI was repeated using new BMI GRS_free subgroups of TDI ("IV free"). These results did not substantially differ from the main analysis (**Supplementary Table 1)**.

Table 1. Baseline Characteristics of the study population

	High SES (N = 111563)	Middle SES (N = 111427)	Low SES (N = 111488)	Very low SES (N= 111487)	Total (N= 446495)
Sex, Female (%)	60302 (54%)	60929 (55%)	61493 (55%)	59736 (54%)	242742 (54%)
Age (years)	57.3 (7.80)	57.2 (7.90)	56.6 (8.07)	55.9 (8.24)	56.8 (8.02)
Townsend Deprivation Index Score	-4.44 [-6.26,-3.71]	-3.01 [-3.70,-2.29]	-1.29 [-2.29, 0.210]	2.44 [0.211, 11.0]	-2.29 [6.26, 11.0]
BMI (kg/m²)	27.0 (4.33)	27.2 (4.49)	27.4 (4.74)	28.0 (5.34)	27.4 (4.75)
Systolic blood pressure (mmHg)	141 (19.6)	141 (19.6)	140 (19.6)	139 (19.7)	140 (19.7)
Diastolic blood pressure (mmHg)	82.3 (10.5)	82.3 (10.6)	82.2 (10.7)	82.0 (10.9)	82.2 (10.7)
LDL Cholesterol (mmol/l)	5.76 (1.13)	5.74 (1.14)	5.71 (1.14)	5.64 (1.16)	5.71 (1.14)
Triglycerides (mmol/l)	1.72 (0.980)	1.73 (0.998)	1.75 (1.03)	1.81 (1.09)	1.75(1.02)

Data are mean (SD) for normally distributed continuous variables or median (IQR) otherwise. Percentages are shown for dichotomous variables. Abbreviations: BMI, Body Mass Index.

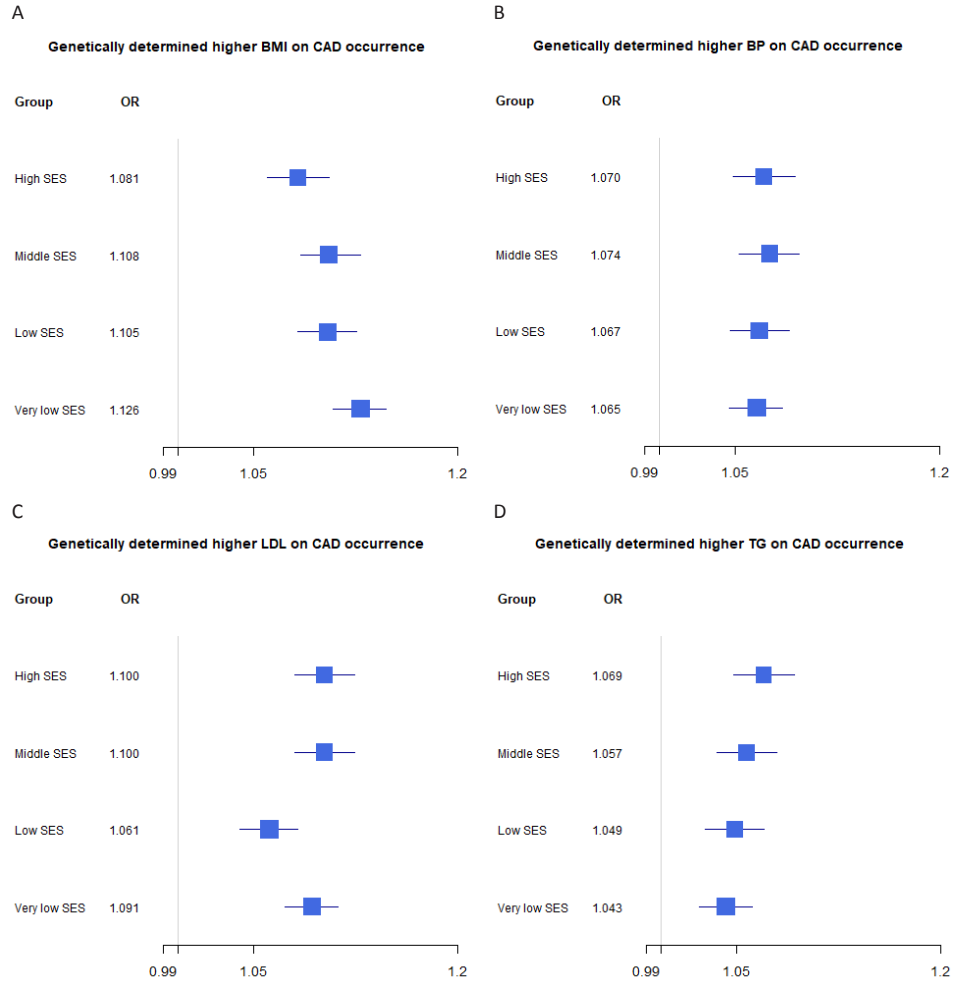

Figure 2. Association between genetically determined risk factors Body Mass Index (A), systolic blood pressure (B), LDL Cholesterol (C), triglyceride (D) and coronary artery disease stratified for Socioeconomic Status

Estimated ORs represent the effect per SD increase in risk factor GRS on CAD. Results obtained using a logistic regression with genetic risk score as exposure, corrected for age, sex, and the first 10 principal components and were stratified for SES. Abbreviations: BMI, Body Mass Index; SBP, systolic blood pressure; TG, triglyceride; SES, Socioeconomic Status.

DISCUSSION

We performed MR analyses using calculated genetic risk scores for CVD risk factors to investigate their association with CAD in different SES groups, using data from 446,495 European-ancestry participants from the UK Biobank. Results indicated that in every SES group, each investigated genetically-influenced risk factor (BMI, SBP, LDL, triglycerides) was associated with an increased risk of CAD, confirming the previously observed effects of these risk factors on CAD. However, for genetically-influenced BMI, the observed effect on CAD differed between SES groups. Specifically, in the lower SES group the increased risk on CAD per SD increase BMI was larger compared with the highest SES group. These results could be an indication that an increased BMI is not only more prevalent in low SES groups, but that the risk associated with a one-unit increased BMI is also higher.

To the best of our knowledge, the current study is the first to investigate the impact of classic CVD risk factors on CAD occurrence in different sub-groups of SES in a MR analysis. Of interest, earlier Mendelian Randomization studies have shown age-specific effects attributable to CVD risk factors [17, 18]. These findings, together with the findings from the present study, further illustrate that the effect of CVD risk factors is not homogeneous, but subgroup specific instead.

To illustrate, a hypothetical intervention in our study population leading to an equal reduction (e.g., 1 SD) in BMI across all groups, would lead to a relatively larger case reduction of CAD in the lower SES group. A 1 SD reduction of BMI would lead to 1.081 lower odds of developing CAD attributable to BMI in the highest SES group. Conversely, the same 1 SD reduction would lead to 1.126 lower odds in the lowest SES group. With both a larger incidence and a bigger effect size, a 1 SD reduction of BMI leads to a larger absolute case reduction in the low SES group. By not recognizing the different effect attributable to BMI in different SES groups and using the overall increased risk (OR: 1.107), a hypothetical intervention would assume an under- and overestimation of case-reduction in high and low SES groups respectively.

Our findings on the subgroup-specific impact of BMI on CAD risk could be caused by different body compositions in different socioeconomic groups. BMI is generally thought to have a linear relationship with CVD incidence [33]. However, it is a measure of overall adiposity that takes into account body weight and length, but not body composition variables such as fat mass and muscle mass. It is possible that the increased CAD risk in groups of lower SES could be due to the fact that groups of lower SES have a higher body fat percentage. However, there is currently not much literature on body composition in different socioeconomic groups. The distribution of fat is another aspect to consider. Literature has shown that compared with subcutaneous fat, visceral fat is associated with a higher risk for CAD [34]. It could be hypothesized that the body fat of individuals of lower SES consists of a larger proportion of visceral fat than the body fat of individuals of higher SES. As there are, to our knowledge, currently not sufficient reliable genetic instruments for visceral fat, we were unable to test this hypothesis. Thus, subsequent studies should aim to explore the potential differences in body composition between SES groups.

It has been shown that low SES is one of the strongest predictors towards engaging in lifestyle risk behavior associated with cardiovascular death [35]. These include smoking, alcohol consumption and an unhealthy diet. All of these lifestyle factors are in turn associated with increased liver fat and/or visceral fat, which are known to increase CAD risk [36, 37]. It is therefore possible that lifestyle risk behavior could lead to different body compositions between SES groups, which in turn could explain our results.

Although SBP, triglyceride and LDL cholesterol are assumed to be causal risk factors for CAD incidence, there does not seem to be a difference in effect between SES groups according to our results. Thus, it is likely that interventions targeting either SBP, triglyceride or LDL cholesterol would have a comparable effect on CAD incidence, independent of SES.

The main strength of this study is the large sample size as well as considerable number of CAD cases. This ensured statistical power for our analyses on the association between CVD risk factors and CAD occurrence. The MR method also aims to prevent possible reverse causation or confounding. Finally, our findings on the associations between known CVD risk factors and CAD are directionally consistent compared with earlier literature, which increases the credibility of our main findings. Some limitations should also be considered. First, we used the Townsend Deprivation Index as an indication of SES. As TDI is only measured at baseline, potential changes in SES during follow-up cannot be taken into account. To add, TDI is calculated based on geographical data and therefore is not a measure of individual SES, but a measure of environmental poverty. Furthermore, using a measure of neighborhood SES could provide suitable target locations for potential tailor-made intervention policies. Second, our study population from the UK Biobank consists of Caucasian participants. Therefore, the generalizability of our results to other ancestry groups is limited. This is especially relevant as prevalence of CVD risk factors differ between ethnic groups [38]. However, limiting the study population to Caucasians greatly reduces the heterogeneity between participants.

In conclusion, our findings indicate that CAD risk attributable to BMI is not homogenous and is modified by SES. Although genetically-influenced BMI was associated with CAD in all SES subgroups, tailor-made approaches for risk reduction dependent on SES should be considered to optimize the reduction in disease risk.

Acknowledgements

The present study has been conducted using the UK Biobank Resource (Application Number 56340) that is available to researchers.

Funding
This study was supported by an innovation grant from the Dutch Heart Foundation (grant number 2019T103; to R Noordam).

Disclosures
None

Conflicts of interest statement
None declared.

REFERENCES

1. Kaplan, G.A. and J.E. Keil, *Socioeconomic factors and cardiovascular disease: a review of the literature.* Circulation, 1993. **88**: p. 1973-98.
2. Mackenbach, J.P., A.E. Cavelaars, A.E. Kunst, and F. Groenhof, *Socioeconomic inequalities in cardiovascular disease mortality; an international study.* Eur Heart J, 2000. **21**: p. 1141-51.
3. Khaing, W., S.A. Vallibhakara, J. Attia, M. McEvoy, and A. Thakkinstian, *Effects of education and income on cardiovascular outcomes: A systematic review and meta-analysis.* Eur J Prev Cardiol, 2017. **24**: p. 1032-1042.
4. Schultz, W.M., H.M. Kelli, J.C. Lisko, T. Varghese, J. Shen, P. Sandesara, A.A. Quyyumi, H.A. Taylor, M. Gulati, J.G. Harold, et al., *Socioeconomic Status and Cardiovascular Outcomes.* Circulation, 2018. **137**: p. 2166-2178.
5. Hiscock, R., L. Bauld, A. Amos, J.A. Fidler, and M. Munafo, *Socioeconomic status and smoking: a review.* Ann N Y Acad Sci, 2012. **1248**: p. 107-23.
6. Kelly, L.A., J.J. Reilly, A. Fisher, C. Montgomery, A. Williamson, J.H. McColl, J.Y. Paton, and S. Grant, *Effect of socioeconomic status on objectively measured physical activity.* Arch Dis Child, 2006. **91**: p. 35-8.
7. Darmon, N. and A. Drewnowski, *Contribution of food prices and diet cost to socioeconomic disparities in diet quality and health: a systematic review and analysis.* Nutr Rev, 2015. **73**: p. 643-60.
8. Molarius, A., *The contribution of lifestyle factors to socioeconomic differences in obesity in men and women--a population-based study in Sweden.* Eur J Epidemiol, 2003. **18**: p. 227-34.
9. O'Donnell, C.J. and R. Elosua, *[Cardiovascular risk factors. Insights from Framingham Heart Study].* Rev Esp Cardiol, 2008. **61**: p. 299-310.
10. Diez-Roux, A.V., F.J. Nieto, C. Muntaner, H.A. Tyroler, G.W. Comstock, E. Shahar, L.S. Cooper, R.L. Watson, and M. Szklo, *Neighborhood environments and coronary heart disease: a multilevel analysis.* Am J Epidemiol, 1997. **146**: p. 48-63.
11. Colhoun, H.M., H. Hemingway, and N.R. Poulter, *Socio-economic status and blood pressure: an overview analysis.* J Hum Hypertens, 1998. **12**: p. 91-110.
12. Morgenstern, M., J.D. Sargent, and R. Hanewinkel, *Relation between socioeconomic status and body mass index: evidence of an indirect path via television use.* Archives of pediatrics & adolescent medicine, 2009. **163**: p. 731-738.
13. Streel, S., A.-F. Donneau, A. Hoge, S. Majerus, P. Kolh, J.-P. Chapelle, A. Albert, and M. Guillaume, *Socioeconomic Impact on the Prevalence of Cardiovascular Risk Factors in Wallonia, Belgium: A Population-Based Study.* BioMed Research International, 2015. **2015**: p. 580849.
14. Carter, A.R., D. Gill, N.M. Davies, A.E. Taylor, T. Tillmann, J. Vaucher, R.E. Wootton, M.R. Munafò, G. Hemani, R. Malik, et al., *Understanding the consequences of education inequality on cardiovascular disease: mendelian randomisation study.* Bmj, 2019. **365**: p. l1855.
15. Pucci, G., R. Alcidi, L. Tap, F. Battista, F. Mattace-Raso, and G. Schillaci, *Sex- and gender-related prevalence, cardiovascular risk and therapeutic approach in metabolic syndrome: A review of the literature.* Pharmacol Res, 2017. **120**: p. 34-42.
16. Uijl, A., S. Koudstaal, K. Direk, S. Denaxas, R.H.H. Groenwold, A. Banerjee, A.W. Hoes, H. Hemingway, and F.W. Asselbergs, *Risk factors for incident heart failure in age- and sex-specific strata: a population-based cohort using linked electronic health records.* Eur J Heart Fail, 2019. **21**: p. 1197-1206.
17. Noordam, R., K. Lall, R.A.J. Smit, T. Laisk, T. Estonian Biobank Research, A. Metspalu, T. Esko, L. Milani, R.J.F. Loos, R. Magi, et al., *Stratification of Type 2 Diabetes by Age of Diagnosis in the UK Biobank Reveals Subgroup-Specific Genetic Associations and Causal Risk Profiles.* Diabetes, 2021. **70**: p. 1816-1825.
18. Jansen, S.A., B. Huiskens, S. Trompet, J. Jukema, S.P. Mooijaart, K. Willems van Dijk, D. van Heemst, and R. Noordam, *Classical risk factors for primary coronary artery disease from an aging perspective through Mendelian Randomization.* Geroscience, 2021.
19. Lawlor, D.A., R.M. Harbord, J.A. Sterne, N. Timpson, and G. Davey Smith, *Mendelian randomization: using genes as instruments for making causal inferences in epidemiology.* Stat Med, 2008. **27**: p. 1133-63.
20. Burgess, S., A. Butterworth, and S.G. Thompson, *Mendelian randomization analysis with multiple genetic variants using summarized data.* Genet Epidemiol, 2013. **37**: p. 658-65.
21. Bowden, J., G. Davey Smith, P.C. Haycock, and S. Burgess, *Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator.* Genet Epidemiol, 2016. **40**: p. 304-14.

22. Haycock, P.C., S. Burgess, K.H. Wade, J. Bowden, C. Relton, and G. Davey Smith, *Best (but oft-forgotten) practices: the design, analysis, and interpretation of Mendelian randomization studies.* Am J Clin Nutr, 2016. **103**: p. 965-78.
23. Sudlow, C., J. Gallacher, N. Allen, V. Beral, P. Burton, J. Danesh, P. Downey, P. Elliott, J. Green, M. Landray, et al., *UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age.* PLoS Med, 2015. **12**: p. e1001779.
24. Locke, A.E., B. Kahali, S.I. Berndt, A.E. Justice, T.H. Pers, F.R. Day, C. Powell, S. Vedantam, M.L. Buchkovich, J. Yang, et al., *Genetic studies of body mass index yield new insights for obesity biology.* Nature, 2015. **518**: p. 197-206.
25. Willer, C.J., E.M. Schmidt, S. Sengupta, G.M. Peloso, S. Gustafsson, S. Kanoni, A. Ganna, J. Chen, M.L. Buchkovich, S. Mora, et al., *Discovery and refinement of loci associated with lipid levels.* Nat Genet, 2013. **45**: p. 1274-1283.
26. Ehret, G.B., P.B. Munroe, K.M. Rice, M. Bochud, A.D. Johnson, D.I. Chasman, A.V. Smith, M.D. Tobin, G.C. Verwoert, S.J. Hwang, et al., *Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk.* Nature, 2011. **478**: p. 103-9.
27. Walter, K., J.L. Min, J. Huang, L. Crooks, Y. Memari, S. McCarthy, J.R.B. Perry, C. Xu, M. Futema, D. Lawson, et al., *The UK10K project identifies rare variants in health and disease.* Nature, 2015. **526**: p. 82-90.
28. Auton, A., G.R. Abecasis, D.M. Altshuler, R.M. Durbin, G.R. Abecasis, D.R. Bentley, A. Chakravarti, A.G. Clark, P. Donnelly, E.E. Eichler, et al., *A global reference for human genetic variation.* Nature, 2015. **526**: p. 68-74.
29. McCarthy, S., S. Das, W. Kretzschmar, O. Delaneau, A.R. Wood, A. Teumer, H.M. Kang, C. Fuchsberger, P. Danecek, K. Sharp, et al., *A reference panel of 64,976 haplotypes for genotype imputation.* Nat Genet, 2016. **48**: p. 1279-83.
30. Townsend, P., *Deprivation.* Journal of Social Policy, 1987. **16**: p. 125-146.
31. Hemani, G., J. Zheng, B. Elsworth, K.H. Wade, V. Haberland, D. Baird, C. Laurin, S. Burgess, J. Bowden, R. Langdon, et al., *The MR-Base platform supports systematic causal inference across the human phenome.* Elife, 2018. **7**.
32. Coscia, C., D. Gill, R. Benitez, T. Perez, N. Malats, and S. Burgess, *Avoiding collider bias in Mendelian randomization when performing stratified analyses.* Eur J Epidemiol, 2022.
33. Zhao, Y., R. Qie, M. Han, S. Huang, X. Wu, Y. Zhang, Y. Feng, X. Yang, Y. Li, Y. Wu, et al., *Association of BMI with cardiovascular disease incidence and mortality in patients with type 2 diabetes mellitus: A systematic review and dose-response meta-analysis of cohort studies.* Nutr Metab Cardiovasc Dis, 2021. **31**: p. 1976-1984.
34. Nakamura, T., K. Tokunaga, I. Shimomura, M. Nishida, S. Yoshida, K. Kotani, A.H. Islam, Y. Keno, T. Kobatake, Y. Nagai, et al., *Contribution of visceral fat accumulation to the development of coronary artery disease in non-obese men.* Atherosclerosis, 1994. **107**: p. 239-46.
35. Meader, N., K. King, T. Moe-Byrne, K. Wright, H. Graham, M. Petticrew, C. Power, M. White, and A.J. Sowden, *A systematic review on the clustering and co-occurrence of multiple risk behaviours.* BMC Public Health, 2016. **16**: p. 657.
36. Koch, M., J. Borggrefe, S. Schlesinger, J. Barbaresko, G. Groth, G. Jacobs, W. Lieb, M. Laudes, M.J. Muller, A. Bosy-Westphal, et al., *Association of a lifestyle index with MRI-determined liver fat content in a general population study.* J Epidemiol Community Health, 2015. **69**: p. 732-7.
37. Molenaar, E.A., J.M. Massaro, P.F. Jacques, K.M. Pou, R.C. Ellison, U. Hoffmann, K. Pencina, S.D. Shadwick, R.S. Vasan, C.J. O'Donnell, et al., *Association of lifestyle factors with abdominal subcutaneous and visceral adiposity: the Framingham Heart Study.* Diabetes Care, 2009. **32**: p. 505-10.
38. Taylor, H.A., Jr., S.A. Coady, D. Levy, E.R. Walker, R.S. Vasan, J. Liu, E.L. Akylbekova, R.J. Garrison, and C. Fox, *Relationships of BMI to cardiovascular risk factors differ by ethnicity.* Obesity (Silver Spring), 2010. **18**: p. 1638-45.

6

GENERAL DISCUSSION

GENERAL DISCUSSION

Despite successful prevention and treatment options, cardiovascular diseases (CVD) remain a leading cause of death worldwide. This thesis investigated potential mechanisms leading to CVD in order to characterize existing and identify novel targets for future preventive strategies to reduce CVD incidence and associated disease burden. For this aim, the causes of mitochondrial dysfunction, resulting in oxidative stress, and the cerebrovascular consequences in the general population were assessed. Additionally, the causality of risk factors for CVD were studied and possible differences in subgroups of the population were investigated to determine the possibility of more personalized cardiovascular disease risk prediction. In the following paragraphs, the main findings as well as their implications and future perspectives will be discussed.

MAIN FINDINGS

Antioxidants

High levels of reactive oxygen species (ROS) have previously been associated with higher risk of developing CVD [1-3]. As antioxidants are the natural scavengers of ROS, in **Chapter 2** we studied how lifestyle factors are associated with antioxidant levels, and its derivatives, in blood and urine. In a study sample of about 500 participants from the Netherlands Epidemiology of Obesity study [4], we measured serum vitamin E levels and urinary vitamin E metabolite levels as a reflection of antioxidant status. We observed that some lifestyle factors associate differently with vitamin E serum levels as compared to urinary vitamin E metabolite levels. For example, smoking was associated with higher urinary vitamin E metabolite levels, but not with serum vitamin E levels. These findings can be interpreted in the context of higher oxidative stress, which has been associated with smoking behavior [5-7], resulting in the increased turnover of serum vitamin E to urinary metabolites. However, oxidative stress apparently does not result in lower serum vitamin E levels given the null association between smoking and vitamin E levels in serum. Other research in the same study sample showed similar results, where urinary vitamin E metabolites were associated with lower insulin resistance, but not serum vitamin E [8]. Together, these findings indicate that measuring only serum vitamin E levels is not an accurate reflection of antioxidant activity. Since serum vitamin E levels are apparently controlled differently compared with secreted vitamin E metabolites, this might also explain why increasing serum vitamin E levels through supplementation does not significantly reduce disease risk [9, 10]. These results raise the hypothesis that higher urinary vitamin E levels might be a better reflection of oxidative stress. However, this remains to be investigated in large prospective cohort studies in combination with large-scale causal inference studies.

As smoking is one of the major risk factors associated with (any) stroke [11, 12], our observed association between smoking and higher urinary vitamin E metabolite levels could speculate towards the hypothesis that smoking affects stroke pathology, at least partly, via increased oxidative stress levels. Several other studies have also linked smoking behavior to increased oxidative stress levels [5-7], but used different measures as reflection of increased oxidative stress levels. Although higher antioxidant intake has been associated with lower stroke risk in observational cohort studies [3, 13-15], the results of **Chapter 3** show that these associations are likely not causal. When studying disease risk, it is vital to provide evidence whether a risk factor is also a causal factor. However, studying causal factors for cardiovascular disease in an RCT setting is challenging due to the generally long duration of disease development and associated exceedingly high economic costs. In Mendelian Randomization studies, genetic variants are used as instrumental variables to approximate a causal association between an exposure and an outcome. This provides a cheaper, more practical alternative approach. In **Chapter 3**, we investigated both circulating antioxidant blood and metabolite levels and their association with ischemic stroke risk using MR. We found that none of 5 genetically-influenced diet-derived antioxidants (Vitamin E, Carotene, Lycopene, Ascorbate, and Retinol) were associated with ischemic stroke risk. In other words, no evidence was found for a possible causal association between diet-derived antioxidant levels and ischemic stroke. These results are in agreement with the findings that increasing circulating levels of antioxidants via supplementation do not reduce the risk of ischemic stroke in RCTs [16, 17]. A similar study investigating antioxidant levels and coronary heart disease using MR also found no evidence for a causal association [18], as having (partly) overlapping pathophysiological mechanisms. Recent reviews covering oxidative stress mechanisms have suggested that circulating antioxidant levels are indeed not reflective of actual antioxidant capacity [9, 10]. This indicates that increasing antioxidant levels via supplementation is clinically irrelevant for CVD risk reduction, at least in the general population.

Commonly suggested methods to reduce stroke risk are physical exercise and smoking cessation [19]. Although these methods do not increase circulating antioxidant levels, **chapter 2** of this thesis has shown they are associated with altered antioxidant metabolites. This relation again highlights the difference between serum antioxidant levels and antioxidant capacity and might reveal why habitual exercise reduces stroke risk but antioxidant supplementation does not.

Oxidants

To investigate both sides of the oxidative stress mechanism and its relation to CVD, we shifted our focus from antioxidants to oxidants, in the form of ROS. Mitochondria are a major source of ROS production and have evolved elaborate means to scavenge these ROS. Mitochondrial dysfunction may thus result in increased ROS production,

and is therefore considered to be a driver and hallmark of the ageing process [20]. Mitochondrial dysfunction can be proxied using the mtDNA copy number (mtDNA-CN). MtDNA-CN can be assessed by estimating the mtDNA abundance relative to genomic DNA. Therefore, we aimed to study mtDNA abundance, as a proxy for mitochondrial dysfunction, and the effect it might have on the development of cardiovascular diseases such as stroke (**Chapter 4**).

In our prospective analyses, after correcting for potential confounders, we observed no association between mtDNA abundance and stroke. Although extensive sensitivity analyses provided weak evidence for lower mtDNA abundance as a causal factor for ischemic stroke, these results were not in line with the prospective multivariable-adjusted cross sectional analyses. Furthermore, the evidence favoring a possible relationship was only found after the exclusion of pleiotropic SNPs associated with platelet count, which were half of the originally included SNPs. Therefore, caution should be taken when interpreting these results. Further studies will have to shed light on the nature of this type of pleiotropy. One study did report an association between low mtDNA-CN and increased risk of incident stroke [21]. However, we could not replicate these results, while using a much larger sample size and applying triangulation by adding MR analyses in over 1 million participants.

Although mitochondrial dysfunction is thought to increase ROS production, which could damage surrounding cell structures [22], it is questionable whether the effect of increased mitochondrial dysfunction as measured in leukocytes reflects damage induced by ROS that could lead to stroke. For instance, mitochondrial dysfunction in endothelial cells might result in ROS-induced cell damage that lies within the direct causal pathway of stroke development. One study found great variation in mtDNA abundance between liver, kidney, brain, lung, muscle and heart samples, which was interpreted as a reflection of tissue-specific differences in mitochondrial activity [23]. However, another study has shown that there is a correlation between leukocyte mitochondrial dysfunction and metabolic health of other, different tissues [24]. In this paper, mitochondrial dysfunction measured in blood was shown to be predictive for neurodegenerative disease incidence [24], indicating that leukocyte mtDNA abundance could be used to study cerebrovascular disease onset. Importantly, tissue-specific mtDNA abundance itself could be a marker for other health conditions. Obesity causes increased mitochondrial dysfunction in multiple tissues [25, 26], which may be reflected by lower mtDNA abundance. However, a recent study has shown that there was no causal association between genetically-influenced BMI and mtDNA-CN in a large sample from European-ancestry participants from UK Biobank. [27]. Thus, future research will have to disentangle the exact causes of mitochondrial dysfunction, and how mitochondrial dysfunction is linked to adverse health outcomes. In order to continue studying the potential contribution of mitochondrial dysfunction to stroke risk, investigating relevant measurement data, such as mitochondrial dysfunction, from endothelial cells might be the next logical step. However, as blood samples tend to be more readily available, discovering and analyzing blood biomarkers that reflect endothelial dysfunction might be a more practical approach.

RISK MODIFICATION IN SUB-POPULATIONS

In **Chapter 5**, we studied whether established risk factors of CVD exert different effects in specific sub-populations. It is known that inhabitants of lower SES neighborhoods have an increased risk of CVD development [28-31]. Our results showed that the increased risk for CAD attributable to increased BMI can be modified by SES. In other words, an increased genetically-influenced BMI resulted in a higher risk on CAD in low SES groups when compared with the same genetically-influenced BMI increase in high SES groups. Therefore, when tackling obesity as a strategy for CAD prevention, a sub-group specific approach taking SES into account might lead to a proportionally higher reduction in CAD cases and should be considered. However, this approach has some drawbacks. From a social perspective, there might be some ethical considerations. Placing individuals, or city areas, into groups of high or low SES and targeting them with different programs might be stigmatizing [32, 33]. Additionally, as BMI was still associated with increased stroke risk in all SES groups, a separate approach may wrongly convey that the high SES group is devoid from disease risk. Nonetheless, from a scientific point of view it might be relevant to continue to study known associations in greater detail by investigating subgroup specific effects. In the end, by optimizing correct risk allocation, the results of these studies could aid in future individualized treatment options.

REFLECTION ON MAIN FINDINGS

Study population characteristics are one of the main considerations to be taken into account when reflecting on the results of this thesis. Since the research in this thesis used European-ancestry cohorts, it may not be generalizable to other (non-European) populations. Therefore, it is relevant to replicate these studies in datasets of non-European ancestry. In addition, the results from **Chapter 5** demonstrated that individuals of low SES have an increased CVD risk. This group is generally underrepresented in population research, including in the UK Biobank. This indicates that the highest risk group in these studies is simultaneously the least accurately represented in the study population [34, 35]. Therefore, it is possible that our results are an underestimation of the difference between SES groups, as the lowest SES group of the general population might not be fully represented by the lowest SES group of our study population. Furthermore, participants of the UK biobank are thought to be relatively healthy [34]. Participants were between 40 and 70 years of age at inclusion of the study. Although the

UK biobank started in 2005, this can still be considered relatively young to study stroke incidence. This is also reflected in our relatively low number of cases (2%), compared with the general life-time stroke incidence of 25% [36]. As a consequence, a number of participants currently in the control group might become cases when more follow-up years are completed.

Since traditional observational association studies cannot determine causal effects, we combined multivariable-adjusted analyses with MR. However, MR has several underlying assumptions. One of these is that MR assesses the association using the assumed life-long exposure of the determinant in its genetic instrument, whereas the data from the studied population reflects only the outcomes up until the point of data collection. There could be uncertainty as to whether outcomes such as disease occurrence are altered in the future. Increasing follow-up years of the study cohort should decrease the likelihood of this possibility, although the final conclusion might not be drawn for years to come. Additionally, the genetic instruments used in these studies accounted for a relatively low amount of variation of the outcome. Possibly a future GWAS could provide stronger instruments. Alternatively, it is imaginable that genetically derived oxidative stress related to stroke incidence is only activated in combination with adverse lifestyle and environmental factors, which can be assessed by future studies of gene-environment interactions.

CONCLUSIONS AND FUTURE PERSPECTIVES

In this thesis, we observed that lifestyle factors smoking and physical exercise are associated with urinary vitamin E metabolite levels, which we hypothesize to be a marker of antioxidant activity. However, we did not find evidence supporting a causal association between either antioxidant levels or mtDNA abundance and stroke risk. Adiposity and glucose homeostasis were found to be associated with urinary vitamin E metabolites, but not with circulating vitamin E blood levels [8, 37]. Combined with the lack of success of vitamin E supplementation in clinical trials [38, 39], these results support the hypothesis that vitamin E levels do not reflect antioxidant status or capacity. In other words, the rate limiting step in the scavenging process of antioxidants such as vitamin E is not the circulating blood levels but rather a downstream process. Discovering this rate limiting step will be crucial in order to study the real effect of antioxidants on cardiovascular disease risk.

A partial contributor to the observed null-findings could be the studied population. Therefore, if these studies were to be repeated, and replicated, in the future, an older study population might provide more accurate results. Nevertheless, the current associations are still representative for this study population, taking into account age group and follow-up duration. Additionally, the study populations used in this thesis were very large. Even after studying up to 1 million participants, no association between determinants and outcomes was observed. Although by increasing the study population statistical significance could be achieved, the found effect would be so small that clinical (and biological) relevance is highly unlikely. Furthermore, our results show that lower SES groups have worse CVD outcomes with identical risk factors. Future studies would have to take into account that their results might be biased towards a healthier volunteer population. Generally, this would lead to an underestimation of the real population risk factor prevalence, as those with on average worse health variables will be underrepresented [40, 41]. However, even with having a bias towards a healthier population, any found correlation within the study population is likely to be confounded by bias. A study has shown that within-subject analyses are not affected by so-called non-response bias [41]. Nonetheless, new recruitment should focus on obtaining a participant distribution to better reflect the general population.

The scientific field continues to evolve and novel ways to accurately measure oxidative stress are being developed. Therefore, it would be premature to state that the results from this thesis entirely answer the question whether oxidative stress is causally associated with stroke. However, similar studies using the same study population, but investigating different phenotypes reflective of higher oxidative stress levels, found similar results. For instance, genetically determined antioxidant levels were not causally associated with coronary heart disease [18]. Additionally, no evidence was found for the causal association between low mtDNA copy numbers and type 2 diabetes risk [27].

As science becomes increasingly data-driven and larger datasets become (publicly) available, future studies should be able to carry out more in-depth research, increasing the potential to provide conclusive evidence of the relation between oxidative stress and cardiovascular disease. A higher number of participants would not only increase the number of cases, but additionally allow for the investigation of subgroup specific effects. Although the current studied population in itself was considerably large, our results in **Chapter 5** have shown that there could be relevant differences in risk effect between two subgroups within a specific population. A previous study has presented similar results where CAD risk attenuated with age [42]. Therefore, it would be naïve to treat such a case group as homogeneous. Future studies could focus on true personalization of risk by further disentangling sub-group specific effects. In turn, this would open up possibilities for individualized medicine and prevention strategies.

This thesis provides novel insights in the relationship between oxidative stress and CVD using both biological and genetic data. Although associations were established, no conclusive evidence supporting a causal association between oxidative stress and CVD was found. However, as CVD is thought to be a complex and heterogenous disease, it is important to keep in mind that there are likely multiple pathways that together affect the progression of this disease. Our results already showed a complex combined relationship between genetically-influenced BMI, low SES, and CVD. When taking into

account the long onset of CVD, future research demands a more sophisticated approach in order to explain the underlying mechanisms of CVD development. For example, data on risk factors should be gathered throughout a person's lifespan, starting from a relatively young age. This would result in a better visualization of the true exposure per risk factor for each participant. This could be combined with individual genetically determined risk factors. Individuals who are at increased risk would then be invited for periodic screening. Although this method might be initially expensive, if effective, the reduction of CVD cases and patients would ultimately lower the burden on our healthcare system [43, 44]. Consequently, in order to truly combat the global healthcare burden caused by CVD, interventions might have to aim for a younger target population and continuous monitoring of risk factors over the life course.

REFERENCES

1. Shirley, R., E.N. Ord, and L.M. Work, *Oxidative Stress and the Use of Antioxidants in Stroke.* Antioxidants (Basel), 2014. **3**: p. 472-501.
2. Allen, C.L. and U. Bayraktutan, *Oxidative stress and its role in the pathogenesis of ischaemic stroke.* Int J Stroke, 2009. **4**: p. 461-70.
3. Rodrigo, R., R. Fernandez-Gajardo, R. Gutierrez, J.M. Matamala, R. Carrasco, A. Miranda-Merchak, and W. Feuerhake, *Oxidative stress and pathophysiology of ischemic stroke: novel therapeutic opportunities.* CNS Neurol Disord Drug Targets, 2013. **12**: p. 698-714.
4. de Mutsert, R., M. den Heijer, T.J. Rabelink, J.W. Smit, J.A. Romijn, J.W. Jukema, A. de Roos, C.M. Cobbaert, M. Kloppenburg, S. le Cessie, et al., *The Netherlands Epidemiology of Obesity (NEO) study: study design and data collection.* Eur J Epidemiol, 2013. **28**: p. 513-23.
5. Karademirci, M., R. Kutlu, and I. Kilinc, *Relationship between smoking and total antioxidant status, total oxidant status, oxidative stress index, vit C, vit E.* Clin Respir J, 2018. **12**: p. 2006-2012.
6. Lourenco, M.A.M., M.G. Braz, A.G. Aun, B.L.B. Pereira, F.H. Fernandes, E.M. Kazmarek, T.F. Bachiega, S.G. Zanati, P.S. Azevedo, B.F. Polegato, et al., *Lipid damage is the best marker of oxidative injury during the cardiac remodeling process induced by tobacco smoke.* BMC Pharmacol Toxicol, 2018. **19**: p. 74.
7. Kosecik, M., O. Erel, E. Sevinc, and S. Selek, *Increased oxidative stress in children exposed to passive smoking.* Int J Cardiol, 2005. **100**: p. 61-4.
8. Luo, J., F.L. Meulmeester, L.G. Martens, N. Ashrafi, R. de Mutsert, D.O. Mook-Kanamori, F.R. Rosendaal, K. Willems van Dijk, S. le Cessie, K. Mills, et al., *Urinary oxidized, but not enzymatic vitamin E metabolites are inversely associated with measures of glucose homeostasis in middle-aged healthy individuals.* Clin Nutr, 2021. **40**: p. 4192-4200.
9. Luo, J., K. Mills, S. le Cessie, R. Noordam, and D. van Heemst, *Ageing, age-related diseases and oxidative stress: What to do next?* Ageing Res Rev, 2020. **57**: p. 100982.
10. Luo, J., Y. Hashimoto, L.G. Martens, F.L. Meulmeester, N. Ashrafi, D.O. Mook-Kanamori, F.R. Rosendaal, J.W. Jukema, K.W. van Dijk, K. Mills, et al., *Associations of metabolomic profiles with circulating vitamin E and urinary vitamin E metabolites in middle-aged individuals.* Nutrition, 2022. **93**: p. 111440.
11. Virani, S.S., A. Alonso, H.J. Aparicio, E.J. Benjamin, M.S. Bittencourt, C.W. Callaway, A.P. Carson, A.M. Chamberlain, S. Cheng, F.N. Delling, et al., *Heart Disease and Stroke Statistics-2021 Update: A Report From the American Heart Association.* Circulation, 2021. **143**: p. e254-e743.
12. Kernan, W.N., B. Ovbiagele, H.R. Black, D.M. Bravata, M.I. Chimowitz, M.D. Ezekowitz, M.C. Fang, M. Fisher, K.L. Furie, D.V. Heck, et al., *Guidelines for the prevention of stroke in patients with stroke and transient ischemic attack: a guideline for healthcare professionals from the American Heart Association/ American Stroke Association.* Stroke, 2014. **45**: p. 2160-236.
13. Yochum, L.A., A.R. Folsom, and L.H. Kushi, *Intake of antioxidant vitamins and risk of death from stroke in postmenopausal women.* Am J Clin Nutr, 2000. **72**: p. 476-83.
14. Sanchez-Moreno, C., A. Jimenez-Escrig, and A. Martin, *Stroke: roles of B vitamins, homocysteine and antioxidants.* Nutr Res Rev, 2009. **22**: p. 49-67.
15. Gillman, M.W., L.A. Cupples, D. Gagnon, B.M. Posner, R.C. Ellison, W.P. Castelli, and P.A. Wolf, *Protective effect of fruits and vegetables on development of stroke in men.* JAMA, 1995. **273**: p. 1113-7.
16. Leppala, J.M., J. Virtamo, R. Fogelholm, D. Albanes, P.R. Taylor, and O.P. Heinonen, *Vitamin E and beta carotene supplementation in high risk for stroke: a subgroup analysis of the Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study.* Arch Neurol, 2000. **57**: p. 1503-9.
17. Yusuf, S., G. Dagenais, J. Pogue, J. Bosch, and P. Sleight, *Vitamin E supplementation and cardiovascular events in high-risk patients.* N Engl J Med, 2000. **342**: p. 154-60.
18. Luo, J., S. le Cessie, D. van Heemst, and R. Noordam, *Diet-Derived Circulating Antioxidants and Risk of Coronary Heart Disease: A Mendelian Randomization Study.* J Am Coll Cardiol, 2021. **77**: p. 45-54.
19. Arnett, D.K., R.S. Blumenthal, M.A. Albert, A.B. Buroker, Z.D. Goldberger, E.J. Hahn, C.D. Himmelfarb, A. Khera, D. Lloyd-Jones, J.W. McEvoy, et al., *2019 ACC/AHA Guideline on the Primary Prevention of Cardiovascular Disease: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines.* Circulation, 2019. **140**: p. e596-e646.
20. Lopez-Otin, C., M.A. Blasco, L. Partridge, M. Serrano, and G. Kroemer, *The hallmarks of aging.* Cell, 2013. **153**: p. 1194-217.

21. Ashar, F.N., Y. Zhang, R.J. Longchamps, J. Lane, A. Moes, M.L. Grove, J.C. Mychaleckyj, K.D. Taylor, J. Coresh, J.I. Rotter, et al., *Association of Mitochondrial DNA Copy Number With Cardiovascular Disease.* JAMA Cardiol, 2017. **2**: p. 1247-1255.
22. Bandy, B. and A.J. Davison, *Mitochondrial mutations may increase oxidative stress: implications for carcinogenesis and aging?* Free Radic Biol Med, 1990. **8**: p. 523-39.
23. D'Erchia, A.M., A. Atlante, G. Gadaleta, G. Pavesi, M. Chiara, C. De Virgilio, C. Manzari, F. Mastropasqua, G.M. Prazzoli, E. Picardi, et al., *Tissue-specific mtDNA abundance from exome data and its correlation with mitochondrial transcription, mass and respiratory activity.* Mitochondrion, 2015. **20**: p. 13-21.
24. Yang, S.Y., C.A. Castellani, R.J. Longchamps, V.K. Pillalamarri, B. O'Rourke, E. Guallar, and D.E. Arking, *Blood-derived mitochondrial DNA copy number is associated with gene expression across multiple tissues and is predictive for incident neurodegenerative disease.* Genome Res, 2021. **31**: p. 349-358.
25. Skuratovskaia, D.A., J.K. Sofronova, P.A. Zatolokin, K.Y. Popadin, M.A. Vasilenko, L.S. Litvinova, and I.O. Mazunin, *Additional evidence of the link between mtDNA copy number and the body mass index.* Mitochondrial DNA A DNA Mapp Seq Anal, 2018. **29**: p. 1240-1244.
26. Yin, X., I.R. Lanza, J.M. Swain, M.G. Sarr, K.S. Nair, and M.D. Jensen, *Adipocyte mitochondrial function is reduced in human obesity independent of fat cell size.* J Clin Endocrinol Metab, 2014. **99**: p. E209-16.
27. Wang, W., J. Luo, K. Willems van Dijk, S. Hagg, F. Grassmann, T.H. LM, D. van Heemst, and R. Noordam, *Assessment of the bi-directional relationship between blood mitochondrial DNA copy number and type 2 diabetes mellitus: a multivariable-adjusted regression and Mendelian randomisation study.* Diabetologia, 2022.
28. Kaplan, G.A. and J.E. Keil, *Socioeconomic factors and cardiovascular disease: a review of the literature.* Circulation, 1993. **88**: p. 1973-98.
29. Khaing, W., S.A. Vallibhakara, J. Attia, M. McEvoy, and A. Thakkinstian, *Effects of education and income on cardiovascular outcomes: A systematic review and meta-analysis.* Eur J Prev Cardiol, 2017. **24**: p. 1032-1042.
30. Schultz, W.M., H.M. Kelli, J.C. Lisko, T. Varghese, J. Shen, P. Sandesara, A.A. Quyyumi, H.A. Taylor, M. Gulati, J.G. Harold, et al., *Socioeconomic Status and Cardiovascular Outcomes.* Circulation, 2018. **137**: p. 2166-2178.
31. Mackenbach, J.P., A.E. Cavelaars, A.E. Kunst, and F. Groenhof, *Socioeconomic inequalities in cardiovascular disease mortality; an international study.* Eur Heart J, 2000. **21**: p. 1141-51.
32. Phelan, J.C., J.W. Lucas, C.L. Ridgeway, and C.J. Taylor, *Stigma, status, and population health.* Soc Sci Med, 2014. **103**: p. 15-23.
33. MacLean, L., N. Edwards, M. Garrard, N. Sims-Jones, K. Clinton, and L. Ashley, *Obesity, stigma and public health planning.* Health Promot Int, 2009. **24**: p. 88-93.
34. Fry, A., T.J. Littlejohns, C. Sudlow, N. Doherty, L. Adamska, T. Sprosen, R. Collins, and N.E. Allen, *Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population.* Am J Epidemiol, 2017. **186**: p. 1026-1034.
35. Andreeva, V.A., B. Salanave, K. Castetbon, V. Deschamps, M. Vernay, E. Kesse-Guyot, and S. Hercberg, *Comparison of the sociodemographic characteristics of the large NutriNet-Sante e-cohort with French Census data: the issue of volunteer bias revisited.* J Epidemiol Community Health, 2015. **69**: p. 893-8.
36. Collaborators, G.B.D.L.R.o.S., V.L. Feigin, G. Nguyen, K. Cercy, C.O. Johnson, T. Alam, P.G. Parmar, A.A. Abajobir, K.H. Abate, F. Abd-Allah, et al., *Global, Regional, and Country-Specific Lifetime Risks of Stroke, 1990 and 2016.* N Engl J Med, 2018. **379**: p. 2429-2437.
37. Meulmeester, F.L., J. Luo, L.G. Martens, N. Ashrafi, R. de Mutsert, D.O. Mook-Kanamori, H.J. Lamb, F.R. Rosendaal, K. Willems van Dijk, K. Mills, et al., *Association of measures of body fat with serum alpha-tocopherol and its metabolites in middle-aged individuals.* Nutr Metab Cardiovasc Dis, 2021. **31**: p. 2407-2415.
38. Sesso, H.D., J.E. Buring, W.G. Christen, T. Kurth, C. Belanger, J. MacFadyen, V. Bubes, J.E. Manson, R.J. Glynn, and J.M. Gaziano, *Vitamins E and C in the prevention of cardiovascular disease in men: the Physicians' Health Study II randomized controlled trial.* JAMA, 2008. **300**: p. 2123-33.
39. Lee, I.M., N.R. Cook, J.M. Gaziano, D. Gordon, P.M. Ridker, J.E. Manson, C.H. Hennekens, and J.E. Buring, *Vitamin E in the primary prevention of cardiovascular disease and cancer: the Women's Health Study: a randomized controlled trial.* JAMA, 2005. **294**: p. 56-65.
40. Melton, L.J., 3rd, P.J. Dyck, J.L. Karnes, P.C. O'Brien, and F.J. Service, *Non-response bias in studies of diabetic complications: the Rochester Diabetic Neuropathy Study.* J Clin Epidemiol, 1993. **46**: p. 341-8.
41. Cheung, K.L., P.M. Ten Klooster, C. Smit, H. de Vries, and M.E. Pieterse, *The impact of non-response bias due to sampling in public health studies: A comparison of voluntary versus mandatory recruitment in a Dutch national survey on adolescent health.* BMC Public Health, 2017. **17**: p. 276.
42. Jansen, S.A., B. Huiskens, S. Trompet, J. Jukema, S.P. Mooijaart, K. Willems van Dijk, D. van Heemst, and R. Noordam, *Classical risk factors for primary coronary artery disease from an aging perspective through Mendelian Randomization.* Geroscience, 2022. **44**: p. 1703-1713.
43. Woolf, S.H., *The power of prevention and what it requires.* JAMA, 2008. **299**: p. 2437-9.
44. Cohen, J.T., P.J. Neumann, and M.C. Weinstein, *Does preventive care save money? Health economics and the presidential candidates.* N Engl J Med, 2008. **358**: p. 661-3.

7

NEDERLANDSE SAMENVATTING

Dankzij continue innovaties in onder andere de medische sector is de levensverwachting sinds 1800 gestaag gegroeid, met toenemende vergrijzing van de bevolking tot gevolg. Hierdoor is de prevalentie van leeftijdsgebonden ziekten zoals cardiovasculaire ziekten enorm gestegen. Hoewel cardiovasculaire ziekten al decennialang wordt onderzocht en er effectieve medicatie is gevonden, blijft het de belangrijkste doodsoorzaak wereldwijd. Daarnaast is leeftijd een van de grootste risicofactoren voor het ontwikkelen van nog vele chronische ziekten. Op biologisch niveau is het verouderingsproces onder te verdelen in een aantal specifieke kenmerken. Deze worden de 9 'Hallmarks of Ageing' genoemd. Mitochondriële dysfunctie is één van deze Hallmarks en ligt mogelijk aan de basis van verschillende negatieve aspecten van veroudering en verhoging van het risico op het krijgen van ziekten. Mitochondriën zijn een belangrijke bron van reactieve oxidatieve componenten ('Reactive Oxigen Species'; ROS), ook wel oxidanten genoemd. Deze worden aangemaakt als bijproduct tijdens de energieproductie. Hoewel een zekere hoeveelheid ROS noodzakelijk is voor een normale cel functie, kan een teveel aan ROS schade veroorzaken aan allerlei celstructuren. Bij mitochondriële dysfunctie ontstaat er te veel ROS. Om te voorkomen dat er een te hoge concentratie ROS ontstaat, beschikt het lichaam over antioxidanten die de reactieve oxidanten kunnen neutraliseren. Deze antioxidanten ruimen ROS op om deze binnen de optimale concentraties te houden voor normale cellulaire functie. Een goede balans tussen de oxidanten en antioxidanten is cruciaal om oxidatieve schade, ook wel oxidatieve stress genoemd, te voorkomen. Dit proefschrift beoogt de oorzaken van mitochondriële dysfunctie, een mogelijke oorzaak van oxidatieve stress, en de gevolgen van deze oxidatieve stress binnen de algemene populatie en in specifieke subpopulaties te onderzoeken.

In **hoofdstuk 2** hebben we onderzocht of vitamine E, een antioxidant, is geassocieerd met de leefstijlfactoren roken, slaap, bewegen, eten, en alcohol. Om de vitamine E status van een persoon te bepalen hebben we bij 530 mensen uit de Nederlandse Epidemiologie van Obesitas (NEO) studie zowel bloedwaarden van vitamine E als de metaboliet waarden van vitamine E in urine gemeten. In NEO waren roken, en te weinig of te veel bewegen geassocieerd met minder vitamine E metabolieten in urine. Opvallend is dat er geen associatie was met de vitamine E bloedwaardes. Deze resultaten laten zien dat alleen de vitamine E bloedwaardes niet voldoende inzicht geven in de antioxidant activiteit in het lichaam. Mogelijk verklaart deze bevinding ook waarom het gebruik van vitamine E supplementen niet effectief blijkt in het voorkomen van leeftijdsgebonden ziektes, een bevinding die ook eerder is gedaan in verschillende klinische studies, aangezien vitamine E supplementen enkel de bloedwaardes, en niet de metabolieten, verhoogt.

In **hoofdstuk 3** van dit proefschrift hebben we de mogelijke causale associatie onderzocht tussen verschillende dieet-afgeleide antioxidanten en het krijgen van een beroerte. Eerder epidemiologisch cohortonderzoek heeft aangetoond dat er een verband is tussen een hogere antioxidant concentratie van bijvoorbeeld vitamine E en een lager risico op het krijgen van een beroerte. Echter in observationele cohortstudies is het niet mogelijk om een oorzakelijk verband aan te tonen tussen antioxidant niveaus in bloed en het risico op een beroerte. Met de Mendeliaanse Randomisatie (MR) methode, waarbij gebruik wordt gemaakt van genetische varianten als proxy voor de onderzochte variabele, is het wel mogelijk om een causale associatie te onderzoeken. Door genetische varianten te gebruiken die samengaan met antioxidant niveaus in bloed, kun je de associaties relatief vrij van residuele verstoring of omgekeerde causaliteit onderzoeken. Bij het doen van dit onderzoek hebben wij gebruik gemaakt van de datasets MEGASTROKE, UK Biobank, en Finngen. In de gecombineerde populatie van meer dan 1 miljoen mensen, hebben we geen bewijs gevonden voor een oorzakelijk verband tussen dieet-afgeleide antioxidanten en het krijgen van een beroerte. Deze resultaten ondersteunen de hypothese dat het innemen van antioxidant supplementen niet bijdraagt aan de cardiovasculaire gezondheid in de algemene populatie.

Omdat naast antioxidanten, ook oxidanten belangrijk zijn bij het behouden van het oxidatieve evenwicht, wilden we in **hoofdstuk 4** kijken naar de associatie tussen mitochondriële dysfunctie, als oorzaak van verhoogde ROS productie, en beroertes. We hebben prospectief de relatie onderzocht tussen mitochondriële dysfunctie en het krijgen van een beroerte door gebruik te maken van beschikbare bloedwaardes van mensen die gevolgd zijn in de tijd. Na het analyseren van gegevens van bijna 300.000 participanten, hebben we geen associatie gevonden tussen mitochondriële dysfunctie en een beroerte. Gecombineerd met MR analyses, waarin we geen oorzakelijk verband tussen mitochondriële dysfunctie en een beroerte hebben kunnen aantonen, suggereren de resultaten dat er geen oorzakelijk verband is. Een mogelijke verklaring voor dit negatieve resultaat is dat de biomarker gebruikt voor dit onderzoek niet representatief is voor mitochondriële dysfunctie in relevante lichaamsdelen. Verder onderzoek zal moeten uitwijzen welke biomarkers hiervoor wel geschikt zijn.

Het is bekend dat inwoners van achterstandswijken, en dus met een lage sociaaleconomische status, een verhoogde kans hebben op het ontwikkelen van cardiovasculaire ziekten (CVD). Daarnaast weten we dat deze populatie meer bloot wordt gesteld aan factoren die oxidatieve stress kunnen verhogen zoals roken en een ongezond dieet. Zo heeft onderzoek aangetoond dat deze mensen gemiddeld een hoger BMI hebben. In **hoofdstuk 5** hebben we onderzocht of bekende risicofactoren van CVD, zoals een hoog BMI, een even sterk effect hebben op het ontwikkelen van CVD in mensen die wel of niet in een achterstandswijk wonen. Ondanks dat een hoog BMI in alle groepen geassocieerd was met een verhoogd risico voor het ontwikkelen van CVD, hebben we kunnen aantonen dat een hoog BMI, onafhankelijk van andere factoren, een groter risico geeft op het ontwikkelen van CVD als men in een achterstandswijk woont. Dit geeft aan dat BMI niet een eendimensionaal risico is, maar op een complexere manier bijdraagt aan de ontwikkeling van CVD. Ook suggereert het dat gepersonaliseerde interventies

nodig zijn in het bestrijden van CVD.

Dit proefschrift heeft antwoord proberen te geven op een aantal vragen omtrent oxidatieve stress en de associaties hiervan met verschillende cardiovasculaire ziekten. Hoewel in de gedane onderzoeken geen causale verbanden zijn gevonden, moeten we er rekening mee houden dat cardiovasculaire ziekten complex en heterogeen zijn. Onze resultaten lieten een ingewikkelde relatie zien tussen BMI, sociaaleconomische status en CVD. Om de mogelijke oorzaken van CVD op de juiste manier te bestuderen, is het onder andere belangrijk om rekening te houden met het lange aanvangstraject. Uit onderzoek weten we dat CVD zich soms wel tientallen jaren kan ontwikkelen voordat de eerste symptomen zichtbaar worden. Door bijvoorbeeld data van personen te verzamelen beginnend vanaf een jongere leeftijd, en dit vervolgens voor de gehele levensduur vol te houden zou je een accuraat beeld kunnen schetsen van de daadwerkelijke blootstelling aan risico factoren over tijd. Hiermee is het mogelijk om op een individueel niveau het risico op het ontwikkelen van CVD te bepalen. Uiteindelijk zullen interventies die CVD bestrijden zich wellicht moeten focussen op een jongere populatie, om zo in een vroeg stadium in te kunnen grijpen.

LIST OF PUBLICATIONS

In this thesis:

Martens LG, Luo J, Meulmeester FL, Ashrafi N, van Eekelen EW, de Mutsert R, Mook-Kanamori DO, Rosendaal FR, van Dijk KW, Mills K, Noordam R, van Heemst D. *Associations between lifestyle factors and vitamin E metabolites in the general population.* Antioxidants (Basel). 2020 Dec 15;9(12):1280. doi: 10.3390/antiox9121280

Martens LG, Luo J, Willems van Dijk K, Jukema JW, Noordam R, van Heemst D. *Diet-derived antioxidants do not decrease risk of ischemic stroke: a Mendelian Randomization Study in 1 million people.* J Am Heart Assoc. 2021 Dec 7;10(23):e022567. doi: 10.1161/JAHA.121.022567. Epub 2021 Nov 19.

Martens LG, Luo J, Wermer MJH, Willems van Dijk K, Hägg S, Grassmann F, Noordam R, van Heemst D. *The association between mitochondrial DNA abundance and Stroke: a combination of multivariable-adjusted survival and mendelian randomization analyses.* Atherosclerosis. 2022 Aug;354:1-7. doi: 10.1016/j.atherosclerosis.2022.06.1012. Epub 2022 Jun 19.

Not in this thesis:

Luo J, Meulmeester FL, **Martens LG**, Ashrafi N, de Mutsert R, Mook-Kanamori DO, Rosendaal FR, Willems van Dijk K, le Cessie S, Mills K, Noordam R, van Heemst D. *Urinary oxidized, but not enzymatic vitamin E metabolites are inversely associated with measures of glucose homeostasis in middle-aged healthy individuals.* Clin Nutr. 2021 Jun;40(6):4192-4200. doi: 10.1016/j.clnu.2021.01.039. Epub 2021 Feb 3.

Meulmeester FL, Luo J, **Martens LG**, Ashrafi N, de Mutsert R, Mook-Kanamori DO, Lamb HJ, Rosendaal FR, Willems van Dijk K, Mills K, van Heemst D, Noordam R. *Association of measures of body fat with serum alpha-tocopherol and its metabolites in middle-aged individuals.* Nutr Metab Cardiovasc Dis. 2021 Jul 22;31(8):2407-2415. doi: 10.1016/j.numecd.2021.05.001. Epub 2021 May 18.

Luo J, Hashimoto Y, **Martens LG**, Meulmeester FL, Ashrafi N, Mook-Kanamori DO, Rosendaal FR, Jukema JW, van Dijk KW, Mills K, le Cessie S, Noordam R, van Heemst D. *Associations of metabolomic profiles with circulating vitamin E and urinary vitamin E metabolites in middle-aged individuals.* Nutrition. 2022 Jan;93:111440. doi: 10.1016/j.nut.2021.111440. Epub 2021 Jul 29.

Meulmeester FL, Luo J, **Martens LG**, Mills K, van Heemst D, Noordam R. *Antioxidant Supplementation in Oxidative Stress-Related Diseases: What Have We Learned from Studies on Alpha-Tocopherol?* Antioxidants (Basel). 2022 Nov 24;11(12):2322. doi: 10.3390/antiox11122322.

PHD PORTFOLIO

Courses	**Years**	**Hours**
PhD Introductory Meeting	2021	5
Epidemiology “The First Round” (Rothman)	2019	84
Basic Methods and Reasoning in Biostatistics	2020	42
Basic Qualification for Education (BKO)	2019-2022	160
Basic Course “Regulations and Organization for Clinical Researchers”	2021	42

Attended Congresses and meetings	**Years**	**Hours**
Annual Dutch diabetes research meeting (Wageningen, the Netherlands	2019	8
The European and international congress on obesity (online)	2020	6
European Atherosclerosis Society Congress (online)	2021	6
International conference on Mendelian Randomization (online)	2021	6
Annual Dutch Diabetes research meeting (Wageningen, the Netherlands)	2022	8
Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) meeting (Philadelphia, United States of America)	2022	24

Teaching Activities	**Years**	**Hours**
Weekly working groups for Master Vitality and Ageing students for courses Biology of Vitality and Ageing, the Older Individual and Organization of the Ageing Society	2019-2021	1440
Annually returning final working group for Bachelor students in medicine regarding gerontology and geriatrics (Vraagstukken Latere Levensfasen)	2021-2022	16
Annually returning supervision of an assignment on mendelian randomization for students of the Master Vitality and Aging	2019-2021	12
Junior Course Coordinator Biology of Vitality and Ageing	2020-2021	80
Junior Course Coordinator Research and Evidence	2020-2021	80

Student Supervision	**Years**	**Hours**
Twan Simons (Master Thesis)	2021	80
Daan van Hamersveld (Bachelor Thesis)	2021	10

CURRICULUM VITAE

Leon Gilbert Martens is geboren op 15 juli 1994 in Haarlem. Hij behaalde zijn diploma in het tweetalig Atheneum in 2012 aan het Schoter Scholengemeenschap in Haarlem. Daarna heeft hij zijn bachelor Gezondheid en Leven, met als afstudeerrichting Biomedische Wetenschappen, behaald aan de Vrije Universiteit in Amsterdam. Vervolgens is hij de Master Vitality and Ageing gaan volgen aan de Universiteit van Leiden. Aansluitend is hij begonnen als junior docent voor deze opleiding. Na 1 jaar gewerkt te hebben als junior docent, heeft hij deze baan gecombineerd met een baan als onderzoeker op de afdeling Ouderengeneeskunde in het Leids Universitair Medisch Centrum. Op deze afdeling is hij uiteindelijk ook begonnen aan zijn promotieonderzoek, lopend van mei 2020 tot mei 2022. Momenteel werkt hij als Research Consultant bij zorgverzekeraar Zorg en Zekerheid.

DANKWOORD

Om te beginnen wil ik alle deelnemers van de verschillende studies waarop de onderzoeken in dit proefschrift zijn gebaseerd bedanken. Zonder jullie, geen onderzoek.

Verder zou ik graag al mijn collega's, vrienden, en familie willen bedanken voor alle steun en hulp de afgelopen jaren. In het bijzonder bedank ik:

Mijn promotor Prof. dr. Ko Willems van Dijk. Beste Ko, bedankt voor je onmisbare adviezen en treffende metaforen, die je altijd zo kenmerkend over kunt brengen.

Mijn co-promotoren dr. Ir. Diana van Heemst en dr. Raymond Noordam. Beste Diana, jouw enthousiasme over het biologisch onderzoek werkt altijd aanstekelijk. Bedankt voor de boeiende gesprekken. Beste Raymond, het is niet overdreven om te zeggen dat dit me zonder jou niet was gelukt. Bedankt voor je begeleiding en toegankelijkheid. Ik heb ontzettend veel van je geleerd.

Collega en co-auteur dr. Jiao Luo. Beste Jiao, ik ken niemand die harder werkt dan jij. Bedankt voor je expertise en je geduld tijdens onze vele samenwerkingen.

Mijn collega's van de afdeling Ouderengeneeskunde. In het bijzonder mijn mede (oud) promovendi. Evie, Laura, Nicolien, Yara, Veerle, Michelle. Bedankt voor het meedenken, sparren en delen in frustraties die alleen jullie kunnen begrijpen.

Alle oud-collega's van de opleiding Vitality and Ageing. Ik wil jullie bedanken voor alle wijsheid en betrokkenheid de afgelopen jaren. Een betere start van mijn carrière had ik me niet kunnen wensen.

Mijn Paranimf Gali. Wij begrijpen elkaar. Bedankt voor je grenzeloze behulpzaamheid, niet alleen als paranimf, maar ook daarbuiten.

Graag wil ik mijn vrienden bedanken. In het bijzonder mijn oud-huisgenoten Ricardo, Max Roelofsen, Max Ahlers en Ronald. Bedankt voor de broodnodige afleiding en de onvergetelijke jaren.

Lieve pap en mam, bedankt dat ik me voor jullie nooit hoef te bewijzen. Bedankt voor het vinden van de relativering die ik soms verlies.

Mijn paranimf en zus Joan. Een broer-zus band als deze is ontzettend zeldzaam. Daar ben ik me elke dag bewust van. Bedankt dat je er ook tijdens mijn promotie voor me bent. Ik hoop dat wij er op deze manier altijd voor elkaar zullen zijn.

Winnie. Lieve Win, jij bent de enige die altijd dwars door mij heen kijkt en ziet wat ik eigenlijk zeggen wil. Bedankt voor de aanmoediging die ik nodig had om het af te maken. Ik kan niet wachten om te zien wat het leven ons zal brengen.