Oxidative stress in chronic diseases: causal inference from observational studies

Luo, J.

Oxidative Stress in Chronic Diseases:

Causal inference from observational studies

Jiao Luo

Oxidative Stress in Chronic Diseases: Causal Inference from Observational Studies
PhD thesis. Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, the Netherlands

J. Luo was supported by the Chinese Scholarship Council (No. 201808500155).

Cover design and layout: Rentao Li and Jiao Luo
Printed by: printsupport4u.nl
ISBN: 978-94-93289-08-6

Oxidative Stress in Chronic Diseases:

Causal inference from observational 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 donderdag 1 september 2022
klokke 13.45 uur

door

Jiao Luo
geboren te Hongya, China
in 1991

Promotor
Prof. dr. S. le Cessie

Co-promotors
Dr. D. van Heemst
Dr. R. Noordam

Leden promotiecommissie
Prof. dr. B. T. Heijmans
Dr. R. de Mutsert
Prof. dr. R. Frikke-Schmidt (University of Copenhagen, Denmark)
Dr. M. Kavousi (Erasmus University Medical Center, the Netherlands)

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

To my family

献给我的家人

“Knowledge is the object of our inquiry,
and men do not think they know a thing
till they have grasped
the ‘why’ of (which is to grasp its primary cause).“

Aristotle, 384-322 BC

TABLE OF CONTENTS

Part I	**General aspects**
Chapter 1	General introduction, study population, and thesis outline
Chapter 2	Ageing, age-related diseases, and oxidative stress: What to do next?
Part II	**Mitochondrial dysfunction in cardiovascular disease**
Chapter 3	Low mitochondrial DNA copy number drives atherosclerotic cardiovascular disease: cohort and genetic studies
Chapter 4	The associations between mitochondrial DNA copy number and metabolomic profiles in the general population
Part III	**Antioxidants in cardiovascular disease**
Chapter 5	Diet-derived circulating antioxidants and risk of coronary heart disease: a Mendelian randomization study
Chapter 6	Urinary oxidized, but not enzymatic vitamin E metabolites are inversely associated with measures of glucose homeostasis in middle-aged healthy individuals
Chapter 7	Associations of metabolomic profiles with circulating vitamin E and urinary vitamin E metabolites in middle-aged individuals

Part IV **Inflammation in neurological disorders**

Chapter 8 Depression and inflammatory bowel diseases:
a bidirectional two-sample Mendelian randomization study
Chapter 9 Systemic inflammatory markers in relation to
cognitive function and measures of brain atrophy:
a Mendelian randomization study

Part V **Summary**

Chapter 10 Main findings and future perspectives
Appendices Nederlandse samenvatting (Summary in Dutch)
List of publications
PhD Portfolio
Curriculum vitae
Acknowledgements

CHAPTER 1

General introduction, study population, and thesis outline

INTRODUCTION

According to WHO's 2019 Global Health Estimates, chronic diseases, also known as non-communicable diseases, make up seven of the top ten causes of death, accounting for 71% of all deaths globally. Among those, cardiovascular disease (CVD) remains the most common cause of mortality[1] and neurological disorders are the leading cause of disability and the second leading cause of death[2] worldwide. The Global Disease Burden study estimated 17.8 million deaths attributable to CVD in 2017, which represents an increase of 21% in population mortality from CVD and an increase of 15% in years of life lost in the decade before 2017[1]. Similarly, approximately 3 million deaths and 26% increment in years of life lost are attributed to neurological disorders[1]. A set of traditional risk factors have been well acknowledged for chronic diseases, including but not limited to, tobacco use, obesity, hypertension, hyperlipidemia[3]. These have been combined into multiple risk assessment tools to estimate individuals' risk for developing diseases such as CVD[4-6] and dementia[7]. Pharmacological and lifestyle interventions targeting several risk factors can also substantially reduce CVD risk[8-11]. Nevertheless, most of these main causes, if not all, share a common pathological mechanism in the development of chronic diseases, and that is oxidative stress[12]. Therefore, oxidative stress is fundamental among the pathophysiological mechanisms leading to chronic diseases, and a better understanding of their associations may provide early opportunities as a preventive and therapeutic target.

Oxidative stress

Oxidative stress was first formulated in 1985 by Helmut Sies and refined afterwards[13-15]. It refers to conditions whenever the generation of reactive oxygen species (ROS) exceeds endogenous antioxidant capacity, leading to disruption of redox signaling and molecular damage. ROS are the by-products of aerobic metabolism predominantly generated in mitochondria through the electron transport chain during oxidative phosphorylation. The heart and brain are high oxygen-consuming organs to produce energy continuously, with large amounts of mitochondria constituting up to one-quarter of cardiomyocyte volume[16]. This inevitably increases the susceptibility of the heart and vasculature and the neural system to oxidative stress and associated damage. Notably, ROS in low to modest levels regulates multiple redox-dependent vascular wall signaling processes[17], whereas maladaptive excessively high ROS levels mediate irreversible and nonspecific macromolecular damage to cellular membranes, proteins, and DNA[18], and induction of inflammation[19] and dysregulation of vascular functions[20], particularly endothelial dysfunction by disrupting the nitric oxide signaling cascade[20,21]. Collectively, high levels of oxidative stress promote the development of atherosclerosis and further CVD, as well as other chronic diseases.

Mitochondrial dysfunction

ROS generated in mitochondria contributes to mitochondrial dysfunction, which in turn stimulate mitochondrial ROS overproduction, forming a vicious circle. Mitochondrial dysfunction has been widely seen as a hallmark of the ageing

process[22], with disruption of energy transduction, and perturbed calcium and redox homeostasis[23,24], and plays a role in the etiology of multiple age-related diseases. Mitochondria have their own circular genome, the cytoplasmic mitochondrial DNA (mtDNA), consisting of 37 genes, 13 of which encode multi-subunit enzymatic components of the electron transport chain. Individual mitochondria may contain two to ten copies of the mitochondrial genome, known as mtDNA copy number (mtDNA-CN). Therefore, considerable variations exist across cells, tissues, and individuals. The alterations of mtDNA-CN per nucleated cell might be indicative of aberrant mitochondrial health and are associated with perturbations of bioenergetics, mitochondrial membrane potential, and oxidative stress[25], and therefore could roughly serve as an easily accessible and minimally invasive surrogate biomarker of mitochondrial dysfunction[26].

In epidemiological cohort studies, lower peripheral leukocyte mtDNA-CN has been associated with several CVDs and CVD-related risk factors. For example, in the prospective Atherosclerosis Risk in Communities (ARIC) study comprising about 20,000 participants with no history of CVD during a follow-up of up to more than 20 years, a low mtDNA-CN was associated with a 1.29- and 1.11- times higher risk of coronary artery disease and stroke, respectively[27]. A low level of mtDNA-CN was cross-sectionally associated with several cardiovascular risk factors such as obesity, hypertension, hyperglycemia, and diabetes in 408,361 participants[28].

Antioxidants

Antioxidants are substances that neutralize oxidants and their actions to maintain a biological redox steady state. Given the crucial role of ROS in the pathogenesis of chronic diseases, particularly CVD, it has been of particular interest for decades to investigate the protective role of antioxidants in CVD. Of those, dietary-derived modifiable ones have been central for a long time, including β-carotene, vitamin C, and E, among others. However, an important paradox of antioxidants exists.

Since the late 1980s, multiple epidemiological studies have been carried out to investigate the role of antioxidants in CVD. In prospective epidemiological cohort studies, high intake of dietary antioxidants at baseline, either as dietary components or supplements, or high blood concentration of antioxidants including vitamin C and E, have been consistently associated with a lower CVD risk[29-35]. For example, two major studies from 1993 concluded that high consumption of vitamin E was associated with a reduced risk of coronary heart disease[36,37]. These promising findings were widely advocated in the mass news media and led to a significant increase in supplement use in the general public. Not surprisingly, a considerable proportion of the population used dietary antioxidative supplements over the last few decades, particularly in the United States[38]. Notwithstanding, the translation of these findings into evidence-based interventions has not been so straightforward. Randomized clinical trials (RCTs) and their meta-analyses generally failed to demonstrate a significant beneficial effect of antioxidants supplementation, including β-carotene, vitamin C, E, B12, B6, and folic acid, on CVD outcomes and related risk factors[39-45]. Currently, there is inadequate evidence on the benefits of several supplementations in the prevention of CVD[46].

Myths or facts

This vitamin paradox raises the importance of correct causal inference to contribute to our etiological understanding of the role of antioxidants in CVD and in public health. Although the associations between oxidative stress, proxied by mitochondrial dysfunction or low levels of antioxidants in the blood, and CVD have been identified in several observational studies, several limitations challenge the causal inference of the original research findings.

Conventional observational epidemiological studies are often plagued by confounding, reverse causation, and other forms of bias that can limit the validity of the findings from these studies. The confounding factors that causally influence both a risk factor and an outcome induce an spurious association between the risk factor and the outcome. Adjustment for known confounders is possible in statistical analyses, but residual confounding may remain, either due to confounders not being measured or measurement errors in the confounder. Reverse causation is the situation in which the observed association between the risk factor and an outcome is because the outcome causally affects the risk factor. For example, people with early symptoms of CVD may adjust their lifestyle to be healthier accordingly, with higher dietary intake of antioxidants and earlier initiation of supplementation than those without symptoms. Therefore, an interpretation of these spurious associations as causality would be misleading.

RCTs are regarded as the "gold standard" to establish causal effects for all intervention studies because imbalances of participants' characteristics are eliminated by randomization and any differences in outcome can therefore only be attributed to the intervention. However, there are substantial challenges to examine antioxidants supplementation in RCTs, and several mechanisms could explain the null findings in RCTs. First, administration strategies are highly heterogenous, such as timing, monotherapy or not, dose, and treatment duration. For example, supplementation started later than irreversible ROS damage could have a negligible clinical beneficial effect. Moreover, a clear limit in circulating levels is reached by antioxidant supplementation that might be below the therapeutic levels[47]. Also, the circulating antioxidants from supplementation may not access the required target sites to scavenge ROS, especially the mitochondria, and could not represent the functional levels. For instance, vitamin E undergoes different catabolism depending on whether radical-dependent or not and generates distinct metabolites that can be detected in the urine[48]. Furthermore, the form of the antioxidants, natural or synthetic, may exert an influence, such as lower bioactivity of synthetic vitamin E than the natural ones, characterized by preferentially non-oxidation metabolites in urine[49].

Therefore, there is an urgent demand to explore innovative and rigorous research designs for examining the effects of antioxidants taking those complexities in antioxidants and administration into account.

Mendelian randomization in causal inference

Mendelian randomization (MR) has emerged in recent years as a powerful study design in causal inference. It uses genetic variants associated with the exposure of interest to investigate the causal effect of modifiable risk factors on outcomes.

Genetic variants inherited from parents to offspring at conception are randomly assorted and segregated at meiosis and are used as proxies of the exposure levels. As a result, individuals are divided into two comparable groups. Those who carry the effect allele (e.g. with an increased level of exposure) are assigned to the group with higher levels of the exposure of interest whereas those who carry the alternative allele are assigned to the group with lower levels of exposure. Through this process of random allocation, genetic variants are independent of environmental and other genetic factors, except for those variants that are in linkage disequilibrium of the variant of interest. Hence, genetic variants will not be related to any potential confounding factors that affect the exposure and the outcome and any difference in outcomes between genetically defined groups can be directly attributed to the exposure. Consequently, the MR design is considered as a "natural experiment", mimicking RCTs, where individuals are randomized to carry a genetic variant rather than to an intervention. In addition, the earlier disposition of genetic variants than the exposure measurements ensures the correct temporal order to eliminate the possibility of reverse causation. Therefore, MR can be used to estimate the causal relationship between exposures and outcomes largely free from confounding, in contrast to conventional epidemiological approaches.

The massively increasing availability of genotyped and high-quality phenotypic data in large study samples and mega biobanks during the past decade facilitates us to apply MR designs and to generate valid and well-powered approximated causal associations between exposures and outcomes. Indeed, the past decade has witnessed a considerable increase of publications using the MR approach50. To obtain reliable causal estimates of the effect of an exposure on an outcome from MR studies, the following three principal assumptions should be fulfilled by the genetic variants, namely that they 1) are associated with the exposure in genetic studies (relevance); 2) are not related to any observed or unobserved confounding factors (independence); 3) are associated with the outcome exclusively through its effect on exposure (exclusion restriction), as illustrated in **Figure 1**.

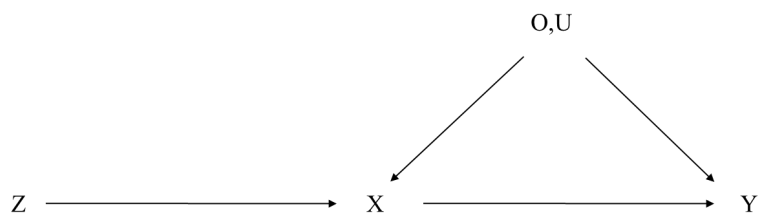


Figure 1 A graphical presentation for causal inference using Mendelian randomization.

Z represents the genetic instruments. To conclude that exposure of interest (X) is a causal risk factor for the outcome (Y), three assumptions need to be satisfied. First, relevance is implied by the arrow from Z to X (Z→X), where Z is causally associated with the exposure. Second, independence is that Z is independent of any observed (O) or unobserved (U) confounding factors. Third, exclusion

restrictions signify that Z is independent of Y conditioning on exposure and confounding factors, namely no path from Z to Y other than via X (Z→X→Y), i.e., no horizontal pleiotropy.

OUTLINE OF THIS THESIS

This thesis aims to provide extensive insights into the role of oxidative stress in the onset of chronic diseases with an emphasis on CVD and related cardiometabolic risk factors via applying conventional epidemiological approaches in combination with Mendelian randomization (MR) designs. The thesis is structured in five. In **Part I** we provide an overview of oxidative stress and its role in ageing and age-related diseases on the progress beyond state-of-the-art. In **part II** we investigate the associations between mitochondrial dysfunction and CVD and lipid profiles. In **Part III** we focus on the role of antioxidants in CVD and cardiometabolic traits. In **Part IV** we expand the scope of this thesis by examining the association of inflammation, which is inextricably interrelated to oxidative stress, with neurological diseases. **Part V** contains a summary of the main findings from this thesis and a general discussion.

Part I: General aspects

Oxidative stress has been put forward for decades, the role of oxidative stress in ageing and age-related diseases is however controversial. In **Chapter 2**, we provide a detailed overview of the generation of reactive oxygen species and their role in redox signaling and oxidative damage. In addition, we review the current evidence regarding the association of oxidative stress and ageing and age-related diseases as CVD and neurodegenerative disease. We then summarize the possible reasons for the inconsistency and put forward some remarks.

Part II: Mitochondrial dysfunction in cardiovascular disease

"Triangulation" in etiological epidemiology could be helpful to strengthen causal inference by integrating several different approaches that are assumed to have different and largely unrelated sources of potential bias[51]. When estimates obtained from different approaches converge on a similar effect, evidence favors that the relationship between exposure and outcome is causal. Hence, we combine a prospective cohort study design using data from the general population from the UK biobank and a Mendelian randomization framework exploiting publicly available summary-level data from the large consortia and mega-biobanks to infer causality different exposures and outcomes. We specifically look into the association between mitochondrial DNA copy number and incident coronary artery disease and heart failure in **Chapter 3** and metabolomic profiles (mostly lipids and lipoprotein fractions) in **Chapter 4**.

Part III: Antioxidants in cardiovascular disease

We still do not know the causal nature of the associations between antioxidants and CVD because of the contradictory results from observational studies and RCTs and considering the limitations of both designs. Here, we focus on dietary-derived antioxidants, predominantly on the most appealing yet debatable well-known chain-breaking antioxidant: vitamin E. In **Chapter 5**, using a Mendelian randomization design, we investigate the associations between genetically predicted circulating antioxidants levels (vitamins E and C, retinol,

β-carotene, and lycopene), both as absolute levels and their metabolites, and the risk of coronary heart disease, using publicly available genetic consortium. Since vitamin E is partly catabolized in a radical-dependent way, even the circulating levels are not representative of the functional levels, which might be the most plausible explanation for the null findings between circulating vitamin E and CVD. Accordingly, there are two forms of vitamin E metabolites in the urine, the oxidized metabolites being indicative of its antioxidative capacity, and enzymatic metabolites. We subsequently link these two metabolites to cardiovascular risk factors and cardiometabolic traits, including glucose metabolism in **Chapter 6** and lipoprotein profiles in **Chapter 7** in the Netherlands Epidemiology of Obesity study (NEO) using an observational study design.

Part IV: Inflammation in neuropsychiatric diseases

Inflammation and oxidative stress are inextricably linked. Inflammation, the "host defense" against pathogens and chemical and physical challenges towards tissue integrity to restore tissue homeostasis through inducing various repair mechanisms, involves an enhanced release and accumulation of ROS by activated immune cells. Evidence has emerged that ROS play a critical role in the pathophysiology of inflammation. Inflammation has been well-acknowledged in atherosclerotic CVD but its association with neurological diseases is less known. For this reason, we expand the scope of this thesis by examining also neurological diseases as additional outcomes with large relevance to population health. In **Chapter 8**, we investigate the associations between inflammatory bowel disease that is characterized by chronic inflammation and depression using a bidirectional Mendelian randomization approach. In **Chapter 9**, we examine the role of multiple systemic inflammatory markers in cognitive function and brain atrophy measures.

Part V: Summary

Finally, **Chapter 10** summarizes the main findings of this thesis and discusses some future perspectives in the field.

MAIN STUDY POPULATIONS

The Netherlands Epidemiology of Obesity study (NEO)

The NEO study is used in **Chapters 6 and 7**. This population-based prospective cohort study started in 2008 and includes 6,671 individuals aged 45-65 years, with an oversampling of individuals with a self-reported body mass index (BMI) of 27 kg/m^2 or higher. All inhabitants aged between 45 to 65 years from the municipality of Leiderdorp were invited irrespective of their BMI. The study was approved by the medical ethical committee of the Leiden University Medical Center (LUMC), and all participants gave written informed consent. Participants were invited to come to the NEO study center of LUMC for one baseline study visit after an overnight fast. Prior to this study visit, participants collected their urine over 24 hours and completed a general questionnaire at home in terms of their demographic, lifestyle, and clinical data in addition to specific questionnaires on diet and physical activity. Patients were asked to bring their medication use within one month prior to the visit and relevant information was recorded by research nurses. Fasting blood samples were drawn for biochemical measurements.

UK biobank

Individual and summary-level data of the UK biobank are used in **Chapters 3, 4, 5, and 8**. The UK Biobank is a prospective cohort with 502,628 participants between the age of 40 and 69 years recruited from the general population at multiple assessment centers across the UK between 2006 and 2010. 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 genotyping. The 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) from England and Wales. All participants provided electronically written informed consent for the study.

References

1. GBD 2017 Causes of Death Collaborators. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018; 392(10159): 1736-88.
2. GBD 2015 Neurological Disorders Collaborator Group. Global, regional, and national burden of neurological disorders during 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet Neurol 2017; 16(11): 877-97.
3. S. Yusuf, P. Joseph, S. Rangarajan, et al. Modifiable risk factors, cardiovascular disease, and mortality in 155 722 individuals from 21 high-income, middle-income, and low-income countries (PURE): a prospective cohort study. Lancet 2020; 395(10226): 795-808.
4. SCORE2 working group and ESC Cardiovascular risk collaboration. SCORE2 risk prediction algorithms: new models to estimate 10-year risk of cardiovascular disease in Europe. Eur Heart J 2021; 42(25): 2439-54.
5. D. C. Goff, Jr., D. M. Lloyd-Jones, G. Bennett, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. Circulation 2014; 129(25 Suppl 2): S49-73.
6. P. W. Wilson, R. B. D'Agostino, D. Levy, A. M. Belanger, H. Silbershatz, W. B. Kannel. Prediction of coronary heart disease using risk factor categories. Circulation 1998; 97(18): 1837-47.
7. I. J. Rasmussen, K. L. Rasmussen, B. G. Nordestgaard, A. Tybjærg-Hansen, R. Frikke-Schmidt. Impact of cardiovascular risk factors and genetics on 10-year absolute risk of dementia: risk charts for targeted prevention. Eur Heart J 2020; 41(41): 4024-33.
8. F. Taylor, M. D. Huffman, A. F. Macedo, et al. Statins for the primary prevention of cardiovascular disease. Cochrane Database Syst Rev 2013; 2013(1): Cd004816.
9. J. Fulcher, R. O'Connell, M. Voysey, et al. Efficacy and safety of LDL-lowering therapy among men and women: meta-analysis of individual data from 174,000 participants in 27 randomised trials. Lancet 2015; 385(9976): 1397-405.
10. The Blood Pressure Lowering Treatment Trialists' Collaboration. Blood pressure-lowering treatment based on cardiovascular risk: a meta-analysis of individual patient data. Lancet 2014; 384(9943): 591-8.
11. D. Ettehad, C. A. Emdin, A. Kiran, et al. Blood pressure lowering for prevention of cardiovascular disease and death: a systematic review and meta-analysis. Lancet 2016; 387(10022): 957-67.
12. B. Niemann, S. Rohrbach, M. R. Miller, D. E. Newby, V. Fuster, J. C. Kovacic. Oxidative Stress and Cardiovascular Risk: Obesity, Diabetes, Smoking, and Pollution: Part 3 of a 3-Part Series. J Am Coll Cardiol 2017; 70(2): 230-51.
13. H. Sies. Oxidative stress: a concept in redox biology and medicine. Redox Biol 2015; 4: 180-3.
14. H. Sies. 1 - Oxidative Stress: Introductory Remarks. In: Sies H, ed. Oxidative Stress. London: Academic Press; 1985: 1-8.
15. H. Sies, C. Berndt, D. P. Jones. Oxidative Stress. Annu Rev Biochem 2017; 86: 715-48.
16. J. Schaper, E. Meiser, G. Stämmler. Ultrastructural morphometric analysis of myocardium from dogs, rats, hamsters, mice, and from human hearts. Circ Res 1985; 56(3): 377-91.

17. J. K. Freed, D. D. Gutterman. Mitochondrial reactive oxygen species and vascular function: less is more. Arterioscler Thromb Vasc Biol 2013; 33(4): 673-5.
18. R. A. Jacob, B. J. Burri. Oxidative damage and defense. Am J Clin Nutr 1996; 63(6): 985s-90s.
19. H. Li, S. Horke, U. Förstermann. Vascular oxidative stress, nitric oxide and atherosclerosis. Atherosclerosis 2014; 237(1): 208-19.
20. N. R. Madamanchi, A. Vendrov, M. S. Runge. Oxidative stress and vascular disease. Arterioscler Thromb Vasc Biol 2005; 25(1): 29-38.
21. U. Förstermann. Nitric oxide and oxidative stress in vascular disease. Pflugers Arch 2010; 459(6): 923-39.
22. C. López-Otín, M. A. Blasco, L. Partridge, M. Serrano, G. Kroemer. The hallmarks of aging. Cell 2013; 153(6): 1194-217.
23. D. E. Handy, J. Loscalzo. Redox regulation of mitochondrial function. Antioxid Redox Signal 2012; 16(11): 1323-67.
24. L. Contreras, I. Drago, E. Zampese, T. Pozzan. Mitochondria: the calcium connection. Biochim Biophys Acta 2010; 1797(6-7): 607-18.
25. D. C. Wallace. A mitochondrial bioenergetic etiology of disease. J Clin Invest 2013; 123(4): 1405-12.
26. A. N. Malik, A. Czajka. Is mitochondrial DNA content a potential biomarker of mitochondrial dysfunction? Mitochondrion 2013; 13(5): 481-92.
27. F. N. Ashar, Y. Zhang, R. J. Longchamps, et al. Association of Mitochondrial DNA Copy Number With Cardiovascular Disease. JAMA Cardiol 2017; 2(11): 1247-55.
28. X. Liu, R. J. Longchamps, K. L. Wiggins, et al. Association of mitochondrial DNA copy number with cardiometabolic diseases. Cell Genom 2021; 1(1).
29. D. Aune, N. Keum, E. Giovannucci, et al. Dietary intake and blood concentrations of antioxidants and the risk of cardiovascular disease, total cancer, and all-cause mortality: a systematic review and dose-response meta-analysis of prospective studies. Am J Clin Nutr 2018; 108(5): 1069-91.
30. P. Knekt, J. Ritz, M. A. Pereira, et al. Antioxidant vitamins and coronary heart disease risk: a pooled analysis of 9 cohorts. Am J Clin Nutr 2004; 80(6): 1508-20.
31. G. Li, Y. Li, X. Chen, H. Sun, X. Hou, J. Shi. Circulating tocopherols and risk of coronary artery disease: A systematic review and meta-analysis. Eur J Prev Cardiol 2016; 23(7): 748-57.
32. B. Song, K. Liu, Y. Gao, et al. Lycopene and risk of cardiovascular diseases: A meta-analysis of observational studies. Mol Nutr Food Res 2017; 61(9).
33. J. Kim, J. Choi, S. Y. Kwon, et al. Association of Multivitamin and Mineral Supplementation and Risk of Cardiovascular Disease: A Systematic Review and Meta-Analysis. Circ Cardiovasc Qual Outcomes 2018; 11(7): e004224.
34. Z. Ye, H. Song. Antioxidant vitamins intake and the risk of coronary heart disease: meta-analysis of cohort studies. Eur J Cardiovasc Prev Rehabil 2008; 15(1): 26-34.
35. H. M. Cheng, G. Koutsidis, J. K. Lodge, A. W. Ashor, M. Siervo, J. Lara. Lycopene and tomato and risk of cardiovascular diseases: A systematic review and meta-analysis of epidemiological evidence. Crit Rev Food Sci Nutr 2019; 59(1): 141-58.

36. E. B. Rimm, M. J. Stampfer, A. Ascherio, E. Giovannucci, G. A. Colditz, W. C. Willett. Vitamin E consumption and the risk of coronary heart disease in men. N Engl J Med 1993; 328(20): 1450-6.
37. M. J. Stampfer, C. H. Hennekens, J. E. Manson, G. A. Colditz, B. Rosner, W. C. Willett. Vitamin E consumption and the risk of coronary disease in women. N Engl J Med 1993; 328(20): 1444-9.
38. R. L. Bailey, J. J. Gahche, P. E. Miller, et al. Why US Adults Use Dietary Supplements. JAMA Intern. Med 2013; 173(5): 355-61.
39. L. Al□Khudairy, N. Flowers, R. Wheelhouse, et al. Vitamin C supplementation for the primary prevention of cardiovascular disease. Cochrane Database Syst Rev 2017; (3).
40. L. Schwingshackl, H. Boeing, M. Stelmach-Mardas, et al. Dietary Supplements and Risk of Cause-Specific Death, Cardiovascular Disease, and Cancer: A Systematic Review and Meta-Analysis of Primary Prevention Trials. Adv Nutr 2017; 8(1): 27-39.
41. S. K. Myung, W. Ju, B. Cho, 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: f10.
42. S. P. Fortmann, B. U. Burda, C. A. Senger, J. S. Lin, E. P. Whitlock. Vitamin and mineral supplements in the primary prevention of cardiovascular disease and cancer: An updated systematic evidence review for the U.S. Preventive Services Task Force. Ann Intern Med 2013; 159(12): 824-34.
43. D. J. A. Jenkins, J. D. Spence, E. L. Giovannucci, et al. Supplemental Vitamins and Minerals for CVD Prevention and Treatment. J Am Coll Cardiol 2018; 71(22): 2570-84.
44. V. A. Moyer, on behalf of the U.S. Preventive Services Task Force. Vitamin, Mineral, and Multivitamin Supplements for the Primary Prevention of Cardiovascular Disease and Cancer: U.S. Preventive Services Task Force Recommendation Statement. Ann Intern Med 2014; 160(8): 558-64.
45. A. W. Ashor, R. Brown, P. D. Keenan, et al. Limited evidence for a beneficial effect of vitamin C supplementation on biomarkers of cardiovascular diseases: an umbrella review of systematic reviews and meta-analyses. Nutr Res 2019; 61: 1-12.
46. V. A. Moyer. Vitamin, mineral, and multivitamin supplements for the primary prevention of cardiovascular disease and cancer: U.S. Preventive services Task Force recommendation statement. Ann Intern Med 2014; 160(8): 558-64.
47. A. Munteanu, J. M. Zingg, A. Azzi. Anti-atherosclerotic effects of vitamin E--myth or reality? J Cell Mol Med 2004; 8(1): 59-76.
48. P. Torquato, O. Ripa, D. Giusepponi, et al. Analytical strategies to assess the functional metabolome of vitamin E. J Pharm Biomed Anal 2016; 124: 399-412.
49. R. S. Parker, T. J. Sontag, J. E. Swanson, C. C. McCormick. Discovery, characterization, and significance of the cytochrome P450 omega-hydroxylase pathway of vitamin E catabolism. Ann N Y Acad Sci 2004; 1031: 13-21.
50. R. C. Richmond, G. Davey Smith. Mendelian Randomization: Concepts and Scope. Cold Spring Harb Perspect Med 2022; 12(1).
51. D. A. Lawlor, K. Tilling, G. Davey Smith. Triangulation in aetiological epidemiology. Int J Epidemiol 2016; 45(6): 1866-86.

CHAPTER 2

Ageing, age-related diseases, and oxidative stress: what to do next?

Jiao Luo, Kevin Mills, Saskia le Cessie, Raymond Noordam, Diana van Heemst

Ageing Res Rev. 2020 Jan; 57:100982.

Abstract

Among other mechanisms, oxidative stress has been postulated to play an important role in the rate of ageing. Oxidative damage contributes to the hallmarks of ageing and essential components in pathological pathways which are thought to drive multiple age-related diseases. Nonetheless, results from studies testing the hypothesis of oxidative stress in ageing and diseases showed controversial results. While observational studies mainly found detrimental effects of high oxidative stress levels on disease status, randomized clinical trials examining the effect of antioxidant supplementation on disease status generally showed null effects. However, re-evaluations of these counterinitiative observations are required considering the lack of reliability and specificity of traditionally used biomarkers for measuring oxidative stress. To facilitate these re-evaluations, this review summarizes the basic knowledge of oxidative stress and the present findings regarding oxidative damage and ageing and age-related diseases. Meanwhile, two approaches are highlighted, namely proper participants selection, together with the development of reliable biomarkers. We propose that oxidized vitamin E metabolites may be used to accurately monitor individual functional antioxidant levels, which might serve as promising key solutions for future elucidating the impact of oxidative stress on ageing and age-related diseases.

Introduction

It has been widely acknowledged that life expectancy has increased over the past centuries as a specific result of improved medical care, vaccination, and hygiene [1,2]. The process of ageing is a dynamic, chronological process characterized by the gradual accumulation of damage to cells, progressive functional decline, and increased susceptibility and vulnerability to diseases. In addition, ageing is closely connected to the onset and progression of multiple age-related diseases, such as cancer, type 2 diabetes mellitus, and cardiovascular and neurodegenerative diseases [3-5]. The ageing process is postulated to originate from several basic molecular changes, better known as the hallmarks of ageing, which include four primary hallmarks, genomic instability, telomere attrition, epigenetic alterations, and loss of proteostasis, three antagonistic hallmarks, deregulated nutrient sensing, mitochondrial dysfunction, and cellular senescence, and two integrative hallmarks stem cell exhaustion, and altered intercellular communication [6].

These hallmarks contributing to the ageing process could be caused by oxidative damage. For example, telomeres are highly sensitive to oxidative damage and their repair capacity is less well than other parts of the chromosome [7,8]. Hence, oxidative damage may result in telomere attrition that accelerates ageing and increases the risk of age-related diseases [9]. The concept of oxidative stress was introduced in 1985 and updated later [10-12]. Oxidative stress refers to "an imbalance between the generation of oxidants and their elimination systems, i.e. antioxidants, in favor of oxidants, leading to disruption of redox signaling and control and/or molecular damage" [12]. Conceptually, the level of oxidative stress ranges from physiological levels for redox signaling to toxic levels of molecular or organelle damage (**Figure 1**). Redox signaling is essential for host defense as well as in a diverse array of signaling pathways [13,14]. Other damages caused by non-physiological high oxidative stress lead to a wide range of phenotypic changes, including altered gene expression, arrested cell proliferation, cell growth, and cellular senescence [15-17].

Antioxidants may act as scavengers of oxidants to maintain the biological redox steady states. Therefore, since the oxidative stress theory was proposed [18], antioxidants were postulated to potentially play a protective role in ageing and age-related diseases. Considering the premise that adverse health consequences caused by oxidative stress can be counteracted by antioxidants, a comprehensive body of studies aiming to examine the beneficial effects of antioxidants on diseases have been carried out in the past three decades. However, results were often disappointing and counterintuitive. The most appealing and well-known example is vitamin E, a well elucidated chain-breaking antioxidant. Although lower disease risks in individuals with higher vitamin E concentration have been found in many observational studies [19-24], as well as protective properties of vitamin E in animal experiments [25,26], most clinical trials examining vitamin E supplementation failed to demonstrate any advantageous effects on the prevention or treatment of various age-related diseases [27-32].

Along with these conflicting evidence, it seems like the controversy about the oxidative stress theory in ageing and age-related diseases has never stopped (**Figure 2**). In addition, over the past 30 years, fluctuations in the use of antiox-

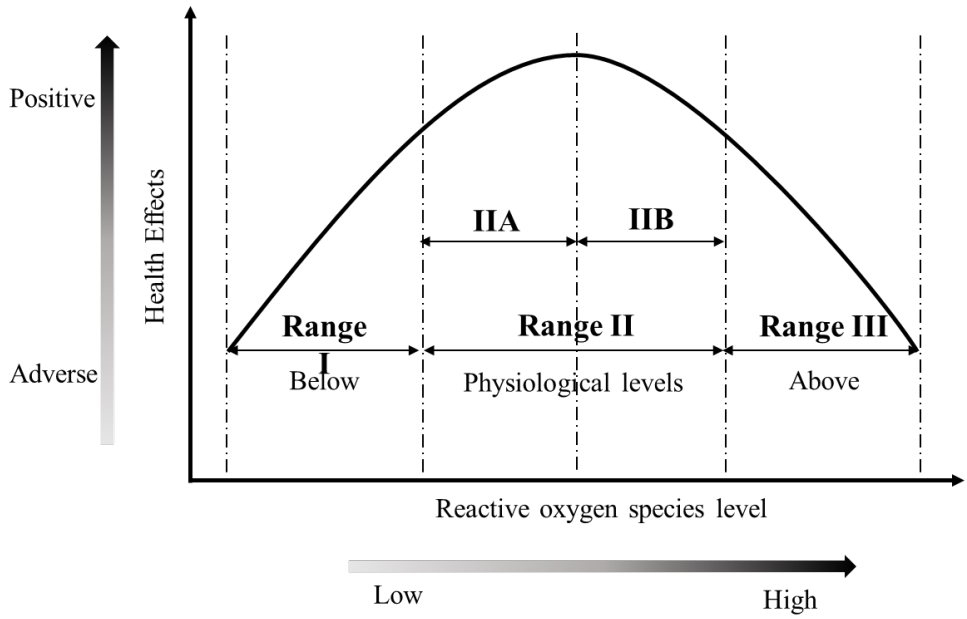


Figure 1 ROS levels and health effects: a dose-response model
Detailed discussion in the texts in part 2.2.

idant supplements were also observed, for example in the US (**Figure 3**) [33,34]. The percentage of individuals using antioxidant supplements gradually increased from the 1980s and peaked in the 1990s. Of note, specifically, the use of vitamin E supplements steeply dropped in the early 21st century, where after decline turned to be stabilized. However, these observations neither imply that any consensus about the effect of antioxidants on diseases has been reached, nor that the oxidative stress theory has been refuted. Conversely, the annual publication count of antioxidant articles steadily increased since the 1990s, and more than 30,000 papers have been published in 2018 alone about this research topic [35].

So far, oxidative damages are thought to play a pivotal role in the pathological processes implicated in ageing and age-related diseases and the underlying biochemical mechanisms have been clarified in detail [6,36]. However, there are still several questions unsettled such as the existing paradox regarding the preventive and therapeutic role of antioxidants (such as vitamin E), the lack of stable and representative biomarkers of oxidative stress, and whether oxidative stress is causally associated with ageing and age-related disease in the general population setting. Therefore, this review is organized as such to provide an overview of the chemical processes involved in oxidative stress and an update on the available evidence about associations with ageing and age-related diseases. In the last part of the review, antioxidants, especially the controversial role of vitamin E will be addressed, together with novel insights and directions for future research.

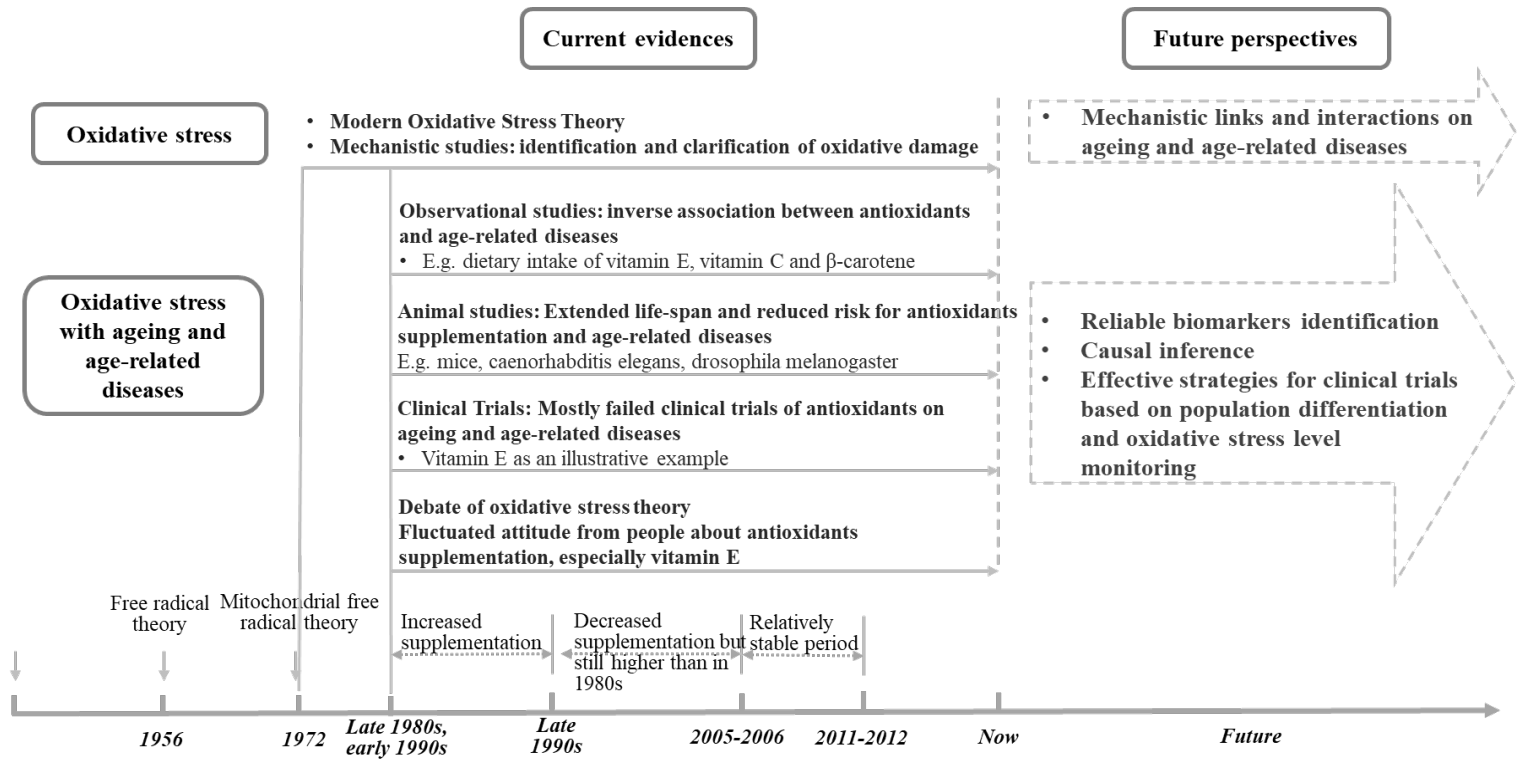


Figure 2 Schematic review of oxidative stress theory in ageing and age-related diseases

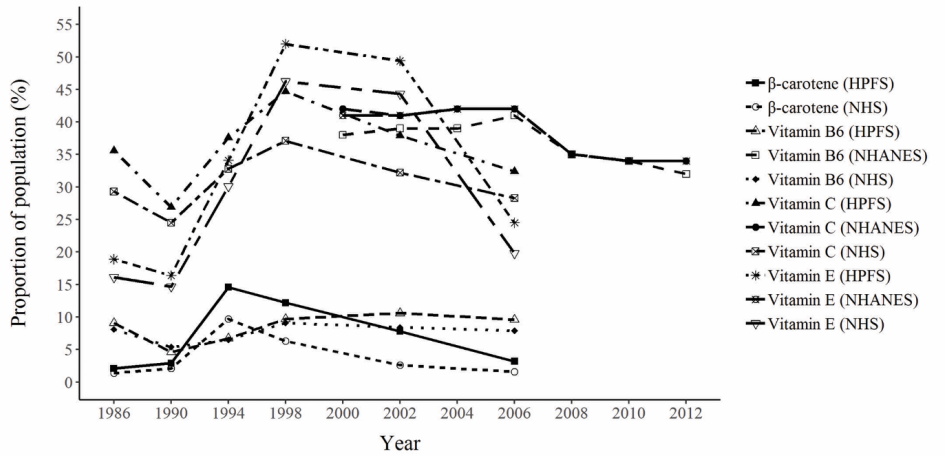


Figure 3 Trends in Antioxidants supplementation in the US

NHS: Nurses' Health Study, HPFS: Health Professionals Follow-up Study, NHANES: National Health and Nutrition Examination Survey. Data presented in the graph of NHS and HPFS are from 1986 to 2006, and data of NHANES are from 1999 to 2012.

The generation and health roles of Reactive Oxygen Species (ROS)

Endogenous generation of ROS

According to the "free radical theory" that was proposed in the 1950s and revised in the 1970s, damages induced by free radicals are the main cause of ageing and a shorter lifespan [18,37,38]. Reactive oxygen species (ROS) are highly reactive molecules, primarily including typical free radicals that contain at least one unpaired electron (superoxide O_2^-•, hydroxyl radical •OH), and hydrogen peroxides (H_2O_2), and have been considered the main source of endogenous oxidative stress damage [39].

It is widely accepted that the bulk of ROS is generated by the mitochondrial electron transport chain during normal oxidative respiration in addition to numerous intracellular pathways (**Figure 4**). It is estimated that about 1-2% of the daily overall oxygen molecules consumed are reduced into O_2^-• with the leak of electrons [40]. This process occurs mainly in two discrete complexes of mitochondrial electron transport chain in the matrix side of inner mitochondrial membrane, notably complex I (NADH-ubiquinone oxidoreductase) and complex III (ubiquinone-cytochrome c reductase) [41]. Iron-sulphur centers are thought as the most likely site of ROS production in complex I [42]. Complex III, also known as the Q cycle, which refers to a set of ubiquinone oxidation reactions, is responsible for the robust production of superoxide, the precursor of other ROS, by non-enzymatic transfer of electrons to molecular oxygen. Once generated, superoxide could be catalyzed by superoxide dismutase (SOD) into H_2O_2, which is unstable and membrane-diffusible peroxide. Subsequently, in the presence of transition cation with reduced form (Fe^{2+} or Cu^+, referred to as the Fenton reaction)

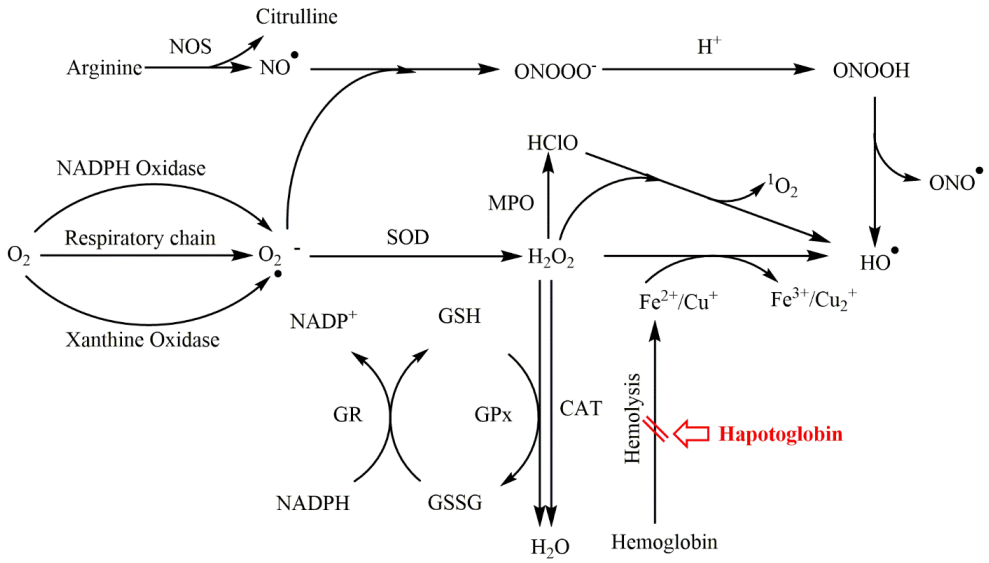


Figure 4 Reactive oxygen species generation

O_2-▪:Super oxide anion, SOD: super oxidize dismutase, H_2O_2: hydrogen peroxide, ▪OH hydroxyl radical, MPO: myeloid peroxide, HClO: hypochlorous acid, $^1O_2$: singlet oxygen, NOS: nitric oxide synthase, NO▪: nitric oxide radical, ONOO-: peroxynitrite, ONOOH: peroxynitrous acid, ONO▪: nitrogen dioxide, GSH: glutathione, GSSG: glutathione disulfide, GPx: glutathione peroxidase, CAT: catalase, NADP+: Nicotinamide adenine dinucleotide phosphate, NADPH: Nicotinamide adenine dinucleotide phosphate reduced form, GR: glutathione reductase.

The bulk of ROS are mainly generated by the mitochondrial electron transport chain during normal aerobic metabolism in addition to multiple ways encompassing cytosolic enzyme systems (NADPH oxidase), monoamine oxidase on the outer membrane of mitochondrial, xanthine oxidase, and uncoupled NOS. In the electron transport chain, oxygen molecules are univalent reduced into O_2-▪ with the leak of electrons. The formation of superoxide is the initial step and the start of a cascade reaction of other ROS generation. Once generated, it could be catalyzed spontaneously by SOD into H_2O_2. Subsequently in the presence of transition cation with reduced form (Fe^{2+} or Cu^+, referred to as the Fenton reaction) or myeloid peroxide (MPO), H_2O_2 further dismutates to •OH. Meanwhile, H_2O_2 can also be reduced into the water by the enzymatic antioxidants such as CAT and GPx. Haptoglobin binds hemoglobin with high affinity and stability, preventing the release of heme iron from hemolysis into circulation, consequently terminating Fenton reaction and preventing the production of •OH.

or myeloid peroxide (MPO) [43], H_2O_2 further dismutates to •OH, the extremely unstable and most reactive ion among all ROS. In summary, the main process of ROS generation in mitochondria could be schematically presented as O_2 → O_2^-• → H_2O_2 → •OH.

Hydroxyl radicals may lead to detrimental damages to macromolecules owing to its chemical properties [44]. Moreover, its formation relies on the presence of a reduced form transition cation, mainly iron generated from the hemoglobin heme group during hemolysis. Hence any component which can stabilize heme iron within hemoglobin could prevent oxidative damage caused by hydroxyl rad-

icals. In recent years, haptoglobin (Hp), an abundant, acute-phase inflammatory glycoprotein, which is predominantly synthesized in the liver and is regulated by cytokines, has been indicated to have an important role in the prevention of the generation of hydroxyl radicals by virtue of its ability to bind hemoglobin with high affinity and stability, thus preventing the release of heme iron from hemolysis into the circulation [45], as shown in **Figure 4**. The Hp gene basically contains two common alleles, namely Hp1 and Hp2, with homozygous (1-1 or 2-2) and heterozygous (2-1) genotypes. In parallel with the theoretical evidence, both haptoglobin concentration and genotype, specifically Hp2-2, are associated with various age-related diseases, such as cancer, cardiovascular disease, etc. [46]. However, the underlying mechanisms related to the pathophysiology of these diseases still remain to be demonstrated.

The complex role of ROS in health maintenance and diseases

The role of ROS in the body is rather complex, and the influences on health vary largely along with changing ROS levels. ROS levels, as a reflection of oxidative stress, are modulated by oxidant generation and their elimination, and are linked to many pathophysiological processes. Within physiological levels, ROS are in a biological redox steady state [12] and facilitate the maintenance of cellular homeostasis and function. However, ROS levels would go toward either side beyond dynamic balance (pathological states). Thus ROS levels have both beneficial and damaging aspects, as put forward in the concept of mitohormesis [47], Consequently, both (too) low and (too) high levels of ROS will have adverse health effects, as illustrated in **Figure 1**.

Physiological levels: beneficial health effects

ROS may act as second messengers owing to the characteristics of having an intricate system for synthesis and removal as well as reversible signaling effects. Both superoxide and hydrogen peroxide could be potential messengers to regulate reduction-oxidation-dependent signaling mechanisms, while hydrogen peroxide has a higher advantage in signaling capacity given its stability and membrane permeability. The major mechanism underlying most redox-dependent signaling has been considered as the reversible modulation of cysteine thiol groups (thiolate anion to sulfenic form Cys-SOH) regulated by hydrogen peroxide [14,48].

Within physiological levels (Range II in **Figure 1**), ROS can promote host defense mechanisms such as for the optimal activity for macrophages against bacteria, as well as in signaling pathways, such as toll-like receptors initiated pathways, Mitogen-Activated Protein Kinase (MAPK) signaling pathways, NF- B signaling pathway and Keap1-Nrf2-ARE signaling pathway [13,14,49,50]. Therefore, ROS levels within the physiological range are critical signaling molecules for many redox-dependent signaling processes including gene expression, metabolic regulation, inflammatory response, stem cell proliferation and differentiation, cancer pathogenesis as well as ageing. Intensive discussion regarding ROS signaling physiological consequences has been provided in previous reviews [13,14].

Given the signaling effect of ROS, we can speculate that a certain increase of ROS in the physiological range would lead to better health effects, from health maintenance to health promotion, such as decreased risks of diseases, or even prolonged lifespan [47], as illustrated in IIA, **Figure 1**. However, with further increase of ROS levels and more damage events, this beneficial effect may decrease or disappear, but with no manifestation of pathological symptoms (IIB, **Figure 1**).

Elevated or Decreased Physiological levels (Pathological state): adverse health effects

When ROS levels are beyond the range of physiological levels, either (too) low or (too) high, adverse health effects can happen. For example, upon inflammatory stimulation, neutrophils are activated and generate large amounts of ROS for host defense. However, Ncf1 (neutrophil cytosolic factor 1) mutated mice, with low production of ROS, have higher type I interferon and develop an accelerated lupus-like disease [51]. Similarly, the Ncf1-339 T allele, related to reduced extracellular ROS production in neutrophils and increased type 1 interferon-regulated genes expression, was found enriched in systemic lupus erythematosus patients compared to healthy controls [52]. Moreover, patients with NADPH oxidase 2 (Nox2) mutations that are associated with lower ROS levels, were more susceptible to develop chronic granulomatous disease [53]. These findings indicated an association between lower ROS generation that is not sufficient to maintain physiological processes which could lead to inferior health outcomes, as shown in Range I, **Figure 1**.

On the other hand, in the situation of high ROS levels (toxicity), excessive ROS might irreversibly react with intracellular macromolecules, including lipids, proteins, and DNA, both in the mitochondria and nucleus. Hazardous products are formed and accumulated, and the functions of macromolecules and organelles are altered. Once these exceed the repair capacity of the body, devastating damage will occur, and accelerated ageing or multiple diseases might manifest or progress (discussion in part 2.3 and part 3), as shown in Range III, **Figure 1**.

What is the physiological level?

Though **Figure 1** shows the possible ranges of ROS levels, several caveats should be noted. Firstly, how to define "physiological levels". ROS generation happens all the time and signaling merits and damaging events occur simultaneously. For example, protein oxidation will produce many harmful byproducts, meanwhile, it is also involved in redox signaling and control, especially cysteine side chain peroxidation. When we are talking about the levels of either ROS or oxidative stress, we should be more cautious of using "high" and "low", given that "normal" is undefined. Possibly, the so-called "high" or "low" ROS levels close to the physiological range should also be regarded as tolerable levels. For example, transgenic animal models with lower SOD production did not show a shortened lifespan in mice [54,55], and even an extended lifespan in C.elegans [56]. This "lower SOD production" may lead to a relatively "higher ROS level", but notably, this level could be tolerated by the experimented animals and they also function well for maintaining or promoting health, thus being potential physiological. In addition, antioxidants would have a similar dose-response manner since the change of

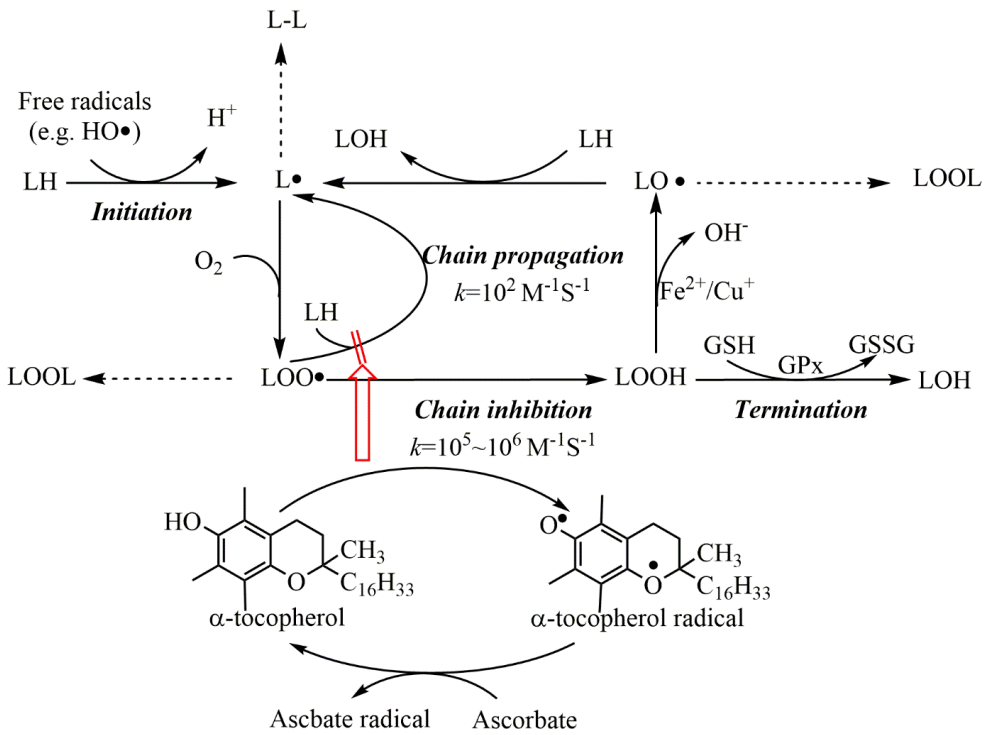


Figure 5 Lipid peroxidation chain reaction induced by reactive oxygen species

L: lipid radical, LOO: lipid peroxyl radical, LOOH: lipid hydroperoxides, LOOL: peroxide bridged dimer, L-L: fatty acid dimer, LOH: aldehydes (e.g., MDA, HNE), α-TocH: α-tocopherol, α-TocH•: α-tocopherol radical, Asc: ascorbic, Asc•: ascorbic radical

Three steps of lipid peroxide chain reaction:

Initiation: ROS (especially hydroxy radicals) initially trigger a reactive hydrogen atom extraction from the methylene group of polyunsaturated fatty acid forming carbon-centered radicals at the allylic position, and then react with oxygen molecule thus forming a peroxyl radical.

Propagation: Peroxyl radical, as an intermittent radical in the reaction chain, may further react with nearby lipids, yielding lipid hydroperoxides (LOOH), as well

antioxidant levels are mostly inversely associated with oxidants, for example, in C elegans, an inverted U-shaped dose-response relationship between antioxidants and lifespan was observed [57]. Consequently, the same situation might also be there for oxidative stress.

Damages caused by oxidative stress

Lipid peroxidation

Polyunsaturated fatty acid (PUFA), particularly with a high number of double bonds, such as arachidonic acid and linoleate, are highly susceptible to ROS and free radicals, known as autocatalytic chain reaction [58]. Lipid peroxidation includes

three steps: initiation, propagation, and termination, as shown in **Figure 5**. Phospholipids peroxidation in lipid membranes may lead to a decrease in membrane fluidity and permeability, and inactivation of receptors, resulting in cell apoptosis. Furthermore, lipid radicals generated during oxidation processes can form a myriad of deleterious end products, including reactive aldehydes, alkanes, and alkenes [59]. Among those products, malondialdehyde and 4-hydroxy-2-nonenal (4-HNE) are two of the most widely studied aldehydes. However, due to the high reactivity of these aldehydes, for example, they can react with proteins through the Michael addition reaction or with DNA to form adducts [60,61], it is difficult to accurately measure their free concentrations as valid oxidative damage levels. Moreover, urinary malondialdehyde levels are affected by dietary malondialdehyde content [62], and glutathione S-transferases genetic polymorphisms could confound the metabolism of 4-HNE [61]. Therefore, both the properties of the products themselves, xenobiotic sources, and genetic predispositions present obstacles for their development and further use as reliable biomarkers.

F_2-Isoprostanes (F_2-IsoPs), another important class of lipid peroxidation end products, can be formed through the non-enzymatic oxidation of arachidonic acid. These are prostaglandin-like compounds initially formed in-situ esterified to phospholipids and subsequently released in free form by phospholipases [63]. Isoprostanes have widely been considered as reliable biomarkers quantitating oxidative stress with their important merits of being (i) stable, both chemically and unaffected by diet or health status; (ii) sensitive to changes in oxidative stress; (iii) present in detectable quantities in many biological fluids, such as urine and plasma; and (iv) accessible to define a normal range in the population level which means the change of their formation can serve as the reflection of different oxidative stress levels and further relate to pathological status or diseases [64-66]. A meta-analysis identified different quantitative levels of 8-iso-prostane-$F_{2\alpha}$ in patients across different pathological conditions, with relatively small increased levels observed in patients with, for example, hypertension and metabolic syndrome, while large increased levels were observed with patients such as kidney related pathologies and respiratory tract disorders [67].

Protein oxidation

Proteins are also important targets for ROS. Protein oxidation involves (i) oxidative modification of site-specific amino acid residues, including oxidation of sulphur-containing amino acid residues, particularly cysteine and methionine, as well as reactions with aromatic amino acid residues and peroxynitrite; (ii) protein fragmentation resulting from oxidative cleavage of the peptide backbone; (iii) generation of protein carbonyl derivatives; and (iv) generation of protein-protein cross-linkages [68-70]. Protein oxidation may lead to a change in its three-dimensional structures, modification of physiological properties such as enzyme activities and signal transduction networks which are related to cellular regulation and function, and further protein proteolytic degradation on one hand, or protein aggregation, partial unfolding, and altered conformation on the other hand [70-72]. Carbonyl derivatives are a large group of end-products of protein oxidation that have been considered as the most widely used biomarkers of oxidative damage to protein. They can be formed through oxidative cleavage of backbones, direct oxidation of amino acid residues including lysine, arginine, proline, and threonine,

as well as the reaction of amino acid residues with aldehyde resulting from lipid peroxidation or carbohydrate [73]. Among the protein carbonyls, advanced glycation end products (AGEs) are a heterogenous group of additive derivatives produced by reactions between proteins and oxidation end products from PUFAs or carbohydrates, such as hazardous aldehyde [74]. Other protein oxidation end-products include advanced oxidation protein products (AOPP) formed mainly by myeloperoxidase-derived chlorinated oxidants [75], oxidized low-density lipoprotein (oxLDL) derived from oxidative modification [76], and nitrotyrosine resulting from tyrosine oxidation [77]. However, because of the structural heterogeneity caused by different biochemical pathways as well as the low specificity and sensitivity of antibodies in available detection assays, it is difficult to use them as stable and efficient biomarkers in large epidemiological studies.

Nuclear DNA oxidation

ROS, in particular the hydroxyl radicals, can cause oxidative damage to the DNA, which includes: (i) base mutation; (ii) strand breaking, both single and double; (iii) DNA-protein cross-links; and (iv) formation of DNA-adducts. In general, hydroxyl radicals could react with DNA bases and sugar-phosphate backbone, leading to erroneous base pairing and further common mutation of G to T and C to A bases [78]. Meanwhile, hydroxyl radicals could also react with the deoxyribose moiety leading to the loss of DNA bases, generating base-free sites, thus single or double DNA strand(s) could break. DNA strand breaks are well-established risk factors of genome instability, cell cycle disruption, as well as cell death [79-81]. DNA-protein cross-links involving thymine and tyrosine in the nucleoprotein complex of histones and DNA (nucleo-histone) can also be induced by the hydroxyl radicals [82]. In addition, adducts to DNA can be formed by the reaction of deoxyguanosine (M(1)G) and deoxyadenosine with other macromolecular modifications trigged by ROS, such as reactive aldehyde products generated during the lipid peroxidation [83].

Specific to the most common base mutations, there are mainly three ways: (i) adding double bonds to heterocyclic DNA bases; or (ii) abstracting a hydrogen atom from the methyl group of thymine; or (iii) forming allyl radical of thymine and carbon-centered sugar radicals [84,85]. For pyrimidines, hydroxyl radicals react with the particularly sensitive C5- and C6- double bond of thymine and cytosine, generating a spectrum of adducts, including thymine glycol, cytosine glycol, 5-hydroxymethyl uracil and 5-formyluracil, 5-hydroxycytosine, uracil glycol, and 5-hydroxyuracil [85,86]. For purines, similarly, C4-, C5, and C8- sites are more sensitive to radicals [85,86]. Among all these DNA oxidative modifications, the mutagenic lesions of saturated imidazole ring 7,8-dihydroxy-8-oxo-7,8-dihydroguanine (8-oxodG), formed from hydroxylation of the C8 residue of guanine, has been the most abundant, best characterized, and widely considered as a potential biomarker of DNA oxidation [85,87]. Theoretically, the concentration of 8-oxodG measured in urine, with its merits of long-term stability in urine and multiple measurement methods, reflects the accumulative DNA oxidation in the whole body and is thus considered being predictive for ageing and multiple diseases. However, unlike other molecules, oxidative damage to DNA might be repaired by the DNA repair systems [88]. So the ROS-induced 8-oxodG concentration is not only related to oxidative damage but also depends on the DNA repair capacity.

Yet, due to the unknown inter-individual difference in DNA repair capacity, it is difficult to quantify the real 8-oxodG concentration induced by ROS accurately in individuals, therefore, the further use of this biomarker has been hampered.

Mitochondrial dysfunction

The mitochondrial respiratory chain is the main intracellular place for endogenous ROS generation, with steady higher concentrations of free radicals in the mitochondrial matrix, and this leads to close proximity of mitochondria to ROS. Evidence showed that an about 5- to 10-fold higher superoxide anion level is present in the mitochondrial matrix than in the cytosolic and nuclear spaces [38], making the mitochondria primary targets for ROS-induced damages.

Mitochondrial DNA (mtDNA) casts a critical role in energy generation during oxidative phosphorylation by the function of encoding important bioenergetic genes including 13 polypeptides, 22 transfer RNA, 2 ribosomal RNA, which are essential for the synthesis of multi-subunit enzymatic components of the electron transport chain [89]. However, there are several factors ascribed that contributed to the higher vulnerability of mtDNA to oxidative damage: (i) close proximity to the oxidants forming site; (ii) the absence of histones in mtDNA; and (iii) limited mtDNA repair activity. This less active repair capacity is partly attributed to the multiple copies of mtDNA and a circular and compact coding arrangement in mtDNA [90]. Therefore, damaged mtDNA would be degraded and replaced by newly produced ones copied from undamaged genomes, which gives a higher ability to mitochondria against detrimental mutation effects. Meanwhile, there are also fewer mtDNA repair mechanisms than that in nuclear DNA and mostly are poorly understood. Though base excision repair, mismatch repair, as well as recombination or nonhomologous end joining repair mechanisms, do exist in mtDNA, nucleotide excision repair is still absent, and the specific roles of each pathway, as well as proteins involved in repair processes, remain to be fully elucidated [90,91]. It is estimated that oxidized bases are 10 times more frequent in mtDNA than nuclear DNA, for instance, the concentration of the common DNA lesion 8-oxodG is much higher in mtDNA than in nuclear DNA [36,92]. When a mutation occurs, mutated and normal mtDNA co-exist in the same mitochondria, which is known as heteroplasmy [93,94]. The percentage of an erroneous sequence of mtDNA changes during cell replication and division, and mitochondria could aggregate different mutations over time, reducing their bioenergetic capacity [95]. When the increasing proportion of mutated mtDNA exceeds the critical threshold level of normal mitochondrial energy generation, bioenergetic defects in cells may occur. Besides, the sustaining existence and accumulation of damaged mtDNA in the mitochondria will inevitably lead to more ROS generation, which in turn causes further damage, making a vicious cycle [92].

The mtDNA damages would lead to loss of redox homeostasis, perturbed Ca^{2+} homeostasis, damage to membrane proteins and lipids, as well as to abnormal mitochondrial energy transduction [95]. All these modifications are the driving force for further mitochondrial dysfunction and loss of integrity, which will affect cell viability and cellular function [96,97]. In addition, organs, especially the brain, heart, and muscles, being high-energy consuming and sensitive to bioenergetic defects, are strongly affected by mitochondrial dysfunctions, resulting in organ-specific

pathologies. Moreover, variations in mtDNA have also been shown to associate with several clinical phenotypes, including cardiovascular diseases, anthropometric and metabolic measures [98].

In addition to mtDNA, mitochondrial sirtuins have been found to be involved in the regulation of redox homeostasis. Sirtuins are a group of proteins (from SIRT1 to SIRT7), acting predominantly as nicotinamide adenine dinucleotide (NAD^+) dependent deacetylases and ADP-ribosyltransferases. SIRT3, 4, and 5 are exclusively localized within mitochondrial, altering the NAD/NADPH ratio in mitochondria to respond to metabolic changes. SIRT3 is supposed as the major mitochondrial deacetylase and is the most thoroughly studied mitochondrial sirtuin. It plays an important role in mitochondrial bioenergetics [99-102], as well as promotes resistance to oxidative stress by reducing ROS via magnesium SOD deacetylation and activation [99,100]. Though SIRT4 and SIRT5 are less well understood, SIRT5 has also been shown to be strongly associated with oxidative stress signaling and protect cells from ROS, suppressing oxidative stress levels [103]. In contrast, SIRT4 might have dual roles, either to induce ROS production or to have antioxidative function [103]. In addition, these mitochondrial sirtuins work in concert and present complicated interaction profiles [104], for example, SIRT3 and SIRT4 may function together to protect cells from stress and DNA damage [105]. A series of reviews regarding sirtuins on mitochondrial and oxidative regulation has been published before [99-102,106-108]. Nevertheless, additional studies are needed to in-depth ascertain and illustrate the mechanism of mitochondrial sirtuins in maintaining mitochondrial biological functions and redox homeostasis.

Causes and consequences of oxidative damage on ageing and age-related diseases

The progressive accumulation of oxidative damage to macromolecules and mitochondria will finally lead to pathophysiological alterations, functional decline, and accelerated ageing. Here we will discuss the harmful consequences of oxidative stress, mainly oxidative damage which are more closely related to accelerated ageing and multiple age-related diseases, with a special focus on ageing and lifespan, cardiovascular diseases (CVD), and neurodegenerative diseases (NNDs).

Ageing and lifespan

The best way to determine the effect of oxidative stress on ageing is to test whether changes in lifespan are dependent on alterations in oxidative stress levels. Numerous studies aimed to investigate this effect have been conducted on animals. Based on the free radical theory of ageing, attempts to extend lifespan have been carried out with a focus on two different targets: alterations of oxidant load (i.e. ROS generation) or alternations of mitochondrial antioxidant capacity. Alternations of oxidant load are usually investigated through calorie restriction, while alternation of mitochondrial antioxidant capacity is often examined through inducing genetic alterations, or through dietary supplementations.

Calorie reduction (CR) is defined as a 10-50% reduction of total energy intake without inducing malnutrition. A large body of evidence suggested that CR

might modulate the mitochondrial activity and lower ROS production resulting in the reduction of oxidative damage through sirtuins regulation, for example through primary activation of sirtuin 1 and further activation of PPARγ coactivator-1α (PGC-1α), leading to a slower rate of aging-related decline and extended lifespan [109-111]. Several reviews indicated that, in lower model organisms ranging from yeast to mammals, CR is capable to extend both average and maximum lifespan and health span [112-116]. Interestingly, studies in humans also detected some favorable biomarkers of longevity induced by CR, mainly hormonal adaptions, such as lower insulin, and lower IGF-1 level, though the latter was mainly in response to protein restriction [117,118].

With respect to induced genetic alternations, antioxidants-related transgenic lower models have been frequently used, which have altered antioxidant capacity or a disrupted oxidative-related signaling pathway, leading to either a reduced or extended lifespan [119-123]. Mice with antioxidant overexpression, for example, catalase [124], or with modifications in other signaling pathway components, such as $p66^{sch-/-}$ [125], $Igf1r^{+/-}$ [126] are generally considered to have increased longevity. However, other results are inconclusive about altered expression levels of, especially SOD [54,55,127,128], thioredoxin [129,130], or methionine sulfoxide reductase A ($MsrA^{-/-}$) that function to protect against protein oxidation [131,132]. Moreover, $Gpx1^{-/-}$ knocked-out mice with reduced glutathione peroxidase-1 surprisingly did not show accelerated ageing [55,128]. Similarly, results on the longevity of transgenic invertebrate models, such as C. elegans or drosophila, with alternations of antioxidants capacity, are also quite divergent, and some studies even found a large extension of lifespan with elevated oxidative stress level [56,133-140]. The inconsistence was also found in experiments with dietary antioxidants supplementation [141].

However, oxidative stress may act as a double-edged sword to human health, as discussed in part 2.2. It is difficult to determine a lifespan-affecting ROS level and this further lead to the ambiguity on the relation of oxidative stress and ageing. Nevertheless, a recent review still concludes that mitochondria play a critical role in the ageing process, but it remains unclear whether it is a cause or consequence [142].

Cardiovascular diseases (CVDs)

Cardiac myocytes have a high number of mitochondria which makes them more susceptible to oxidative damage. In addition to the mitochondrial pathway of ROS generation, crosstalk exists between mitochondria and NADPH oxidase, the major source of ROS in blood vessels, affecting vascular function by dysregulation or uncoupling eNOS, leading to endothelium dysfunction [143]. Together with other ROS-induced damages involved in mitochondrial dysfunction, and ROS-induced macromolecule damages (e.g. ox-LDL), they all contribute to the pathogenesis and progress of CVD [143,144]. To date, several lines of evidence have found an association between onset and progression of cardiovascular diseases (CVD), which include atherosclerosis [145], hypertension [146], heart failure [147], and peripheral artery disease [148].

In rodents, deletion of the NADPH oxidases gene (Nox1 and Nox2) or p47phx deficiency resulted in markedly lower ROS generation, and yielded a lower risk

of atherosclerosis, while Nox4 deletion gave an increased risk [149]. Similarly, over/under-expression of different antioxidant enzymes modulated atherogenesis in different genetically-altered mouse models [147,149]. In human studies, increased expression and activity of NADPH oxidase and xanthine oxidase subunits were associated with higher ROS production and an increased risk of coronary artery disease [150]. Moreover, mtDNA copy number, the increase of which is indicative of early molecular changes compensating for oxidative stress-induced mitochondrial defects, has been shown to be associated with CVD. A study composed of 21870 participants from three large cohorts found that the CVD incidence increased 23% with a one-standard deviation decrease in mtDNA copy number after adjustment for traditional CVD risk factors [151]. In addition, epidemiological cohort studies consistently showed that higher dietary intake of antioxidants, either as diet components or supplements, including flavonoids, vitamin E, and C, are associated with lower CVD risk, such as coronary heart disease, ischemic stroke as well as CVD mortality [152,153]. However, large randomized controlled clinical trials generally failed to show any beneficial effect of antioxidants supplementation, including vitamin E, A, C, B12, B6, and folic acid, on CVD risk, especially in atherosclerosis [153,154]. Meta-analyses also concluded that there is insufficient evidence to support vitamins or other antioxidants supplementation to decrease CVD incidence [30,155]. Interestingly, a very recent Mendelian Randomization study found there are causal associations between higher circulating vitamin E and the risk of coronary artery disease or myocardial infarction [156]. However, the three vitamin E-associated instrumental SNPs used in this study are all associated with blood lipid levels or coronary artery disease itself. This violates the important condition for conducting Mendelian Randomization that no horizontal pleiotropy for instruments. Though sensitivity analysis was performed, this is only for one SNP but not for others, thus the results are of limited value. Therefore, despite the biological evidence from consistent results, the premise of interventions with the intention to decrease oxidative levels and consequently CVD risk is still limited.

Neurodegenerative diseases (NDDs)

NDDs are characterized by dynamic and progressive neuronal cell damage and loss of neurons, typically including Alzheimer's disease (AD) and Parkinson's disease (PD). Ageing is considered to be the predominant risk factor for NDDs, and accumulative oxidative damages during ageing are the main culprits of neurological deterioration [96,157]. Neurological systems are extremely sensitive and vulnerable to ROS-induced damage due to their higher energy demand, lower rate of cellular renewal, membrane PUFA levels, as well as less active oxidative defense mechanisms [158]. Notably, studies suggested that mitochondrial dysfunction plays a causal role in the pathogenesis of NDDs [157,159]. Sirtuins, which provide a key role in mitochondrial function and oxidative stress regulation (part 2.3) are also associated with neuron fates, being either neuroprotective or neurotoxic [102,160,161]. A large number of proteins that are implicated in NDDs are found to have mitochondrial involvement. AD is pathologically characterized by extracellular deposition of amyloid- β peptide (Aβ) and intracellular neurofibrillary tangles composed of hyper-phosphorylated tau. Oxidative stress may activate signaling pathways, such as Jun amino-terminal kinase and p38 mitogen-activated protein kinase, and give rise to amyloid precursor protein and β-secretases to form Aβ,

as well as increase tau phosphorylation [96,162]. In turn, Aβ proteins would directly induce ROS generation through activating NADPH oxidase (as shown in Figure 2), further exacerbating neuronal damage, amplifying neurotoxicity [163]. Besides, mtDNA control-region adopted more mutations in 23 pathologically confirmed AD brain samples compared with 40 controls, and heteroplasmic mtDNA control-region mutations increased 63% on average in all AD brains, while markedly increased 130% in patients older than 80 years [164]. 1-methyl-4-phenylpyridinium (MPP+), the oxidized product of 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP), and rotenone, both of which are mitochondrial complex I inhibitors, have been demonstrated to result in clinical PD phenotype by producing pathological degeneration of dopaminergic neurons of the substantia nigra [165,166]. Other PD-related proteins including α-Synuclein, Parkin, DJ-1, PINK1, LRRK2, and HTRA2 were associated with mitochondrial dysfunction as well [96,167]. In addition, ROS-induced macromolecule damage, for example, MDA and HNE, lipid peroxidation end products, may also play an important role in NDDs [168-170]. This can be evidenced by the fact that increased levels of oxidative stress biomarkers and decreased levels of antioxidative biomarkers were observed in pharmacologically induced or transgenic animal models with AD or PD [168]. Several papers have speculated oxidative stress as the key component in the etiology of NNDs [96,157,158,167,168,171-174].

Many, but not all, observational studies indicated that a higher intake of antioxidants or supplements such as vitamin E could reduce the risk of NDDs [175-181]. Plasma vitamin E levels are associated with better cognitive performance [182], and patients with AD were found to have a lower nutrient status of vitamin E, C, B12 in blood and brain or cerebrospinal fluid [183,184]. Animal experimental studies also showed a preventive effect of vitamin E supplementation on AD risk [185]. However, most clinical trials have shown disappointing results in terms of the benefits on cognitive decline or dementia of antioxidants supplementation for NDDs, including vitamin E, C, B12, B6, and coenzyme Q10. Meta-analyses also suggested that there is no evidence of vitamin supplementation for prevention of cognitive decline or dementia in cognitively healthy adults [186], or for prevention of progression from mild cognitive impairment (MCI) to dementia, or for treatment of MCI, or for improvement of cognitive function in people with MCI or dementia due to AD [32,187]. Similar results were partly observed for PD [188,189]. A series of seminal papers on antioxidant therapy as well as the challenges in AD have been reviewed thoroughly [190-192].

In order to disentangle the causality of antioxidants and AD, two Mendelian Randomization studies were conducted recently. In European-ancestry individuals, there is no causal association between circulating vitamin E levels [193], as well as ascorbate, β-carotene, retinol [194,] and AD risk. However, the SNPs selected as vitamin E genetic instruments are also associated with lipid metabolism, and only every single SNP for ascorbate and retinol was selected, limiting the power of the analyses. Meanwhile, these instruments only explained a small portion of the total variance thus may not fully represent the circulating vitamin levels. It remains unclear to what extent vitamin levels are associated with oxidative stress. Therefore, nevertheless the compelling biological plausibility and observational evidence, there is also a lack of convincing data for the antioxidant supplementation aiming to prevent or treat NDDs.

Future Perspectives and Concluding Remarks

Reasons for the failure of antioxidant supplements in clinical trials

Based on the free radical theory and positive results from both experimental and observational studies, numerous clinical trials examining antioxidants, particularly vitamin E, have been conducted for the prevention and/or treatment of age-related morbidity and mortality. However, results are apparently disappointing as partly discussed earlier in part 3. Here, we will have an in-depth summary of the critical points in clinical trials with a focus on the widely-studied vitamin E supplementation.

Vitamin E is a well-known chain-breaking antioxidant, as shown in **Figure 5**. Alpha-tocopherol competitively reacts with lipid peroxyl radical with a higher reaction rate to inhibit chain propagation. Theoretically, abundant vitamin E will ameliorate oxidative damage by preventing lipid peroxidation. Not surprisingly, most observational studies specifically found detrimental/protective effects of high oxidative stress levels/antioxidants levels on disease status. However, oral vitamin E supplementation in middle-aged participants failed to demonstrate protective effects on neither primary and secondary prevention, nor on the treatment of various age-related diseases. Several potential reasons for this failure have been discussed extensively [195-198]. Data reported that the plasma vitamin E levels of participants with normal baseline blood tocopherol levels could only increase less than 2-3 times no matter the amount or duration of supplementation [199]. Therefore, when individuals have a relatively low oxidative level and/or sufficient or high antioxidative level at enrollment in a study, it is highly unlikely to generate many benefits via extra supplementation, which might be explained as a potential "functional concentration limit". This hypothesis is supported by studies in vitamin E supplementation stratified by haptoglobin genotype. Vitamin E supplementation provided cardiovascular-protective effects only in discriminatingly selected subgroups of individuals with both diabetes and Hp2-2 genotype [200-202], who had high levels of oxidative stress and inferior antioxidant protection, in whom supplementation decreased cardiovascular events by 34% [203]. Besides, a larger improvement was observed in nonalcoholic steatohepatitis patients carrying the Hp2 allele after vitamin E treatment compared with Hp1 allele carriers [204]. Still, biological mechanisms underlying these observations should be explored in greater detail in future studies.

Another potential reason for the failure of vitamin E supplementation in clinical trials might lie in the difference of forms of tocopherols. In many studies, for example, in observational studies, vitamin E has been used, but this "vitamin E" was not specific to α-tocopherol and it is more about the combination of complicated tocopherol or tocotrienol isomers, while only α-tocopherol can be called "vitamin E" given its proved ability to avoid vitamin E deficiency [205]. Natural vitamin E has the highest bioactivity in the human body compared to other tocopherols or tocotrienols. However, most of the oral supplementation of vitamin E is synthetic, with lower bioactivity than the natural one, characterized by preferentially non-oxidation metabolites in urine [206].

Besides, administration heterogeneity could result from various reasons and become critical in clinical trials, such as timing, monotherapy or not, dose and duration. Normally, it is hard to define a standardized strategy for supplement administration since chronic diseases often have an ambiguous onset time and the progression can take several years. Administration of the antioxidant supplements after irreversible ROS damage could have a negligible beneficial effect at that moment in time. Therefore, there might be a certain "critical window", and antioxidants administrated at an optimal timing before disease onset or during the early stages of disease could be able to ameliorate ROS-induced damage. Similarly, the duration of intervention needs to be sufficiently long to observe these effects.

Moreover, there is still a lack of consensus about the proper dosage regimen for antioxidant supplements. The doses used in clinical trials are derived from observational studies and are relatively safe and low. However, the dose for vitamin E to suppress plasma F_2-isoprostane concentration is 1600 IU or 3200 IU without observing any adverse effect after 16 weeks of supplementation, while the suppression of isoprostane were 35% and 49% respectively [207]. Importantly, the authors suggested that a 49% reduction with the largest dose of 3200 IU/day was not so profound. Davies et al also concluded from two modeling techniques, that only a sufficiently high concentration of non-enzymatic antioxidants would have a chance for collision and interaction with free radicals [208]. However, oral intake of vitamin E only modestly increases the plasma and tissue vitamin E levels, and this decreases the concentration of vitamin E at target cells or organelles and slows down the reaction rate with ROS, which was quite unlikely to affect health outcomes [209]. And also, endogenously present defense systems will weaken the effects of non-enzymatic antioxidants when the reaction rate with free radicals is below a necessary threshold [208]. Besides, some studies suggested vitamin E could also act as pro-oxidant after denoting a hydrogen atom to a lipid radical [210], so regenerating the reduced form from the oxidized form guarantees the lower depletion rate as well as the antioxidant function of vitamin E. Theoretically, combined administration of vitamin E and other antioxidants, such as ascorbic acid, coenzyme Q10 which can reduce oxidized vitamin E, might be protective. However, trials regarding this point are also inconsistent [28,177].

Ageing and age-related diseases are multifactorial, heterogeneous, and multidimensional. They do not respond to the organ-centered paradigm of "one cause - one mechanism - one disease/condition - one therapy". Therefore, a single therapeutic component that targets only one aspect of the several factors contributing to aging and age-related diseases, even if highly weighted like oxidative stress, might be less effective as the solution by itself. Similarly, it is highly possible that only one component (vitamin E) might not be enough to show physiological significant effects on the complex, heterogeneous pathways of organ dysfunction with increasing age.

In addition to these potential explanations for the lack of significant findings of vitamin E supplementation, some opinions argued that oxidative damage could merely play a role in the pathophysiology instead of being a direct cause of diseases [171,198,211]. However, emerging evidence points to the possible etiological role of oxidative stress, especially ROS-induced mitochondrial dysfunction in

NDDs [157,159,162]. To demonstrate the causality, the best way is to directly measure the association of ROS, or ROS-induced damage change with pathological disorders or diseases, or the improvement of clinical phenotypes led by prevention of oxidative damage [84,198,212]. Yet, it is almost impossible to measure total ROS levels in the human body because of their chemical properties of short-existence and high reactivity. Similarly, it remains a challenge to accurately measure oxidative damage because of the lack of reliable biomarkers as well as technical difficulties, further leading to the uncertainty of causality.

Potential oxidative stress biomarkers in clinical trials

After oral supplementation, vitamin E is digested and absorbed into the circulation by the body. However, plasma vitamin E concentrations after supplementation in a healthy population with a normal range of baseline vitamin E could largely vary from one individual to another which may arise from variations in the regulation of vitamin E uptake and metabolism between subjects, such as genetic factors, particularly apolipoprotein polymorphism [213]. Besides, though no causal association was shown in the previous Mendelian Randomization of circulating vitamin E and Alzheimer's disease [193], this may raise another argument that even circulating levels, especially induced by synthetic vitamin E, should not be considered the same as functional levels. First, vitamin E has two metabolic ways, to go either through non-enzymatic, namely free radical-dependent metabolism, opening the hydroxy-chromanol ring, or through enzymatic (mainly CYP4F2 of cytochrome P-450 family) metabolism with only successive shortening of phytyl side chain [214] (**Figure 6**). As synthetic vitamin E preferentially goes through the non-oxidation pathway, it is essential to determine whether or not, it acts as an antioxidant in the body. Second, oxidative damage is often located and gathered in certain tissues, specific cell types, or organelles. Although plasma circulating vitamin E is incorporated into lipoproteins and non-specifically transported to tissues, the acquisition of vitamin E seems to be largely mediated via the delivery and selective uptake of vitamin E after lipoprotein particles binding to receptor [215]. However, the mechanisms of the delivery and uptake of tissue vitamin E have been relatively poorly-characterized in other tissues except only for the hepatic tissue, and it appears to be essentially related to the expression and function of lipoprotein receptors. It is still hard to define the uptake level by many tissues precisely. Thus the accumulation of vitamin E could be different in various tissues, for instance, an only limited amount of vitamin E in the brain was observed after supplementation, or could be different for regions within the same organ, for example, the striatum contains the lowest amount of vitamin E compared with other areas in the brain [216]. In addition, once taken up by cells, the intracellular distribution of vitamin E to organelles is regulated by different transport proteins, but these proteins and their activity could differ from one tissue to another [215]. Hence the vitamin E concentration at different target tissues or organelles largely differ, and there might be insufficient concentration in certain target locations to counteract with overwhelmed oxidative stress damage, particularly in the high energy-consuming organs such as the brain. Obviously, the real functional vitamin E might be completely different with, to be more precisely, far less than the circulating vitamin E. Therefore, due to the differences in susceptibility to oxidative damage and response to vitamin E supplementation among individuals, even identical supplementation strategies

(e.g. vitamin E type, administration protocol) may possibly not lead to same health outcomes. What matters most should be the ability to take advantage of circulating vitamin E by the body, namely to what extent, it acts as antioxidants. In brief, these issues can be simplified as the detection and monitoring of the antioxidative function of vitamin E.

Ideally, there are two aspects to monitor the effect of vitamin E, one is to measure robust oxidative stress levels in a range of pathologies of specific targets, as well as its alterations corresponding to vitamin E supplementation, and the other is to measure the direct level change of vitamin E acting as antioxidants. However, the major limitation of the first approach is the identification of reliable biomarkers. An oxidative damage biomarker that relates to the pathology could be promising, such as MDA, HNE, carbonyls, and particularly isoprostanes. These biomarkers have been identified in experimental and population-based studies [67,217-219], together with the potential associations with ageing and age-related diseases. Nevertheless, the measured oxidative damages are often the result of the complex network of both endogenous and exogenous antioxidants systems. Moreover, these biomarkers are exclusive to certain macromolecule damage which cannot reflect the whole oxidative damage of the body. Although isoprostanes have been regarded as the most reliable biomarker with many merits mentioned in part 2.2.1 so far, several possible drawbacks impede its further use. In the pathway of isoprostane formation, direct oxygen addition would compete with the second 5-exo cyclization involving a carbon-centered reaction, so the formation relies on low oxygen tension which gives priority to the cyclization reaction [63,220]. Besides, due to the ubiquitous existence of F_2-isoprostane in normal tissues, the products derived from the kidney may confound the urine isoprostane concentration [63,66]. This could also be an explanation why a larger increase of isoprostane was found in the kidney pathologies than other disease conditions [67]. Apart from the generation pathway, quantitation of isoprostanes remains a challenge due to multiple steps in the detection method, which is labor-intensive, as well as due to the impossibility to specify each isomer [61]. All these drawbacks hindered the possibility as a biomarker in large epidemiological studies or clinical trials to assess the specific oxidative damage change caused by vitamin E accurately.

In recent years, vitamin E metabolites have been proposed to help better elucidate vitamin E roles [221]. Those metabolites reveal the biological process and reflect the authentic utility of vitamin E. Notably, identification of metabolites, especially free radical-dependent vitamin E metabolites, along with the determination of their levels may provide insights into oxidative stress-related physiological functions, as well as disease mechanisms and therapeutic strategies. When vitamin E acts as peroxyl lipid radical scavenger, it forms vitamin E radicals, which further react with lipid peroxides into α-tocopherol quinone (α-TQ) (**Figure 6**) [222]. In presence of NAD(P)H, α-TQ captures hydrogens converting into α-tocopherol hydroquinone (α-THQ), following the β-oxidation and cyclization of the phytyl side chain, α-tocopheronic acid and α-tocopherol lactone (α-TL) are generated [223]. Finally, α-TL are excreted as polar glucuronidated and sulfated conjugates of α-TL hydroquinone (α-TLHQ) glucuronide and α-TLHQ sulphate [224]. Theoretically, disease states that are associated with increased oxidative stress are likely to have higher antioxidants requirements, which would result in depletion of circu-

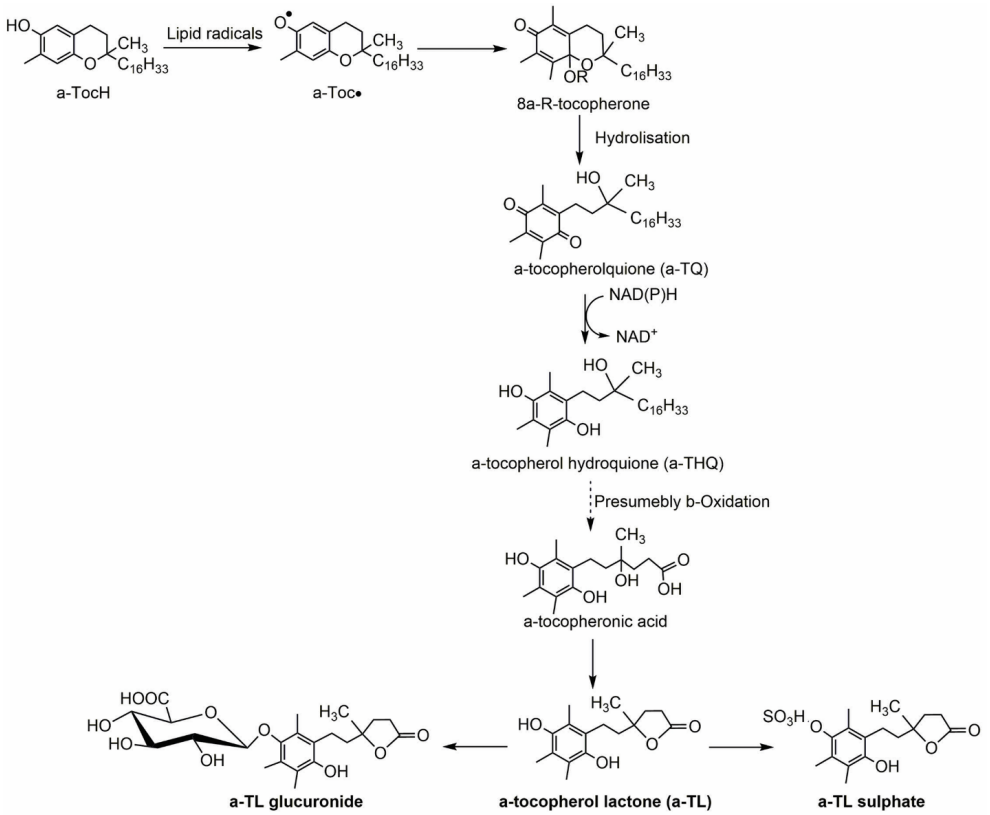


Figure 6 Alpha-tocopherol catabolism

The left panel shows the non-enzymatic metabolic pathway of α-tocopherol. Alpha-tocopherol reacts with lipid peroxyl radicals, forming α-tocopherol radical, which further reacts with lipid peroxides, following hydrolyzation, generating α-tocopherol quinone (α-TQ). In presence of NAD(P)H, α-TQ captures hydrogens converting into α-tocopherol hydroquinone (α-THQ), following the β-oxidation and cyclisation of the phytyl side chain, α-tocopheronic acid and α-tocopherol lactone (α-TL) are generated. Finally, α-TL are excreted as polar glucuronidate and sulfate conjugates of α-TL hydroquinone (α-TLHQ). The right panel presents the enzymatic metabolic pathway of α-tocopherol. This process is initiated in the liver, with the hydroxylation of the methyl group by hepatic CYP enzyme. Subsequently, the phytyl side chain shortens successively with the removal of carbon units in β-oxidation. Finally, α-carboxyethyl-hydroxychroman (α-CEHC) are generated and excreted as their polar glucuronide or sulfate conjugates.

lating vitamin E, and consequently, a higher concentration of oxidized vitamin E metabolites in urine, as a reflection of real antioxidant capacity.

Few studies have shown that oxidized vitamin E has been associated with diseases. A level of about 3% to 11% of α-tocopherol as an oxidized form of α-TQ was found in all lipoprotein density fractions prepared from advanced human atherosclerotic plaque [225]. Older individuals without baseline cognitive impairment and with the highest tertile of α-TQ/cholesterol had a higher risk of prevalent

dementia [226]. A recent study also indicated that both plasma α-TQ and 4-HNE were higher in participants with non-alcoholic fatty liver diseases, with lipid peroxidation as one of the earliest pathogenic events, compared with healthy participants [227]. Though it has been claimed that the formation of α-tocopheronic acid and α-TL could be artificial products from the oxidation of α carboxyethyl hydroxy chromans during the analytical process [228], Sharma and colleagues developed a feasible way to measure these conjugates by avoiding the artifacts [229], and found that the mean concentrations of α-TL conjugates were significantly higher in children with type 1 diabetes compared with age-matched controls [224]. Notably, all these studies were carried out in a small sample size, and it remains unclear whether the levels of the metabolites are consistent with the oxidative damage. Efforts are still required to validate the potential biomarkers in large population-based studies, and further use as clinical monitoring at the individual level after supplementation of vitamin E.

Remarks and Conclusions

There is an ongoing controversy in terms of antioxidant supplementation for the prevention and/or therapy of ageing and age-related diseases. However, large amounts of money are still spent, and tens of trials with different population sets and various administration strategies are in the plan, regarding single antioxidants clinical trials. Of note, “multifactor - multi treatments” might be more effective since there might be an additional or a synergistic benefit from two agents with acceptable safety and efficacy, for example, the combination of treatments (traditional treatment plus antioxidants treatment) to achieve better effects than antioxidants only. Interestingly, a recent trial showed that middle-aged type 2 diabetic patients who received metformin treatment plus vitamin E and/or vitamin C had significant improvement of glucose measures as well as lipid profiles compared to patients with metformin treatment alone, indicating that antioxidants use might be an adjuvant therapy in the management in type 2 diabetic patients [230]. But the combination use still requires additional efforts.

In summary, there are two critical points in clinical trials related to oxidative stress that need to be further discussed: (i) Target population. Participants should be selected by the stratification of certain features (e.g., genotype) which could induce significantly heterogeneous responses. So the selected participants subgroup are most likely to benefit from vitamin E supplementation. For example, vitamin E supplementation has been shown to be associated with an approximately 35% reduction in CVD specifically in individuals with both diabetes and Hp 2-2 [202]. Besides, a recent study also suggested that alpha-tocopherol supplementation was beneficial for cancer prevention only among carriers of homozygous low-activity alleles in the catechol-O-methyltransferase gene [231]. Therefore, there is still a need for further studies aiming to identify potential geno-/pheno- types that could be a determinant of the possible benefits of supplementations, as well as the elucidation of underlying mechanisms. (ii) Reliable biomarkers. To date, there is still little consensus about the best systemic biomolecular measurements of oxidative stress. An optimal biomarker should be an authentic reflection of overall redox status, easily accessible, reasonably stable, easy to be detected accurately. Moreover, since ROS has properties of the second messenger which is necessary for the maintenance of cellular homeo-

stasis, as well as normal physiological function, it remains obscure to define an optimal range of an oxidative damage-induced biomarker from physiological to pathological disorders, i.e. which level is detrimental for health or promotes longevity and metabolic health outcomes.

There is rather convincing evidence that oxidative damage has an important role in aging and the pathogenesis of multiple age-related diseases. Yet, clinical trials did not demonstrate the preventive or therapeutic role of antioxidants supplementation. However, this does not mean the failure of the oxidative stress theory and the use of antioxidants. On the contrary, it compels us to rethink this issue with several open-ended questions: (i) Are these associations reliable in observational studies which could be biased by confounders and reverse causality. (ii) How do we infer the causality of oxidative stress and diseases at population levels. (iii) What should be the best biomarker of oxidative stress in clinical trials. (iv) How to develop proper preventive or therapeutic strategies taking all the critical points into account. Therefore, research in the coming years should be devoted to more in-depth studies examining the role of a more direct measure of oxidative stress in biofluids in the pathogenesis of age-related diseases and the ageing process, and to clarify the causality lying between oxidative stress and age-related diseases and the ageing process.

References

1. K. N. Eggleston, V. R. Fuchs. The New Demographic Transition: Most Gains in Life Expectancy Now Realized Late in Life. J Econ Perspect 2012; 26(3): 137-56.
2. R. Rappuoli, M. Pizza, G. Del Giudice, E. De Gregorio. Vaccines, new opportunities for a new society. Proc Natl Acad Sci U S A 2014; 111(34): 12288-93.
3. T. Wyss-Coray. Ageing, neurodegeneration and brain rejuvenation. Nature 2016; 539(7628): 180-6.
4. T. Finkel, M. Serrano, M. A. Blasco. The common biology of cancer and ageing. Nature 2007; 448(7155): 767-74.
5. N. J. Samani, P. van der Harst. Biological ageing and cardiovascular disease. Heart 2008; 94(5): 537-9.
6. C. López-Otín, M.A. Blasco, L. Partridge, et al. The hallmarks of aging. Cell 2013; 153(6): 1194-217.
7. T. von Zglinicki. Role of oxidative stress in telomere length regulation and replicative senescence. Ann N Y Acad Sci 2000; 908: 99-110.
8. T. von Zglinicki. Oxidative stress shortens telomeres. Trends Biochem. Sci 2002; 27(7): 339-44.
9. E. H. Blackburn, E. S. Epel, J. Lin. Human telomere biology: A contributory and interactive factor in aging, disease risks, and protection. Science 2015; 350(6265): 1193-8.
10. H. Sies. 1 - Oxidative Stress: Introductory Remarks. In: Sies H, ed. Oxidative Stress. London: Academic Press; 1985: 1-8.
11. H. Sies. Oxidative stress: a concept in redox biology and medicine. Redox Biol 2015; 4: 180-3.
12. H. Sies, C. Berndt, D. P. Jones. Oxidative Stress. Annu Rev Biochem 2017; 86: 715-48.
13. M. Schieber, N. S. Chandel. Chandel. ROS function in redox signaling and oxidative stress. Curr Biol; 24(10): R453-R62.
14. K.M. Holmström, T. Finkel. Cellular mechanisms and physiological consequences of redox-dependent signalling. Nat Rev Mol Cell Biol 2014; 15: 411.
15. R. H. Burdon. Superoxide and hydrogen peroxide in relation to mammalian cell proliferation. Free Radic Biol Med 1995; 18(4): 775-94.
16. K. Hensley, K. A. Robinson, S. P. Gabbita, et al. Floyd. Reactive oxygen species, cell signaling, and cell injury. Free Radic Biol Med 2000; 28(10): 1456-62.
17. S. Kreuz, W. Fischle. Oxidative stress signaling to chromatin in health and disease. Epigenomics 2016; 8(6): 843-62.
18. D. Harman. The biologic clock: the mitochondria? J Am Geriatr Soc 1972; 20(4): 145-7.
19. R. M. Bostick, J. D. Potter, D. R. McKenzie, et al. Reduced risk of colon cancer with high intake of vitamin E: the Iowa Women's Health Study. Cancer Res 1993; 53(18): 4230-7.
20. E. B. Rimm, M. J. Stampfer, A. Ascherio, et al. Willett. Vitamin E consumption and the risk of coronary heart disease in men. N Engl J Med 1993; 328(20): 1450-6.
21. M. J. Stampfer, C. H. Hennekens, J. E. Manson, et al. Vitamin E consumption and the risk of coronary disease in women. N Engl J Med 1993; 328(20): 1444-9.

22. E. J. Mayer-Davis, T. Costacou, I. King, D. et al. Plasma and dietary vitamin E in relation to incidence of type 2 diabetes: The Insulin Resistance and Atherosclerosis Study (IRAS). Diabetes Care 2002; 25(12): 2172-7.
23. M. E. Wright, K. A. Lawson, S. J. Weinstein, 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): 1200-7.
24. F. J. Li, L. Shen, 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): 253-8.
25. R. Banks, J. R. Speakman, C. Selman. Vitamin E supplementation and mammalian lifespan. Mol Nutr Food Res 2010; 54(5): 719-25.
26. F. L. Muller, M. S. Lustgarten, Y. Jang, et al. Trends in oxidative aging theories. Free Radic Biol Med 2007; 43(4): 477-503.
27. E. R. Miller 3rd, R. Pastor-Barriuso, D. Dalal, et al. Meta-analysis: high-dosage vitamin E supplementation may increase all-cause mortality. Ann Intern Med 2005; 142(1): 37-46
28. G. Bjelakovic, D. Nikolova, L. L. Gluud, et al. Mortality in randomized trials of antioxidant supplements for primary and secondary prevention: systematic review and meta-analysis. JAMA 2007; 297(8): 842-57.
29. E. A. Klein, I. M. Thompson, Jr., C. M. Tangen, et al. Vitamin E and the risk of prostate cancer: the Selenium and Vitamin E Cancer Prevention Trial (SELECT). JAMA 2011; 306(14): 1549-56.
30. S. K. Myung, W. Ju, B. Cho, 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: f10.
31. L. Wang, H. D. Sesso, R. J. Glynn, 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): 915-23.
32. N. Farina, D. Llewellyn, Mgekn Isaac, et al. Vitamin E for Alzheimer's dementia and mild cognitive impairment. Cochrane Database of Syst Rev 2017; (4).
33. H. J. Kim, E. Giovannucci, B. Rosner, et al. Longitudinal and secular trends in dietary supplement use: Nurses' Health Study and Health Professionals Follow-Up Study, 1986-2006. J Acad Nutr Diet 2014; 114(3): 436-43.
34. E. D. Kantor, C. D. Rehm, M. Du, et al. Giovannucci. Trends in Dietary Supplement Use Among US Adults From 1999-2012. JAMA 2016; 316(14): 1464-74.
35. A. W. K. Yeung, N. T. Tzvetkov, Osama S. El-Tawil, et al. Abdel-Daim, Atanas G. Atanasov. Antioxidants: Scientific Literature Landscape Analysis. Oxid Med Cell Longev 2019; 2019: 11.
36. H. Cui, Y. Kong, H. Zhang. Oxidative stress, mitochondrial dysfunction, and aging. J Signal Transduct 2012; 2012: 646354.
37. D. Harman. Aging: a theory based on free radical and radiation chemistry. J Gerontol 1956; 11(3): 298-300.
38. E. Cadenas, K. J. Davies. Mitochondrial free radical generation, oxidative stress, and aging. Free Radic Biol Med 2000; 29(3-4): 222-30.
39. S. I. Liochev. Reactive oxygen species and the free radical theory of aging. Free Radic Biol Med 2013; 60: 1-4.

40. J. F. Turrens. Mitochondrial formation of reactive oxygen species. J Physiol 2003; 552(Pt 2): 335-44.
41. T. Finkel, N. J. Holbrook. Oxidants, oxidative stress and the biology of ageing. Nature 2000; 408(6809): 239-47.
42. A. Y. Andreyev, Y. E. Kushnareva, A. A. Starkov. Mitochondrial metabolism of reactive oxygen species. Biochemistry (Mosc) 2005; 70(2): 200-14.
43. R. Castagna, J. P. Eiserich, M. S. Budamagunta, et al. Hydroxyl radical from the reaction between hypochlorite and hydrogen peroxide. Atmos Environ 2008; 42(26): 6551-4.
44. B. Halliwell. Free radicals, reactive oxygen species and human disease: a critical evaluation with special reference to atherosclerosis. Br J Exp Pathol 1989; 70(6): 737-57.
45. C. B. F. Andersen, K. Stodkilde, K. L. Saederup, et al. Haptoglobin. Antioxid Redox Signal 2017; 26(14): 814-31.
46. A. P. Levy, R. Asleh, S. Blum, et al. Haptoglobin: basic and clinical aspects. Antioxid Redox Signal 2010; 12(2): 293-304.
47. M. Ristow, K. Schmeisser. Mitohormesis: Promoting Health and Lifespan by Increased Levels of Reactive Oxygen Species (ROS). Dose Response 2014; 12(2): 288-341.
48. C. R. Reczek, N. S. Chandel. ROS-dependent signal transduction. Curr Opin Cell Biol 2015; 33: 8-13.
49. J. Zhang, X. Wang, V. Vikash, et al. ROS and ROS-Mediated Cellular Signaling. Oxid Med Cell Longev 2016; 2016: 4350965.
50. S. Hekimi, J. Lapointe, Y. Wen. Taking a “good” look at free radicals in the aging process. Trends Cell Biol 2011; 21(10): 569-76.
51. V. Urbonaviciute, H. Luo, C. Sjowall, et al. Low Production of Reactive Oxygen Species Drives Systemic Lupus Erythematosus. Trends Mol Med 2019 Oct;25(10):826-835
52. L. M. Olsson, Å. C. Johansson, B. Gullstrand, et al. A single nucleotide polymorphism in the NCF1 gene leading to reduced oxidative burst is associated with systemic lupus erythematosus. Ann Rheum Dis 2017; 76(9): 1607.
53. T. Kelkka, D. Kienhöfer, M. Hoffmann, et al. Reactive Oxygen Species Deficiency Induces Autoimmunity with Type 1 Interferon Signature. Antioxid Redox Signal 2014; 21(16): 2231-45.
54. Y. C. Jang, V. I. Perez, W. Song, et al. Overexpression of Mn superoxide dismutase does not increase life span in mice. J Gerontol A Biol Sci Med Sci 2009; 64(11): 1114-25.
55. Y. Zhang, Y. Ikeno, W. Qi, et al. Mice deficient in both Mn superoxide dismutase and glutathione peroxidase-1 have increased oxidative damage and a greater incidence of pathology but no reduction in longevity. J Gerontol A Biol Sci Med Sci 2009; 64(12): 1212-20.
56. J. M. Van Raamsdonk, S. Hekimi. Deletion of the mitochondrial superoxide dismutase sod-2 extends lifespan in Caenorhabditis elegans. PLoS Genet 2009; 5(2): e1000361.
57. D. Desjardins, B. Cacho-Valadez, J. L. Liu, et al. Antioxidants reveal an inverted U-shaped dose-response relationship between reactive oxygen species levels and the rate of aging in Caenorhabditis elegans. Aging Cell 2017; 16(1): 104-12.
58. H. Yin, L. Xu, N. A. Porter. Free radical lipid peroxidation: mechanisms and analysis. Chem Rev 2011; 111(10): 5944-72.
59. B. C. Sousa, A. R. Pitt, C. M. Spickett. Chemistry and analysis of HNE and other prominent carbonyl-containing lipid oxidation compounds. Free Radic Biol Med 2017; 111: 294-308.

60. C. M. Spickett. The lipid peroxidation product 4-hydroxy-2-nonenal: Advances in chemistry and analysis. Redox Biol 2013; 1(1): 145-52.
61. D. Il'yasova, P. Scarbrough, I. Spasojevic. Urinary biomarkers of oxidative status. Clin. Chim. Acta 2012; 413(19-20): 1446-53.
62. E. D. Brown, V. C. Morris, D. G. Rhodes, et al. Urinary malondialdehyde-equivalents during ingestion of meat cooked at high or low temperatures. Lipids 1995; 30(11): 1053-6.
63. G. L. Milne, E. S. Musiek, J. D. Morrow. F2-isoprostanes as markers of oxidative stress in vivo: an overview. Biomarkers 2005; 10 Suppl 1: S10-23.
64. P. Montuschi, P. J. Barnes, L. J. Roberts, 2nd. Isoprostanes: markers and mediators of oxidative stress. Faseb j 2004; 18(15): 1791-800.
65. N. K. Gopaul, B. Halliwell, E. E. Anggard. Measurement of plasma F2-isoprostanes as an index of lipid peroxidation does not appear to be confounded by diet. Free Radic Res 2000; 33(2): 115-27.
66. L. J. Roberts, J. D. Morrow. Measurement of F(2)-isoprostanes as an index of oxidative stress in vivo. Free Radic Biol Med 2000; 28(4): 505-13.
67. T. J. van't Erve, M. B. Kadiiska, S. J. London, et al. Classifying oxidative stress by F2-isoprostane levels across human diseases: A meta-analysis. Redox Biol 2017; 12: 582-99..
68. E. R. Stadtman. Protein oxidation and aging. Free Radic Res 2006; 40(12): 1250-8.
69. B. S. Berlett, E. R. Stadtman. Protein oxidation in aging, disease, and oxidative stress. J Biol Chem 1997; 272(33): 20313-6.
70. M. J. Davies. Protein oxidation and peroxidation. Biochem J 2016; 473(7): 805-25.
71. A. Hohn, T. Jung, T. Grune. Pathophysiological importance of aggregated damaged proteins. Free Radic Biol Med 2014; 71: 70-89.
72. K. J. Davies. Protein oxidation and proteolytic degradation. General aspects and relationship to cataract formation. Adv Exp Med Biol 1990; 264: 503-11.
73. E. R. Stadtman, R. L. Levine. Free radical-mediated oxidation of free amino acids and amino acid residues in proteins. Amino Acids 2003; 25(3-4): 207-18.
74. R. Singh, A. Barden, T. Mori, et al. Advanced glycation end-products: a review. Diabetologia 2001; 44(2): 129-46.
75. V. Witko-Sarsat, M. Friedlander, T. Nguyen Khoa, et al. Advanced oxidation protein products as novel mediators of inflammation and monocyte activation in chronic renal failure. J Immunol 1998; 161(5): 2524-32.
76. A. Trpkovic, I. Resanovic, J. Stanimirovic, et al. Oxidized low-density lipoprotein as a biomarker of cardiovascular diseases. Crit Rev Clin Lab Sci 2015; 52(2): 70-85.
77. S. Bartesaghi, G. Ferrer-Sueta, G. Peluffo, et al. Protein tyrosine nitration in hydrophilic and hydrophobic environments. Amino Acids 2007; 32(4): 501-15.
78. A. P. Grollman, M. Moriya. Mutagenesis by 8-oxoguanine: an enemy within. Trends Genet 1993; 9(7): 246-9.
79. T. Yan, S. E. Berry, A. B. Desai, et al. DNA mismatch repair (MMR) mediates 6-thioguanine genotoxicity by introducing single-strand breaks to signal a G2-M arrest in MMR-proficient RKO cells. Clin Cancer Res 2003; 9(6): 2327-34.
80. V. Dehennaut, I. Loison, M. Dubuissez, et al. DNA double-strand breaks lead to activation of hypermethylated in cancer 1 (HIC1) by SUMOylation to regulate DNA repair. J Biol Chem 2013; 288(15): 10254-64.

81. P. A. Johnson, P. Clements, K. Hudson, et al. The mitotic spindle and DNA damage-induced apoptosis. Toxicol Lett 2000; 112-113: 59-67.
82. M. Dizdaroglu, E. Gajewski, P. Reddy, et al. Structure of a hydroxyl radical induced DNA-protein cross-link involving thymine and tyrosine in nucleohistone. Biochemistry 1989; 28(8): 3625-8.
83. L. J. Marnett. Oxy radicals, lipid peroxidation and DNA damage. Toxicology 2002; 181-182: 219-22.
84. C. C. Winterbourn. Reconciling the chemistry and biology of reactive oxygen species. Nat Chem Biol 2008; 4(5): 278-86.
85. M. Dizdaroglu, P. Jaruga, M. Birincioglu, et al. Free radical-induced damage to DNA: mechanisms and measurement. Free Radic Biol Med 2002; 32(11): 1102-15.
86. A. P. Breen, J. A. Murphy. Reactions of oxyl radicals with DNA. Free Radic Biol Med 1995; 18(6): 1033-77.
87. J. Cadet, T. Douki, D. Gasparutto, et al. Oxidative damage to DNA: formation, measurement and biochemical features. Mutat Res 2003; 531(1-2): 5-23.
88. N. Chatterjee, G. C. Walker. Mechanisms of DNA damage, repair, and mutagenesis. Environ Mol Mutagen 2017; 58(5): 235-63.
89. S. Anderson, A. T. Bankier, B. G. Barrell, et al. Sequence and organization of the human mitochondrial genome. Nature 1981; 290: 457.
90. N. M. Druzhyna, G. L. Wilson, S. P. LeDoux. Mitochondrial DNA repair in aging and disease. Mech Ageing Dev 2008; 129(7): 383-90.
91. D. L. Croteau, R. H. Stierum, V. A. Bohr. Mitochondrial DNA repair pathways. Mutat Res 1999; 434(3): 137-48.
92. B. N. Ames, M. K. Shigenaga, Tory M. Hagen. Mitochondrial decay in aging. Biochim Biophys 1995; 1271(1): 165-70.
93. T. Ozawa, M. Tanaka, W. Sato, et al. Mitochondrial DNA Mutations as an Etiology of Human Degenerative Diseases. In: Kim CH, Ozawa T, eds. Bioenergetics: Molecular Biology, Biochemistry, and Pathology. Boston, MA: Springer US; 1990: 413-27.
94. J. B. Stewart, P. F. Chinnery. The dynamics of mitochondrial DNA heteroplasmy: implications for human health and disease. Nat Rev Genet 2015; 16(9): 530-42.
95. D. C. Wallace, D. Chalkia. Mitochondrial DNA genetics and the heteroplasmy conundrum in evolution and disease. Cold Spring Harb Perspect Biol 2013; 5(11): a021220.
96. M. T. Lin, M. F. Beal. Mitochondrial dysfunction and oxidative stress in neurodegenerative diseases. Nature 2006; 443(7113): 787-95.
97. D. C. Wallace. A mitochondrial bioenergetic etiology of disease. J Clin Invest 2013; 123(4): 1405-12.
98. A. T. Kraja, C. Liu, J. L. Fetterman, et al. Associations of Mitochondrial and Nuclear Mitochondrial Variants and Genes with Seven Metabolic Traits. Am J Hum Genet 2019; 104(1): 112-38.
99. L. Guarente. Sirtuins, Aging, and Medicine. N Engl J Med 2011; 364(23): 2235-44.
100. M. N. Sack, T. Finkel. Mitochondrial metabolism, sirtuins, and aging. Cold Spring Harb Perspect Biol 2012; 4(12), pii: a013102.
101. S. I. Imai, L. Guarente. It takes two to tango: NAD(+) and sirtuins in aging/longevity control. NPJ Aging Mech Dis2016; 2: 16017.

102. C. H. Westphal, M. A. Dipp, L. Guarente. A therapeutic role for sirtuins in diseases of aging? Trends Biochem Sci 2007; 32(12): 555-60.
103. C. K. Singh, G. Chhabra, M. A. Ndiaye, et al. The Role of Sirtuins in Antioxidant and Redox Signaling. Antioxid Redox Signal 2018; 28(8): 643-61.
104. W. Yang, K. Nagasawa, C. Münch, et al. Mitochondrial Sirtuin Network Reveals Dynamic SIRT3-Dependent Deacetylation in Response to Membrane Depolarization. Cell 2016; 167(4): 985-1000.e21.
105. S. M. Jeong, C. Xiao, L. W. S. Finley, et al. SIRT4 Has Tumor-Suppressive Activity and Regulates the Cellular Metabolic Response to DNA Damage by Inhibiting Mitochondrial Glutamine Metabolism. Cancer Cell 2013; 23(4): 450-63.
106. L. Guarente. Mitochondria--a nexus for aging, calorie restriction, and sirtuins? Cell 2008; 132(2): 171-6.
107. W. Grabowska, E. Sikora, A. B. Zmijewska. Sirtuins, a promising target in slowing down the ageing process. Biogerontology 2017; 18(4): 447-76.
108. L. Guarente. Sirtuins, aging, and metabolism. Cold Spring Harb Symp Quant Biol 2011; 76: 81-90.
109. G. López-Lluch, Pl. Navas. Calorie restriction as an intervention in ageing. The J Physiol 2016; 594(8): 2043-60.
110. A. Zullo, E. Simone, M. Grimaldi, et al. Sirtuins as Mediator of the Anti-Ageing Effects of Calorie Restriction in Skeletal and Cardiac Muscle. Int J Mol Sci 2018; 19(4): 928.
111. L. Guarente. Calorie restriction and sirtuins revisited. Genes Dev 2013; 27(19): 2072-85.
112. G. Testa, F. Biasi, G. Poli, et al. Calorie restriction and dietary restriction mimetics: a strategy for improving healthy aging and longevity. Curr Pharm Des 2014; 20(18): 2950-77.
113. L. Fontana, L. Partridge, V. D. Longo. Extending healthy life span--from yeast to humans. Science 2010; 328(5976): 321-6.
114. D. K. Ingram, R. de Cabo. Calorie restriction in rodents: Caveats to consider. Ageing Res Rev 2017; 39: 15-28.
115. J. Most, V. Tosti, L. M. Redman, et al. Calorie restriction in humans: An update. Ageing Res Rev 2017; 39: 36-45.
116. L. K. Heilbronn, E. Ravussin. Calorie restriction and aging: review of the literature and implications for studies in humans. Am J Clin Nutr 2003; 78(3): 361-9.
117. L. K. Heilbronn, L. de Jonge, M. I. Frisard, et al. Effect of 6-month calorie restriction on biomarkers of longevity, metabolic adaptation, and oxidative stress in overweight individuals: a randomized controlled trial. JAMA 2006; 295(13): 1539-48.
118. D. Lettieri-Barbato, E. Giovannetti, K. Aquilano. Effects of dietary restriction on adipose mass and biomarkers of healthy aging in human. Aging (Albany NY) 2016; 8(12): 3341-55.
119. C. Y. Liao, B. K. Kennedy. Mouse models and aging: longevity and progeria. Curr Top Dev Biol 2014; 109: 249-85.
120. H. Liang, E. J. Masoro, J. F. Nelson, et al. Genetic mouse models of extended lifespan. Exp Gerontol 2003; 38(11-12): 1353-64.
121. A. B. Salmon, A. Richardson, V. I. Perez. Update on the oxidative stress theory of aging: does oxidative stress play a role in aging or healthy aging? Free Radic Biol Med 2010; 48(5): 642-55.

122. Y. Chen, C. Wu, C. Kao, et al. Longevity and lifespan control in mammals: Lessons from the mouse. Ageing Res Rev 2010; 9: S28-S35.
123. C. Selman, D. J. Withers. Mammalian models of extended healthy lifespan. Philos Trans R Soc Lond B Biol Sci 2011; 366(1561): 99-107.
124. D. F. Dai, Y. A. Chiao, G. M. Martin, et al. Mitochondrial-Targeted Catalase: Extended Longevity and the Roles in Various Disease Models. Prog Mol Biol Transl Sci 2017; 146: 203-41.
125. E. Migliaccio, M. Giorgio, S. Mele, et al. The p66shc adaptor protein controls oxidative stress response and life span in mammals. Nature 1999; 402(6759): 309-13.
126. R. K. Junnila, E. O. List, D. E. Berryman, et al. The GH/IGF-1 axis in ageing and longevity. Nat Rev Endocrinol 2013; 9(6): 366-76.
127. V. I. Perez, H. Van Remmen, A. Bokov, et al. The overexpression of major antioxidant enzymes does not extend the lifespan of mice. Aging Cell 2009; 8(1): 73-5.
128. H. Van Remmen, Y. Ikeno, M. Hamilton, et al. Life-long reduction in MnSOD activity results in increased DNA damage and higher incidence of cancer but does not accelerate aging. Physiol Genomics 2003; 16(1): 29-37.
129. G. M. Cunningham, L. C. Flores, M. G. Roman, et al. Thioredoxin overexpression in both the cytosol and mitochondria accelerates age-related disease and shortens lifespan in male C57BL/6 mice. Geroscience 2018; 40(5-6): 453-68.
130. V. I. Perez, L. A. Cortez, C. M. Lew, et al. Thioredoxin 1 overexpression extends mainly the earlier part of life span in mice. J Gerontol A Biol Sci Med Sci 2011; 66(12): 1286-99.
131. J. Moskovitz, S. Bar-Noy, W. M. Williams, et al. Methionine sulfoxide reductase (MsrA) is a regulator of antioxidant defense and lifespan in mammals. Proc Natl Acad Sci U S A 2001; 98(23): 12920-5.
132. A. B. Salmon, V. I. Perez, A. Bokov, et al. Lack of methionine sulfoxide reductase A in mice increases sensitivity to oxidative stress but does not diminish life span. Faseb j 2009; 23(10): 3601-8.
133. R. Doonan, J. J. McElwee, F. Matthijssens, et al. Against the oxidative damage theory of aging: superoxide dismutases protect against oxidative stress but have little or no effect on life span in Caenorhabditis elegans. Genes Dev 2008; 22(23): 3236-41.
134. W. Yang, J. Li, S. Hekimi. A Measurable increase in oxidative damage due to reduction in superoxide detoxification fails to shorten the life span of long-lived mitochondrial mutants of Caenorhabditis elegans. Genetics 2007; 177(4): 2063-74.
135. C. Yee, W. Yang, S. Hekimi. The Intrinsic Apoptosis Pathway Mediates the Pro-Longevity Response to Mitochondrial ROS in C.elegans. Cell 2014; 157(4): 897-909.
136. M. Uno, E. Nishida. Lifespan-regulating genes in C. elegans. NPJ Aging Mech Dis 2016; 2: 16010.
137. T. Oberacker, J. Bajorat, S. Ziola, et al. Enhanced expression of thioredoxin-interacting-protein regulates oxidative DNA damage and aging. FEBS Letters 2018; 592(13): 2297-307.
138. I. Reveillaud, J. Phillips, B. Duyf, et al. Phenotypic rescue by a bovine transgene in a Cu/Zn superoxide dismutase-null mutant of Drosophila melanogaster. Mol Cell Biol 1994; 14(2): 1302-7.
139. W. Yang, S. Hekimi. A mitochondrial superoxide signal triggers increased longevity in Caenorhabditis elegans. PLoS Biol 2010; 8(12): e1000556.

140. A. B. Hwang, E. A. Ryu, M. Artan, et al. Feedback regulation via AMPK and HIF-1 mediates ROS-dependent longevity in Caenorhabditis elegans. Proc Natl Acad Sci U S A 2014; 111(42): E4458.
141. I. Sadowska-Bartosz, G. Bartosz. Effect of antioxidants supplementation on aging and longevity. Biomed Res Int 2014; 2014: 404680.
142. A. Sanz. Mitochondrial reactive oxygen species: Do they extend or shorten animal lifespan?Biochim Biophys Acta 2016; 1857(8): 1116-26.
143. A. Daiber, F. Di Lisa, M. Oelze, et al. Crosstalk of mitochondria with NADPH oxidase via reactive oxygen and nitrogen species signalling and its role for vascular function. Br J Pharmacol 2017; 174(12): 1670-89.
144. A. Daiber. Redox signaling (cross-talk) from and to mitochondria involves mitochondrial pores and reactive oxygen species. Biochim Biophys Acta 2010; 1797(6-7): 897-906.
145. A. J. Kattoor, N. V. K. Pothineni, D. Palagiri, et al. Oxidative Stress in Atherosclerosis. Curr Atheroscler Rep 2017; 19(11): 42.
146. S. Chrissobolis, Q. N. Dinh, G. R. Drummond, et al. Role of Oxidative Stress in Hypertension. In: Rodriguez-Porcel M, Chade AR, Miller JD, eds. Studies on Atherosclerosis. Boston, MA: Springer US; 2017: 59-78.
147. T. Munzel, G. G. Camici, C. Maack, et al. Impact of Oxidative Stress on the Heart and Vasculature: Part 2 of a 3-Part Series. J Am Coll Cardiol 2017; 70(2): 212-29.
148. S. Steven, A. Daiber, J. F. Dopheide, et al. Peripheral artery disease, redox signaling, oxidative stress - Basic and clinical aspects. Redox Biol 2017; 12: 787-97.
149. U. Forstermann, N. Xia, H. Li. Roles of Vascular Oxidative Stress and Nitric Oxide in the Pathogenesis of Atherosclerosis. Circ Res 2017; 120(4): 713-35.
150. T. J. Guzik, J. Sadowski, B. Guzik, et al. Coronary artery superoxide production and nox isoform expression in human coronary artery disease. Arterioscler Thromb Vasc Biol 2006; 26(2): 333-9.
151. F. N. Ashar, Y. Zhang, R. J. Longchamps, et al. Association of Mitochondrial DNA Copy Number With Cardiovascular Disease. JAMA Cardiol 2017; 2(11): 1247-55.
152. Z. Ye, H. Song. Antioxidant vitamins intake and the risk of coronary heart disease: meta-analysis of cohort studies. Eur J Cardiovasc Prev Rehabil 2008; 15(1): 26-34.
153. H. N. Siti, Y. Kamisah, J. Kamsiah. The role of oxidative stress, antioxidants and vascular inflammation in cardiovascular disease (a review). Vascul Pharmacol 2015; 71: 40-56.
154. T. Gori, T. Munzel. Oxidative stress and endothelial dysfunction: therapeutic implications. Ann Med 2011; 43(4): 259-72.
155. H. Boeing, M. Gottschald, S. Dietrich, et al. Dietary Supplements and Risk of Cause-Specific Death, Cardiovascular Disease, and Cancer: A Systematic Review and Meta-Analysis of Primary Prevention Trials. Adv Nutr 2017; 8(1): 27-39.
156. T. Wang, L. Xu. Circulating Vitamin E Levels and Risk of Coronary Artery Disease and Myocardial Infarction: A Mendelian Randomization Study. Nutrients 2019; 11(9), 2153(9).
157. M. T. Islam. Oxidative stress and mitochondrial dysfunction-linked neurodegenerative disorders. Neurol Res 2017; 39(1): 73-82.
158. A. C. Rego, C. R. Oliveira. Mitochondrial dysfunction and reactive oxygen species in excitotoxicity and apoptosis: implications for the pathogenesis of neurodegenerative diseases. Neurochem Res 2003; 28(10): 1563-74.

159. C. Cheignon, M. Tomas, D. Bonnefont-Rousselot, et al. Oxidative stress and the amyloid beta peptide in Alzheimer's disease. Redox Biol 2018; 14: 450-64.
160. L. Sansone, V. Reali, L. Pellegrini, et al. SIRT1 silencing confers neuroprotection through IGF-1 pathway activation. J Cell Physiol 2013; 228(8): 1754-61..
161. A. Satoh, S. I. Imai, L. Guarente. The brain, sirtuins, and ageing. Nat Rev Neurosci 2017; 18: 362.
162. Y. Zhao, B. Zhao. Oxidative stress and the pathogenesis of Alzheimer's disease. Oxid Med Cell Longev 2013; 2013: 316523.
163. Z. Liu, T. Zhou, A. C. Ziegler, et al. Oxidative Stress in Neurodegenerative Diseases: From Molecular Mechanisms to Clinical Applications. Oxid Med Cell Longev 2017; 2017: 2525967.
164. P. E. Coskun, M. F. Beal, D. C. Wallace. Alzheimer's brains harbor somatic mtDNA control-region mutations that suppress mitochondrial transcription and replication. Proc Natl Acad Sci U S A 2004; 101(29): 10726-31.
165. R. Betarbet, T. B. Sherer, G. MacKenzie, et al. Chronic systemic pesticide exposure reproduces features of Parkinson's disease. Nat Neurosci 2000; 3(12): 1301-6.
166. F. Fornai, O. M. Schluter, P. Lenzi, et al. Parkinson-like syndrome induced by continuous MPTP infusion: convergent roles of the ubiquitin-proteasome system and alpha-synuclein. Proc Natl Acad Sci U S A 2005; 102(9): 3413-8.
167. T. Jiang, Q. Sun, S. Chen. Oxidative stress: A major pathogenesis and potential therapeutic target of antioxidative agents in Parkinson's disease and Alzheimer's disease. Prog Neurobiol 2016; 147: 1-19.
168. E. Niedzielska, I. Smaga, M. Gawlik, et al. Oxidative Stress in Neurodegenerative Diseases. Mol Neurobiol 2016; 53(6): 4094-125.
169. D. A. Butterfield, M. L. Bader Lange, R. Sultana. Involvements of the lipid peroxidation product, HNE, in the pathogenesis and progression of Alzheimer's disease. Biochim Biophys Acta 2010; 1801(8): 924-9.
170. F. Coppede, L. Migliore. DNA damage in neurodegenerative diseases. Mutat Res 2015; 776: 84-97.
171. G. H. Kim, J. E. Kim, S. J. Rhie, S. Yoon. The Role of Oxidative Stress in Neurodegenerative Diseases. Exp Neurobiol 2015; 24(4): 325-40.
172. E. Khusnutdinova, I. Gilyazova, E. Ruiz-Pesini, et al. A Mitochondrial Etiology of Neurodegenerative Diseases: Evidence from Parkinson's Disease. Ann N Y Acad Sci 2008; 1147(1): 1-20.
173. E. Mariani, M. C. Polidori, A. Cherubini, et al. Oxidative stress in brain aging, neurodegenerative and vascular diseases: an overview. J Chromatogr B Analyt Technol Biomed Life Sci 2005; 827(1): 65-75.
174. J. Blesa, I. Trigo-Damas, A. Quiroga-Varela, et al. Oxidative stress and Parkinson's disease. Front Neuroanat 2015 Jul 8; 9(91).
175. M. C. Morris, L. A. Beckett, P. A. Scherr, et al. Vitamin E and vitamin C supplement use and risk of incident Alzheimer disease. Alzheimer Dis Assoc Disord 1998; 12(3): 121-6.
176. M. C. Morris, D. A. Evans, J. L. Bienias, et al. Dietary intake of antioxidant nutrients and the risk of incident Alzheimer disease in a biracial community study. JAMA 2002; 287(24): 3230-7.

177. P. P. Zandi, J. C. Anthony, A. S. Khachaturian, et al. Reduced risk of Alzheimer disease in users of antioxidant vitamin supplements: the Cache County Study. Arch Neurol 2004; 61(1): 82-8.
178. S. M. Zhang, M. A. Hernan, H. Chen, et al. Intakes of vitamins E and C, carotenoids, vitamin supplements, and PD risk. Neurology 2002; 59(8): 1161-9.
179. M. J. Engelhart, M. I. Geerlings, A. Ruitenberg, et al. Dietary intake of antioxidants and risk of Alzheimer disease. JAMA 2002; 287(24): 3223-9.
180. J. A. Luchsinger, M. X. Tang, S. Shea, et al. Antioxidant vitamin intake and risk of Alzheimer disease. Arch Neurol 2003; 60(2): 203-8.
181. K. Yaffe, T. E. Clemons, W. L. McBee, et al. Impact of antioxidants, zinc, and copper on cognition in the elderly: a randomized, controlled trial. Neurology 2004; 63(9): 1705-7.
182. V. Boccardi, M. Baroni, F. Mangialasche, et al. Vitamin E family: Role in the pathogenesis and treatment of Alzheimer's disease. Alzheimers Dement 2016; 2(3): 182-91.
183. M. C. de Wilde, B. Vellas, E. Girault, et al. Lower brain and blood nutrient status in Alzheimer's disease: Results from meta-analyses. Alzheimers Dement 2017; 3(3): 416-31.
184. S. Lopes da Silva, B. Vellas, S. Elemans, et al. Plasma nutrient status of patients with Alzheimer's disease: Systematic review and meta-analysis. Alzheimers Dement 2014; 10(4): 485-502.
185. A. Gugliandolo, P. Bramanti, E. Mazzon. Role of Vitamin E in the Treatment of Alzheimer's Disease: Evidence from Animal Models. Int J Mol Sci 2017; 18(12).
186. A. W. S. Rutjes, D. A. Denton, M. Di Nisio, et al. Vitamin and mineral supplementation for maintaining cognitive function in cognitively healthy people in mid and late life. Cochrane Database Syst Rev 2018; (12).
187. J. McCleery, R. P. Abraham, D. A. Denton, et al. Vitamin and mineral supplementation for preventing dementia or delaying cognitive decline in people with mild cognitive impairment. Cochrane Database Syst Rev 2018; (11).
188. Z. Zhu, M. Sun, W. Zhang, et al. The efficacy and safety of coenzyme Q10 in Parkinson's disease: a meta-analysis of randomized controlled trials. Neurol Sci 2017; 38(2): 215-24.
189. A. Negida, A. Menshawy, G. El Ashal, et al. Coenzyme Q10 for Patients with Parkinson's Disease: A Systematic Review and Meta-Analysis. CNS Neurol Disord Drug Targets 2016; 15(1): 45-53.
190. P. Mecocci, V. Boccardi, R. Cecchetti, et al. A Long Journey into Aging, Brain Aging, and Alzheimer's Disease Following the Oxidative Stress Tracks. J Alzheimers Dis 2018; 62(3): 1319-35.
191. P. Mecocci, M. C. Polidori. Antioxidant clinical trials in mild cognitive impairment and Alzheimer's disease. Biochim Biophys Acta 2012; 1822(5): 631-8.
192. M. C. Polidori, G. Nelles. Antioxidant clinical trials in mild cognitive impairment and Alzheimer's disease - challenges and perspectives. Curr Pharm Des 2014; 20(18): 3083-92.
193. G. Liu, Y. Zhao, S. Jin, et al. Circulating vitamin E levels and Alzheimer's disease: a Mendelian randomization study. Neurobiol Aging 2018; 72: 189.e1-.e9.
194. D. M. Williams, S. Hägg, N. L. Pedersen. Circulating antioxidants and Alzheimer disease prevention: a Mendelian randomization study. Am J Clin Nutr 2018; 109(1): 90-8.
195. G. J. Brewer. Why vitamin E therapy fails for treatment of Alzheimer's disease. J Alzheimers Dis 2010; 19(1): 27-30.

196. S. R. Steinhubl. Why have antioxidants failed in clinical trials? Am J Cardiol 2008; 101(10a): 14d-9d.
197. I. Robinson, D. G. de Serna, A. Gutierrez, et al. Vitamin E in humans: an explanation of clinical trial failure. Endocr Pract 2006; 12(5): 576-82.
198. M. P. Murphy. Antioxidants as therapies: can we improve on nature? Free Radic Biol Med 2014; 66: 20-3.
199. A. Munteanu, J. M. Zingg, A. Azzi. Anti-atherosclerotic effects of vitamin E--myth or reality? J Cell Mol Med 2004; 8(1): 59-76.
200. M. Vardi, N. S. Levy, A. P. Levy. Vitamin E in the prevention of cardiovascular disease: the importance of proper patient selection. J Lipid Res 2013; 54(9): 2307-14.
201. J. P. Seferovic, M. Asanin, P. M. Seferovic. Haptoglobin and haptoglobin genotypes in diabetes: A silver bullet to identify the responders to antioxidant therapy? Eur J Prev Cardiol 2018; 25(14): 1498-501.
202. I. Hochberg, E. M. Berinstein, U. Milman, et al. Interaction Between the Haptoglobin Genotype and Vitamin E on Cardiovascular Disease in Diabetes. Curr Diab Rep 2017; 17(6): 42.
203. R. Asleh, A. Briasoulis, E. M. Berinstein, et al. Meta-analysis of the association of the haptoglobin genotype with cardiovascular outcomes and the pharmacogenomic interactions with vitamin E supplementation. Pharmgenomics Pers Med 2018; 11: 71-82.
204. B. A. Banini, S. C. Cazanave, K. P. Yates, et al. Haptoglobin 2 Allele is Associated With Histologic Response to Vitamin E in Subjects With Nonalcoholic Steatohepatitis. J Clin Gastroenterol 2018; 53(10):750-758.
205. A. Azzi. Many tocopherols, one vitamin E. Mol Aspects Med 2018; 61: 92-103.
206. M. G. Traber, A. Elsner, R. Brigelius-Flohé. Synthetic as compared with natural vitamin E is preferentially excreted as α-CEHC in human urine: studies using deuterated α-tocopheryl acetates. FEBS Letters 1998; 437(1): 145-8.
207. L. J. Roberts 2nd, J. A. Oates, M. F. Linton, et al. The relationship between dose of vitamin E and suppression of oxidative stress in humans. Free Radic Biol Med 2007; 43(10): 1388-93.
208. A. M Davies, A. Holt. Why antioxidant therapies have failed in clinical trials. J Theor Biol 2018; 457: 1-5.
209. N. G. Stephens, A. Parsons, P. M. Schofield, et al. Randomised controlled trial of vitamin E in patients with coronary disease: Cambridge Heart Antioxidant Study (CHAOS). Lancet 1996; 347(9004): 781-6.
210. A. Kontush, B. Finckh, B. Karten, et al. Antioxidant and prooxidant activity of alpha-tocopherol in human plasma and low density lipoprotein. J Lipid Res 1996; 37(7): 1436-48.
211. E. Niki. Evidence for beneficial effects of vitamin E. Korean J Intern Med 2015; 30(5): 571-9.
212. M. P. Murphy, A. Holmgren, N. G. Larsson, et al. Unraveling the biological roles of reactive oxygen species. Cell Metab 2011; 13(4): 361-6.
213. H. E. Roxborough, G. W. Burton, F. J. Kelly. Inter- and intra-individual variation in plasma and red blood cell vitamin E after supplementation. Free Radic Res 2000; 33(4): 437-45.
214. P. Torquato, O. Ripa, D. Giusepponi, et al. Analytical strategies to assess the functional metabolome of vitamin E. J Pharm Biomed Anal 2016; 124: 399-412.
215. M. Hacquebard, Y. A. Carpentier. Vitamin E: absorption, plasma transport and cell uptake. Curr Opin Clin Nutr Metab Care 2005; 8(2): 133-8.

216. B. Shukitt-Hale, G. Cao, J. A. Joseph, et al. Effect of Fruits, Vegetables, or Vitamin E–Rich Diet on Vitamins E and C Distribution in Peripheral and Brain Tissues: Implications for Brain Function. J Gerontol A Biol Sci Med Sci 2000; 55(3): B144-B51.
217. M. L. Bartoli, F. Novelli, F. Costa, et al. Malondialdehyde in exhaled breath condensate as a marker of oxidative stress in different pulmonary diseases. Mediators Inflamm 2011; 2011: 891752.
218. S. Dalleau, M. Baradat, F. Guéraud, et al. Cell death and diseases related to oxidative stress:4-hydroxynonenal (HNE) in the balance. Cell Death Differ 2013; 20: 1615.
219. N. A. Strobel, R. G. Fassett, S. A. Marsh, et al. Oxidative stress biomarkers as predictors of cardiovascular disease. Int J Cardiol 2011; 147(2): 191-201.
220. J. P. Fessel, N. A. Porter, K. P. Moore, et al. Discovery of lipid peroxidation products formed in vivo with a substituted tetrahydrofuran ring (isofurans) that are favored by increased oxygen tension. Proc Natl Acad Sci U S A 2002; 99(26): 16713-8.
221. F. Galli, A. Azzi, M. Birringer, et al. Vitamin E: Emerging aspects and new directions. Free Radic Biol Med 2017; 102: 16-36.
222. E. Herrera, C. Barbas. Vitamin E: action, metabolism and perspectives. J Physiol Biochem 2001; 57(1): 43-56.
223. R. Brigelius-Flohe, M. G. Traber. Vitamin E: function and metabolism. Faseb j 1999; 13(10): 1145-55.
224. G. Sharma, D. P. Muller, S. M. O'Riordan, et al. Urinary conjugated α-tocopheronolactone—a biomarker of oxidative stress in children with type 1 diabetes. Free Radic Biol Med 2013; 55: 54-62.
225. X. Niu, V. Zammit, J. M. Upston, et al. Coexistence of oxidized lipids and alpha-tocopherol in all lipoprotein density fractions isolated from advanced human atherosclerotic plaques. Arterioscler Thromb Vasc Biol 1999; 19(7): 1708-18.
226. G. Ravaglia, P. Forti, A. Lucicesare, et al. Plasma tocopherols and risk of cognitive impairment in an elderly Italian cohort. Am J Clin Nutr 2008; 87(5): 1306-13.
227. P. Torquato, D. Bartolini, D. Giusepponi, 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: 115-25.
228. S. A. Pope, P. T. Clayton, D. P. Muller. A new method for the analysis of urinary vitamin E metabolites and the tentative identification of a novel group of compounds. Arch Biochem Biophys 2000; 381(1): 8-15.
229. S. Gayatri, D. Muller, S. O'Riordan, et al. A novel method for the direct measurement of urinary conjugated metabolites of α-tocopherol and its use in diabetes. Mol Nutr Food Res 2010; 54(5): 599-600.
230. A. A. El-Aal, E. A. A. El-Ghffar, A. A. Ghali, M. R. Zughbur, M. M. Sirdah. The effect of vitamin C and/or E supplementations on type 2 diabetic adult males under metformin treatment: A single-blinded randomized controlled clinical trial. Diabetes Metab Syndr 2018; 12(4): 483-9.
231. K. T. Hall, J. E. Buring, K. J. Mukamal, et al. COMT and Alpha-Tocopherol Effects in Cancer Prevention: Gene-Supplement Interactions in Two Randomized Clinical Trials. J Natl Cancer Inst 2019.

CHAPTER 3

Low mitochondrial DNA copy number drives atherogenic cardiovascular disease: cohort and genetic studies

Jiao Luo, Raymond Noordam, J Wouter Jukema, Ko Willems van Dijk, Sara Hägg, Felix Grassmann, Saskia le Cessie, Diana van Heemst

In revision

Abstract

Aim: Mitochondrial DNA dysfunction has been implicated in the pathogenesis of cardiovascular disease. We aimed to investigate the associations of mitochondrial DNA copy number (mtDNA-CN), as a proxy of mitochondrial function, and coronary artery disease (CAD) and heart failure (HF) in a cohort study and test the causal nature of any potential associations using Mendelian randomization (MR).

Methods and Results: Multivariable-adjusted regression analyses were conducted in 273,619 unrelated participants of European ancestry from UK Biobank (UKB). In two-sample MR analyses, single nucleotide polymorphisms (SNPs) associated with mtDNA-CN were retrieved from published genome-wide association studies. SNP-disease associations were obtained for CAD from CARDIoGRAMplusC4D, UKB, and FinnGen, comprising 902,538 participants (134,759 cases), and for HF from the HERMES consortium and FinnGen, comprising 1,195,531 participants (70,706 cases). MR analyses were performed per database and results were subsequently meta-analyzed using fixed-effects models. During a median follow-up of 11.8 years, cox regression restricted cubic spline analyses showed associations between lower mtDNA-CN and higher risk of CAD and HF. Hazard ratios for participants in the lowest quintile of mtDNA-CN compared with those in the highest quintile were 1.08 (95% confidence interval: 1.03, 1.14) and 1.15 (1.05, 1.24) for CAD and HF. Genetically, the pooled odds ratios from two-sample MR of genetically predicted per one-SD decrease in mtDNA-CN were 1.16 (1.05, 1.27) for CAD and 1.00 (0.90, 1.10) for HF, respectively.

Conclusion: Our findings support a possible causal role of lower mtDNA-CN in higher CAD risk, but not in HF.

Introduction

Cardiovascular disease (CVD) is the leading cause of death worldwide. The heart is high oxygen-consuming, with large amounts of mitochondria constituting up to one-quarter of cardiomyocytes volume[1], Mitochondrial dysfunction, a hallmark of the aging process[2], leads to reduced bioenergetic capacity and disrupted redox homeostasis, and is therefore hypothesized to be a critical component in the pathogenesis of CVD[3,4]. Mitochondria have their own circular genome, the mitochondrial DNA (mtDNA), consisting of 37 genes, 13 of which encode proteins on the electron transport chain. Individual mitochondrion may contain several copies of the mitochondrial genome, known as mtDNA copy numbers (mtDNA-CN). The mtDNA-CN are associated with bioenergetics, mitochondrial membrane potential, and oxidative stress[5], and therefore could serve as a surrogate biomarker of mitochondrial dysfunction[6]. A better understanding on the role of mtDNA-CN may provide early opportunities in the prevention and treatment of CVD.

Recent epidemiological studies have assessed the associations between peripheral blood mtDNA-CN and multiple cardiovascular endpoints. These studies so far unequivocally indicated lower mtDNA-CN as an independent risk factor of prevalent CVD in case-control and retrospective cohort studies[7-10] and of incident cardiovascular disease and risk of sudden cardiac death in the prospective Atherosclerosis Risk in Communities (ARIC) study[11-14]. Nevertheless, apart from the ARIC study, other studies comprised a small sample size and/or a limited number of cases, which might have resulted in insufficient statistical power. In a recent cross-sectional study integrating multiple studies, mtDNA-CN was associated with a cluster of cardiometabolic traits that increase the risk of CVD, including obesity, hypertension, and hyperlipidemia[15]. However, due to the vague onset and long-term progression of CVD pathogenesis, it is not possible to fully eliminate reverse causation and residual confounding in studies with observational study designs. Whether these associations are of a causal nature, therefore, remains unclear.

Triangulation of causal inference in etiological epidemiology has been proposed, which integrates results from different methodological approaches to enhance the reliability of a research study[16]. The confidence of the findings will be strengthened if results from different approaches are consistent with each other. Based on earlier studies, we hypothesized that a lower mtDNA-CN is associated with an increased risk of incident CVD. Consequently, we first examined the associations between mtDNA-CN and incidence of coronary artery disease (CAD) and heart failure (HF) in participants of European ancestry in the UK Biobank (UKB) using cox-proportional hazard regression models. Second, we exploited publicly available data to perform a two-sample Mendelian randomization (MR) to investigate whether genetically predicted low mtDNA-CN were causally associated with increased risk of diseases.

Methods

Prospective study

Study population

The UKB cohort is a prospective cohort with 502,628 participants between the age of 40 and 69 years recruited from the general population at multiple assessment centers across the UK between 2006 and 2010[17]. More detailed information about the recruitment of participants is available in **Supplementary methods**. The UKB study was approved by the North-West Multi-center Research Ethics Committee (MREC) and conducted according to the Declaration of Helsinki. All participants provided written informed consent. We used genotype data from 488,377 individuals of the full genetic data release (July 2017) in the present study.

Participants who did not pass the sample quality control were initially excluded according to the parameters presented in the sample quality control file of the UKB, including participants who were: 1) not used to compute principal components; 2) identified as outliers in heterozygosity and missing rates, which is indicative of poor-quality genotype data for these samples; 3) identified as putatively sex chromosome aneuploidy; 4) sex inferred from genotype data did not match their self-reported sex; 5) have an excessive number (more than 10) of relatives in the database. To minimize the variation resulting from population substructures, we restricted the study population to unrelated white British individuals. Participant relatedness was available in the UKB by estimating kinship coefficients for all pairs. White British ancestry was identified based on self-reported ethnic background, and further refined the population definition in a principal component analysis of the genotype data that were tightly clustered as performed and provided by the UKB. This resulted in a primary study cohort comprising 302,685 unrelated European ancestry participants. A flowchart on the different exclusions is provided in **Figure S1**.

mtDNA-CN computation

Somatic mtDNA-CN were assessed from the intensities of genotyping probes on the mitochondrial chromosome on the Affymetrix Array. The method for computing mtDNA-CN has been described in detail before[18]. We followed the same pipeline to calculate mtDNA-CN in the available data of UKB (*https://github.com/GrassmannLab/MT_UKB*). In brief, the relative amount of mtDNA hybridized to the array at each probe was 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. The median L2R values across all 265 variants passing quality control on the MT chromosome were used as an initial raw measure of mtDNA-CN. To correct for the confounding induced by poorly performing probes, we weighted the L2R values of each probe multiplied by the weight of the probe that was 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-CN estimate for each individual. To eliminate the plate effect, we subsequently

standardized the CN to a mean of zero and standard deviation (SD) of one within each genotyping plate comprising 96 wells. An additional quality control step was performed by eliminating individuals with high standard deviation (SD) (two SD from the mean) of autosomal probes log2 ratio (L2R). Consequently, 293,245 individuals remained in the cohort.

Outcome definition

Outcomes in the analysis were incident cardiovascular diseases during the time period from recruitment to January 1st, 2021. Incident disease status was ascertained by linkage with hospital admissions data and national death register data to identify the date of the first known CVD or CVD-related death after the date of baseline assessment. The linkage details are presented in the original study protocol (*https://www.ukbiobank.ac.uk/media/gnkeyh2q/study-rationale.pdf*, accessed April 2021). The outcomes were incident CAD and HF, separately. Incident disease diagnoses are coded according to the International Classification of Diseases edition 10 (ICD-10); CAD cases were defined as angina pectoris (I20), myocardial infarction (MI) (I21 and I22), and acute and chronic ischemic heart disease (IHD) (I24 and I25); Incident HF cases were defined as I50. In addition, we analyzed acute myocardial infarction (MI) and chronic ischemic heart disease (IHD) as separate outcomes in sensitivity analyses. Follow-up time is computed from the baseline visit to the diagnosis of incident disease, death, or the censoring, whichever occurred first.

Covariates

Covariates were from baseline measurements, which included demographic parameters (age at recruitment, sex, deprivation index); the first ten principal components (PCs) to correct for possible remaining population stratification; genotyping batch; cell numbers (white blood cell counts and platelet counts); anthropometric measure of body mass index (BMI) in kg/m^2; self-reported lifestyle factors (smoking status [never, past and current], alcohol consumption frequency [twice or less per week/ more than three times per week], physical activity [MET hours per week for moderate-vigorous activity], sleep duration in hours and insomnia symptoms [yes/no]); familial CVD history (yes/no), lipid levels (mmol/l) (total and LDL [low-density lipoprotein] cholesterol) lipid-lowering medication, blood pressure (mmHg, average of the two measurements taken a few moments apart when applicable) and blood pressure-lowering medication, as well as baseline type 2 diabetes mellitus (T2DM, yes/no) from the medical records.

Statistical analysis

After further exclusion of participants with any prevalent cardiovascular disease or withdrawn informed consent, the study cohort comprised an analytic sample of 273,619 individuals (**Figure S1**). Baseline characteristics of the study population were described in quintiles of mtDNA-CN and presented as mean (SD) or median (interquartile range, IQR) for continuous variables and frequency (proportion) for categoric variables. Cumulative incidence for competing risks (CICR) was used to plot the cumulative incidence of both CAD and HF against follow-up time by mtDNA-CN quintiles, accounting for death as a competing event.

Cox proportional hazards models were used to estimate hazard ratios (HR) and corresponding 95% confidence intervals (CI) for the association between mtDNA-CN and incident CAD and HF separately. Two multivariable-adjusted regression models were fitted: Model 1 was adjusted for age, sex, the first ten PCs, genotyping batch, white blood cell count, and platelet count; Model 2 was additionally adjusted for BMI, smoking, alcohol consumption, sleep duration, insomnia, physical activity, familial CVD history, lipid levels and lipid-lowering medication, blood pressure, and blood pressure-lowering medication and T2DM. In the primary analysis, we treated mtDNA-CN as a continuous variable and assessed the risk of incident diseases associated with per one-SD decrease in mtDNA-CN using cox restricted cubic spline curves, with knots located at 5th, 50th, and 95th percentiles. Subsequently, mtDNA-CN was categorized into quintiles, and hazard ratios (HR) compared the 1st to 4th quintiles with the 5th quintile (reference). The proportional hazard assumption was graphically assessed by plotting log(-log[survival]) versus log(follow-up time) and was tested using Schoenfeld residuals.

Missing data were present in the covariates, and were imputed using multiple imputation by chain equations (MICE)[19], setting the number of imputed datasets to 10. We used predictive mean matching for continuous variables, logistic regression for binary variables, and polytomous regression for categorical variables. The imputation model included mtDNA-CN, all covariates, the Nelson-Aalen estimator of cumulative hazard and incident disease status. Cox proportional hazards models were fitted within each imputed dataset and were subsequently pooled according to Rubin's rules.
As sensitivity analyses, firstly, interaction terms between mtDNA-CN and age and sex were added to Model 2 to test for the presence of effect modification by sex or age. Subgroup analyses were also performed in each stratum of sex and age (< 50 years, 50~60 years, > 60 years), respectively. Secondly, all analyses were performed for the CAD subtypes, i.e. MI and IHD. Thirdly, analyses were repeated restricting to participants without missing data on covariates, i.e. complete cases (N = 162,002)

Mendelian randomization analyses

Instrumental variables
We retrieved 129 independent (linkage disequilibrium < 0.05) nuclear single-nucleotide polymorphisms (SNPs) on autosomes as genetic instruments that were associated with continuous mtDNA-CN at a genome-wide significance threshold ($p < 5e-08$), as identified in a recent genome-wide association study (GWAS) by Longchamps et al20. The study was performed in 465,809 individuals of White European ancestry combining the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium and UKB. Genetic associations were adjusted for age, sex, and covariates that were specific in each cohort, such as PCs, blood collection sites, family structure, and cell composition. F statistics [$(\beta/se)^2$] were computed to evaluate instrumental strength, and SNPs with a value of less than 10 were considered weak instruments. Furthermore,

we calculated the proportion of total variance in the exposure explained by each instrument (R^2) separately[21].

Outcome data source

Summary statistics for instrument-CAD associations were extracted from 3 large databases separately, the CARDIoGRAMplusC4D (Coronary Artery Disease Genome-Wide Replication and Meta-analysis plus the Coronary Artery Disease Genetics from Nikpay et al.[22], where UKB data were not included) consortium, UKB, and FinnGen study (freeze 5, released in May 2021). Similarly, summary statistics for SNP-HF associations were drawn from HERMES Consortium (Heart Failure Molecular Epidemiology for Therapeutic Targets Consortium, which included data from UKB in the meta-analysis) and the FinnGen study, respectively. The descriptions, number of cases and controls, cases definition as well as covariates used for associations tests of each of the databases are presented in detail in **Supplementary methods** and **Table S1**.

Mendelian randomization analysis

SNP-exposure and SNP-outcome data were harmonized to make alignment on effect alleles. Palindromic SNPs were eliminated[23]. The primary MR analysis was performed using Inverse-variance weighted (IVW) method to combine the SNP-specific estimates calculated using Wald ratios, assuming all instrumental variables are valid[24]. Results were expressed as an odds ratio (OR) on disease risk for a one-SD decrease in genetically predicted mtDNA-CN. When the MR assumptions were met, this OR approximated the causal effect of the exposure on the outcome. Sensitivity analyses accounting for pleiotropy were conducted, including Weighted-Median Estimator (WME) and MR-Egger regression[25,26], both of which assume that at least half of the instrumental variable had to be valid. The intercept from MR-Egger represents the average pleiotropic effect; when the intercept deviates from zero, estimates from IVW might be biased. MR-PRESSO (MR Pleiotropy RESidual Sum and Outlier) was applied to detect and correct for horizontal pleiotropy through removing outliers[27]. Moreover, we examined the heterogeneity using Cochran's Q statistic among all SNPs within each outcome database.

For outcomes derived from the UKB, despite the gene-exposure associations being from the same population, it has been shown that Two-sample MR methods can be reliably used for one-sample MR performed within large biobanks, such as UKB, with the exception of the MR-Egger sensitivity analysis[28].

Meta-analysis of estimates from different databases

The effects of mtDNA-CN on CAD/HF in MR analyses were separately estimated in different outcome databases separately, CARDIoGRAMplusC4D consortium (CAD) or HERMES (HF), UKB (CAD only), and FinnGen (both), and derived estimates were subsequently pooled using fixed-effects meta-analysis.

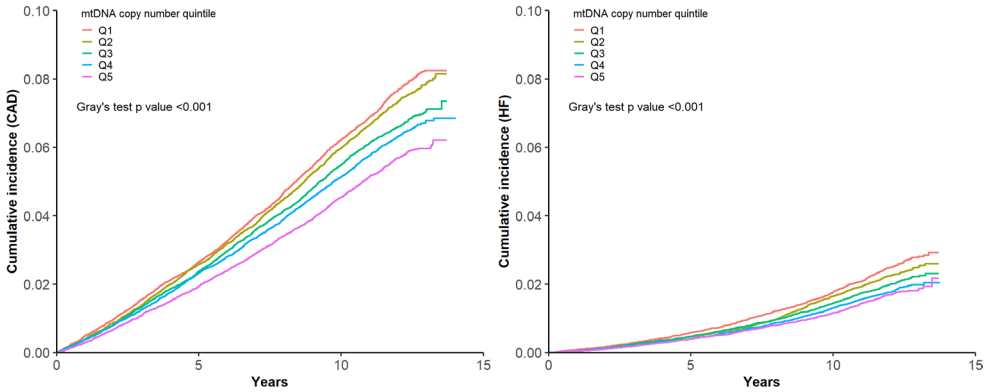


Figure 1 Cumulative Incidence of CAD and HF by quintiles of mtDNA copy number

We calculated Cumulative incidence for CAD and HF, accounting for death as a competing event. Differences in cumulative incidence between mtDNA copy number quintiles were assessed using Gray's test. CAD: coronary artery disease; HF: heart failure.

Sensitivity analysis

Despite the large sample size of the GWAS used for the selection of instrumental variables in the Longchamps et al study, which increased the statistical power, the assessments of mtDNA-CN among cohorts that contributed data to the meta-analysis were very different. To account for this measurement heterogeneity, we additionally performed sensitivity analyses restricting to genetic instruments identified in the UKB only. Therefore, 66 independent (linkage disequilibrium < 0.1) SNPs were used that were associated with mtDNA-CN at a genome-wide significance threshold ($p < 5e\text{-}08$) from 295,150 participants conducted by Hägg et al.18. Genetic associations were adjusted for PCs, age, sex, genotyping batch, genotyping missingness/call rate, and cell composition. All MR analyses were repeated with the substitution of the 69 genetic instruments for mtDNA-CN.

All the analyses were performed using R (v3.6.3) statistical software (The R Foundation for Statistical Computing, Vienna, Austria). Packages used in the analyses included "cmprsk" for cumulative incidence for competing risk analyses, "mice" for multiple imputations, "survival" and "survminer" for cox proportional hazard regression, "rms" for nonlinear dose-response associations, "TwoSampleMR" for MR analyses, and "meta" for meta-analyses. All results were reported as odds ratio with accompanied 95% confidence interval.

Results from the multivariable-adjusted prospective analyses

Prospective results

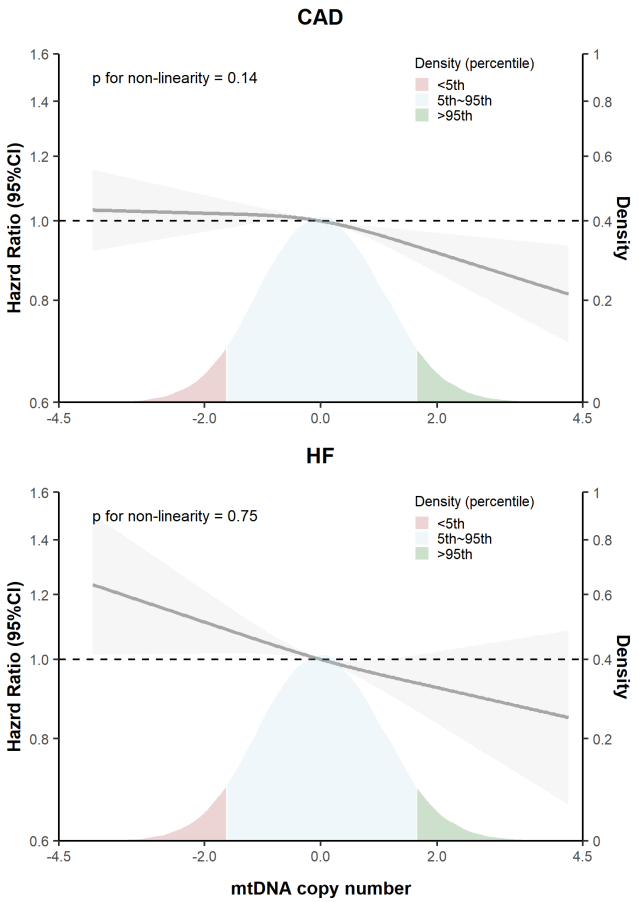


Figure 2 Hazard ratios for incident CAD and HF by levels of mtDNA copy number.

Solid lines represented hazard ratios (derived from model 2 adjusted for age, sex, genotyping batch, the first two principal components, white blood cell count and platelet count, body mass index, physical activity, smoking status, alcohol consumption frequency, blood pressure and blood pressure-lowering medication, cholesterol, triglycerides and lipid-lowering medication, sleep duration and insomnia, type 2 diabetes status, and familial history of cardiovascular disease) and corresponding 95% confidence intervals (gray shadowed area) using restricted cubic splines for mtDNA copy number with knots at distribution of 5th, 50th, and 95th percentiles. The density on the right y-axis shows distribution of baseline mtDNA copy number. Since mtDNA copy number has been standardized during coputation, the distribution is approximately normal. Since mtDNA copy number was scaled before analyses, the distribution is normal. CAD: coronary artery disease; HF: heart failure.

Main analyses

A total of 273,619 participants were eligible for analyses after exclusion (**Figure S1**). Compared with the highest quintile of mtDNA-CN (**Table 1**), participants in the lower quintiles were more likely to have more unfavorable CVD risk factors, including older age, male sex, higher BMI, higher blood pressure and more

Twice or less per week	29,751 (54.4%)	29,851 (54.5%)	29,381 (53.7%)	29,287 (53.5%)	29,015 (53.0%)	
Data missing	34 (0.1%)	37 (0.1%)	39 (0.1%)	37 (0.1%)	30 (0.1%)	
Smoking status						<0.001
Current	6124 (11.2%)	5643 (10.3%)	5369 (9.8)	5046 (9.2%)	4521 (8.3%)	
Previous	18,646 (34.1%)	18,831 (34.4%)	18,588 (34.0%)	18,586 (34.0%)	18702 (34.2%)	
Never	29,732 (54.3%)	30,079 (55.0%)	30,594 (55.9%)	30,914 (56.5%)	31350 (57.3%)	
Data missing	222 (0.4%)	171 (0.3%)	172 (0.3%)	178 (0.3%)	151 (0.3%)	
Sleep duration (hours)	7.1 (1.2)	7.1 (1.2)	7.1 (1.2)	7.1 (1.2)	7.1 (1.2)	0.6
Insomnia						0.02
Sometimes	26,161 (47.8%)	26,273 (48.0%)	26,420 (48.3%)	26,243 (48.0%)	26,321 (48.1%)	
Usually	15,488 (28.3%)	15,145 (27.7%)	15,031 (27.5%)	14,982 (27.4%)	14,991 (27.4%)	
Never/rarely	13,038 (23.8%)	13,271 (24.3%)	13,237 (24.2%)	13,454 (24.6%)	13,384 (24.5%)	
Data missing	37 (0.1%)	35 (0.1%)	35 (0.1%)	45 (0.1%)	28 (0.1%)	
Familial CVD history						<0.001
Yes	21,539 (39.4%)	21,716 (39.7%)	21,690 (39.6%)	21,687 (39.6%)	21,359 (39.0%)	
No	27,791 (50.8%)	27,822 (50.8%)	28,005 (51.2%)	28,026 (51.2%)	28,420 (51.9%)	
Data missing	5394 (9.8%)	5186 (9.5%)	5028 (9.2%)	5011 (9.2%)	4945 (9.0%)	
T2DM history (yes)						<0.001
Yes	1184 (2.2%)	1093 (2.0%)	969 (1.8%)	914 (1.7%)	879 (1.6%)	
No	53,540 (97.8%)	53,631 (98.0%)	53,755 (98.2%)	53,810 (98.3%)	53,845	

Data are mean (SD) or median (interquartile range, IQR) for continuous variables and frequency (percentage) for categorical variables. Some percentages do not add up to 100 because of rounding. P values were obtained from *Kruskal-Wallis H* test or chi-square test as appropriate. BMI: Body mass index; HDL: high-density lipoprotein cholesterol; LDL: low-density lipoprotein cholesterol; CVD: cardiovascular disease; CAD: coronary artery disease; HF: heart failure.

Table 1 Baseline charactertistics of the study participants by quintiles of mtDNA copy number

Variable	Q1	Q2	Q3	Q4	Q5	p value
N	54,724	54,724	54,723	54,724	54,724	-
Age (years)	57.2 (8.0)	56.8 (8.0)	56.5 (8.0)	56.2 (8.0)	55.9 (7.9)	<0.001
Sex						<0.001
Male	25,413 (46.4%)	24,898 (45.5%)	24,478 (44.7%)	24,320 (44.4%)	23,622 (43.2%)	
Female	29,311 (53.6%)	29,826 (54.5%)	30,245 (55.3%)	30,404 (55.6%)	31,102 (56.8%)	
BMI (kg/m^2)	27.6 (5.0)	27.4 (4.8)	27.3 (4.6)	27.1 (4.6)	26.9 (4.5)	<0.001
Deprivation index	-1.5 (2.9)	-1.6 (2.9)	-1.7 (2.9)	-1.7 (2.9)	-1.7 (2.9)	<0.001
Diastolic blood pressure (mmHg)	82.9 (10.2)	82.6 (10.0)	82.4 (10.0)	82.3 (10.1)	81.8 (10.1)	<0.001
Systolic blood pressure (mmHg)	139.5 (18.9)	138.7 (18.7)	138.2 (18.7)	137.6 (18.4)	136.6 (18.3)	<0.001
Blood pressure-lowering medication						<0.001
Yes	10,826 (19.8%)	10,085 (18.4%)	9598 (17.5%)	9089 (16.6%)	8524 (15.6%)	
No	43,898 (80.2%)	44,637 (81.6%)	45,128 (82.5%)	45,635 (83.4%)	46,199 (84.4%)	
Total cholesterol (mmol/L)	5.8 (1.1)	5.8 (1.1)	5.8 (1.1)	5.8 (1.1)	5.7 (1.1)	<0.001
HDL (mmol/L)	1.5 (0.4)	1.5 (0.4)	1.5 (0.4)	1.5 (0.4)	1.5 (0.4)	<0.001
LDL (mmol/L)	3.7 (0.9)	3.6 (0.8)	3.6 (0.9)	3.6 (0.8)	3.6 (0.8)	<0.001
Triglycerides (mmol/L)	1.8 (1.0)	1.8 (1.0)	1.7 (1.0)	1.7 (1.0)	1.7 (1.0)	<0.001
Cholesterol lowering medication						<0.001
Yes	7596 (13.9%)	7394 (13.5%)	7148 (13.1%)	6822 (12.5%)	6663 (12.2%)	
No	47128 (86.1%)	47330 (86.5%)	47575 (86.9%)	47902 (87.5%)	48061 (87.8%)	
Physical activity (moderate-vigorous MET hours/week)	26.5 (33.8)	27.0 (34.3)	27.1 (34.6)	27.1 (33.8)	27.1 (33.8)	0.054
Alcohol consumption frequency						<0.001
At least three times per week	24,939 (45.6%)	24,836 (45.4%)	25,303 (46.2%)	25,400 (46.4%)	25,679 (46.9%)	

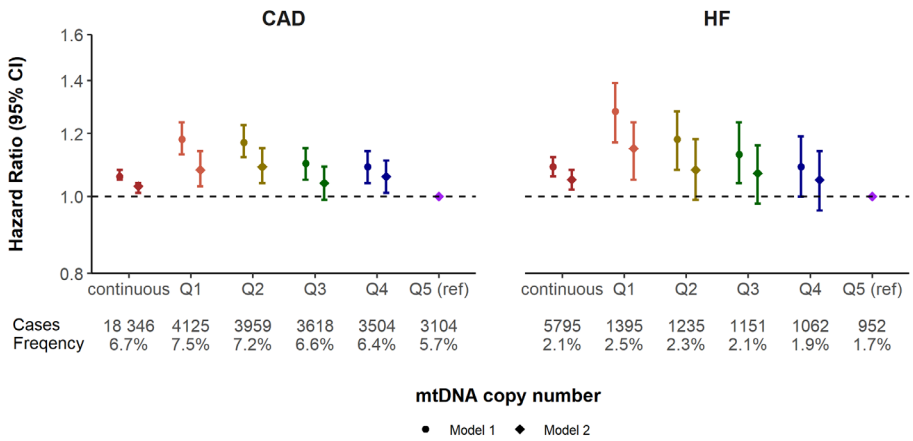


Figure 3 Hazard ratios for incident CAD and HF by quintiles of mtDNA copy number

Estimated hazard ratios for the effect of per-SD decrease in mtDNA copy number (continuous), or for the 1st to the 4th quintile compared to the 5th (reference) quintile (categorical) on CAD and HF. Model 1 was adjusted for age, sex, genotyping batch, the first two principal components, white blood cell count and platelet count. Model 2 was model 1 additionally adjusted for body mass index, physical activity, smoking status, alcohol consumption frequency, blood pressure and blood pressure-lowering medication, cholesterol, triglycerides and lipid-lowering medication, sleep duration and insomnia, type 2 diabetes status, and familial history of cardiovascular disease. CAD: coronary artery disease; HF: heart failure.

blood pressure-lowering medication, higher lipids (total cholesterol and LDL) and more cholesterol-lowering medication, less physical activity, more current smokers and a higher percentage of familial history of CVD or prevalent T2DM.

During a median follow-up of 11.8 (interquartile range: 11.0, 12.5) years, 18,346 participants developed CAD and 5795 participants developed HF. Cumulative incidence of both CAD and HF increased stepwise with the decrease in mtDNA-CN, accounting for death as a competing risk (p for Gray's test < 0.001) (**Figure 1**). In multivariable-adjusted cox proportional hazard models, restricted cubic spline analyses showed an approximately linear dose-response relationship between lower mtDNA-CN with the higher risk of CAD (p for non-linearity = 0.14) and HF (p for non-linearity = 0.73), as shown in **Figure 2**. Categorically (model 1), a one-SD decrease in mtDNA-CN was associated with 1.06-fold (95% confidence interval, CI: 1.05, 1.08) and 1.09-fold (95%CI: 1.06, 1.12) higher hazard of CAD and HF, respectively; adjusted HRs for the first versus the fifth (reference) quintile of mtDNA-CN were 1.18 (95%CI: 1.23, 1.24) for CAD and 1.28 (95%CI: 1.17, 1.39) for HF. Additional adjustment for CVD risk factors only minimally attenuated the estimates of CAD and HF (**Figure 3** and **Table S2**).

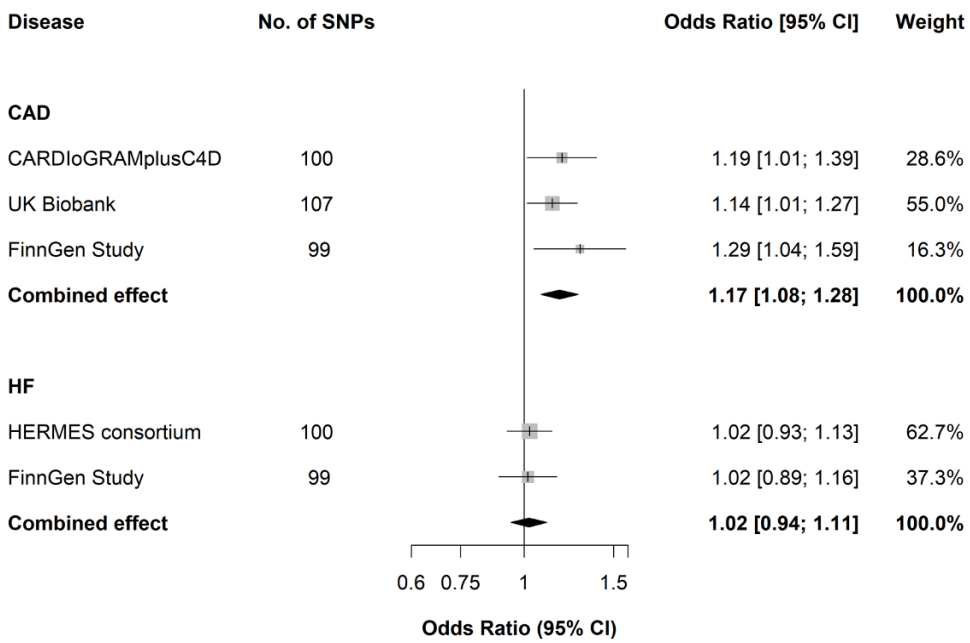


Figure 4 Mendelian randomization study of mtDNA copy number on the risk of CAD and HF

Estimated ORs for the effect of per-SD decrease in mtDNA copy number on CAD and HF, obtained from an MR inverse-variance weighted method, per outcome database separately and combined over the different databases using fixed-effect meta-analyses. CARDIoGRAMplusC4D: Coronary Artery Disease Genome-Wide Replication and Meta-analysis plus the Coronary Artery Disease Genetics; HERMES consortium: Heart Failure Molecular Epidemiology for Therapeutic Targets consortium. UK Biobank data of heart failure was already integrated into HERMES consortium.

Sensitivity analysis

We observed no evidence favoring an interaction between mtDNA-CN and sex (p for interaction = 0.2 for CAD, 0.7 for HF); in line, in sex-stratified analyses, the estimates between men and women were similar (**Table S3**). The interaction was observed between mtDNA-CN and age at baseline for CAD (p for interaction < 0.001). After stratification by age groups, HRs obtained from model 2 for CAD slightly attenuated from the younger group (<50 years) to older groups (50~60 years and >60 years) (HR: 1.06, 1.04 and 1.02, respectively) (**Table S4**). However, no interaction was detected between mtDNA-CN and age at baseline for HF (p for interaction = 0.2); though HR in the younger group was also higher for HF, this may be due to the very limited number of cases in this group

When analyses were conducted for MI and IHD separately, cumulative incidences were higher in lower quintiles compared with the highest quintile (**Figure S2**) for MI and IHD; estimates from cox proportional hazard regression models did not differ considerably from when all CADs were considered (**Table S2-S4**).

In addition, missing data in covariates were present (**Table S5**), and 162,002 (59%) of 273,619 individuals included in the current study provided complete data for all variables. The absolute difference in the baseline characteristics between these participants with and without complete data was very limited (**Table S6**). Furthermore, the main results from sensitivity analyses restricting to complete cases did not differ materially from the results obtained after imputation (**Table S7**).

Mendelian randomization

Main analyses

For the included 129 genetic instrumental variables, 4 of which with an F-statistics below 10 were discarded to avoid weak instrumental bias. In total, 108 distinct SNPs were present in at least one of the outcome databases. F-statistics for each SNP were higher than 10 and ranged from 15.6 to 634.4, and a total of 2.0% variation were explained by the instruments (**Table S8**).

For CAD, the pooled OR of the primary IVW estimates from CARDIoGRAMplusC4D, UKB, and FinnGen of a one-SD decrease in mtDNA-CN was 1.17 (95%CI: 1.08, 1.28) (**Figure 4**). Estimates from WME and MR-Egger generally did not differ substantially with the exception of UKB where the point estimates attenuated to some extent. No pleiotropy was detected by the intercept of MR-Egger ($p > 0.05$). Though outliers were identified by MR PRESSO in each database, estimates after outlier removal remained similar to those obtained from IVW (**Table S9**).

For HF, the combined OR from IVW obtained in the HERMES consortium and FinnGen of per one-SD decrease in mtDNA-CN was 1.02 (95%CI: 0.94, 1.11) (**Figure 4**). Results from WME were similar, and we observed no evidence for horizontal pleiotropy from MR-Egger intercept ($p > 0.05$); outliers were spotted in the HERMES consortium assessed by MR PRESSO, but outlier-corrected estimates did not materially differ with those generated from IVW (**Table S10**).

Sensitivity analyses

When we used genetic instrumental variables from Hägg et al.[18], 64 distinct SNPs were included. F statistics for each SNPs were higher than 10 and ranged from 29.8 to 277.4, and a total of 1.2% variation were explained by the. Detailed full information of the used genetic instruments is presented in **Table S11**. A one-SD decrease in mtDNA-CN was associated with 1.16-fold (95%CI:1.05, 1.27), 1.00-fold (95%CI:0.90, 1.10) higher risk of CAD and HF in meta-analysis, respectively (**Figure S3**). MR sensitivity analyses including WME, MR-Egger, and MR PRESSO are presented in **Table S12-S13**.

Discussion

In the present study, we implemented a prospective cohort study design and MR study to assess the relationship of mtDNA-CN with the risk of incident CAD and HF. Results from the multivariable-adjusted prospective analyses suggested an association between lower mtDNA-CN with a higher risk of CAD and HF, whereas

findings from MR analyses only confirmed an association between genetically predicted lower mtDNA-CN with a higher risk of CAD, possibly reflecting evidence of causality for CAD.

Consistent with our observational findings, previous studies showed that lower mtDNA-CN measured from peripheral blood was related to an increased risk of CVD and its risk factors [7-15]. The only prospective study that assessed the relationship between mtDNA content and either CAD or HF used the Atherosclerosis Risk in Communities (ARIC) Study[12,13]. In the ARIC study, composed of 20,163 participants (2460 incident CHD) with a mean follow-up of 13.5 years, a lower mtDNA-CN was associated with an increased risk of incident CHD. Similarly, with 10,802 participants (2227 incident HF cases) followed-up for a mean of 23.1 years, lower mtDNA-CN was linked to an increased risk of HF. Residual confounding, in particular factors relevant to both mitochondrial function and CVD such as physical activity and insomnia, was not taken into account. However, in our multivariable-adjusted analysis, additional adjustment for these covariates did not further attenuate the estimates substantially.

To the best of our knowledge, the current study is the first to evaluate the causal nature of the association between mtDNA-CN and risk of CVD. MR analyses with the genetic instruments for mtDNA-CN confirmed the detrimental effect of lower mtDNA content on the risk of CAD observed in the cohort studies. Mitochondrial dysfunction, indicating by low mtDNA-CN, would lead to increased production of reactive oxygen species (ROS) in mitochondria[5]. Those maladaptive overproduced mitochondrial ROS mediate irreversible damage to macromolecules, such as increased oxidation of low-density lipoprotein and dysfunction of endothelial cells that are critical factors to promote atherosclerosis, and further CAD events[29]. Nevertheless, several factors merit thoughtful consideration in terms of the interpretation of the null effect on HF in MR analyses. HF has substantial phenotypic heterogeneity, which can be defined by ejection fraction (EF) and diastolic function; more than half of patients have preserved EF while over 40% of cases have isolated diastolic dysfunction[30]. Moreover, a large degree of variation has been described even within patients with preserved EF[31,32]. It has also been shown previously that the associations between mtDNA-CN and HF with preserved and reduced EF were different and possibly would make the association into the direction of zero when we combined the two subgroups in a single analysis[13]. However, stratification by cause of HF in the UKB ended up with a low number of cases and insufficient statistical power. In addition, cause-specific GWAS summary-level data of HF are currently not available. For these reasons, the lack of a clear association between mtDNA-CN and HF should be interpreted with caution, and more follow-up analyses are required to investigate the cause-specific HF in more detail.

Study strengths and limitations

The main strength of our study is that we adopted the triangulation of causal inference in etiological epidemiology[16]. The consistency between biochemically measured and genetically determined mtDNA-CN in relation to CAD increased the credibility of the results. Given the absence of randomized clinical trials with respect to mtDNA-CN and CAD to date, the analyses that have been performed

in the present study provide the foremost evidence on the association between mtDNA content and CAD. Other important strengths of our prospective cohort study include the large sample size and the considerable number of incident cases from UKB, comprehensive assessment of confounding factors, and subtype analyses of MI and IHD within CAD. In MR studies, we meta-analyzed three large databases where SNP-outcome associations were derived, comprising a substantial size of overall participants and cases. The results are consistent across different databases, and the precision of the pooled MR estimates obtained from different databases increased significantly.

Several limitations should be acknowledged. First, mtDNA content was measured in peripheral blood cells, which may be different from cells in the vasculature or in the heart. Nevertheless, it has been shown that the blood cell and cardiomyocyte mtDNA-CN were significantly correlated within individuals, with a Pearson correlation coefficient of 0.72[33]. In addition, the initial calculation of mtDNA-CN from chip arrays might have introduced noise due to the small number of variants. To this end, a weighted mtDNA-CN was implemented, which approximates what would be estimated from exome sequencing and has been validated[18]. Second, despite the large number of instrumental variables in the MR analyses, the variation of mtDNA-CN explained by these SNPs was small. Notwithstanding, we had more than sufficient power to detect the true causal effect in MR analyses (**Figure S4**). Moreover, while we acknowledge the possibility of pleiotropic effects of included genetic instruments, this is likely to be vertical (**Supplementary discussion**). When we additionally stringently excluded the 22 SNPs on chromosome 19, where APOE and LDLR gene locate, the estimates were slightly attenuated in each database but remained statistically significant in the meta-analysis (**Figure S5**). Third, since the population of non-Europeans was highly heterogeneous in UKB, we restricted the prospective analyses to White European populations; furthermore, MR analyses were also performed predominantly in European-descent individuals, except for 23% of individuals with a non-European background in CARDIoGRAMplusC4D. It is therefore inappropriate to extrapolate our findings to other populations with different ethnic backgrounds. Lastly, we were not able to dissect the potential impact of other mtDNA alterations, such as mtDNA mutations or deletions which have been proposed to contribute to the initiation and progression of atherosclerosis[34]. Consequently, there is a need for accurate deep sequencing to simultaneously analyze the entire mitochondrial genome in order to better understand the relationships between mtDNA-CN function, germline and acquired mutations, and CVD.

Conclusion

This study provides the first evidence of a possible causal association between mtDNA-CN and the risk of CAD. Further studies are required to fully understand how mtDNA affect atherogenic risk development.

References

1. J. Schaper, E. Meiser, G. Stämmler. Ultrastructural morphometric analysis of myocardium from dogs, rats, hamsters, mice, and from human hearts. Circ Res 1985; 56(3): 377-91.
2. C. López-Otín, M. A. Blasco, L. Partridge, M. Serrano, G. Kroemer. The hallmarks of aging. Cell 2013; 153(6): 1194-217.
3. C. A. Castellani, R. J. Longchamps, J. Sun, E. Guallar, D. E. Arking. Thinking outside the nucleus: Mitochondrial DNA copy number in health and disease. Mitochondrion 2020; 53: 214-23.
4. D. F. Dai, P. S. Rabinovitch, Z. Ungvari. Mitochondria and cardiovascular aging. Circ Res 2012; 110(8): 1109-24.
5. D. C. Wallace. A mitochondrial bioenergetic etiology of disease. J Clin Invest 2013; 123(4): 1405-12.
6. A. N. Malik, A. Czajka. Is mitochondrial DNA content a potential biomarker of mitochondrial dysfunction? Mitochondrion 2013; 13(5): 481-92.
7. R. Wei, Y. Ni, P. Bazeley, et al. Mitochondrial DNA Content Is Linked to Cardiovascular Disease Patient Phenotypes. J Am Heart Assoc 2021; 10(4): e018776.
8. S. Chen, X. Xie, Y. Wang, et al. Association between leukocyte mitochondrial DNA content and risk of coronary heart disease: a case-control study. Atherosclerosis 2014; 237(1): 220-6.
9. X. B. Wang, N. H. Cui, S. Zhang, Z. J. Liu, J. F. Ma, L. Ming. Leukocyte telomere length, mitochondrial DNA copy number, and coronary artery disease risk and severity: A two-stage case-control study of 3064 Chinese subjects. Atherosclerosis 2019; 284: 165-72.
10. A. Koller, F. Fazzini, C. Lamina, et al. Mitochondrial DNA copy number is associated with all-cause mortality and cardiovascular events in patients with peripheral arterial disease. J Intern Med 2020; 287(5): 569-79.
11. Y. Zhang, E. Guallar, F. N. Ashar, et al. Association between mitochondrial DNA copy number and sudden cardiac death: findings from the Atherosclerosis Risk in Communities study (ARIC). Eur Heart J 2017; 38(46): 3443-8.
12. F. N. Ashar, Y. Zhang, R. J. Longchamps, et al. Association of Mitochondrial DNA Copy Number With Cardiovascular Disease. JAMA Cardiol 2017; 2(11): 1247-55.
13. Y. S. Hong, R. J. Longchamps, D. Zhao, et al. Mitochondrial DNA Copy Number and Incident Heart Failure: The Atherosclerosis Risk in Communities (ARIC) Study. Circulation 2020; 141(22): 1823-5.
14. D. Zhao, T. M. Bartz, N. Sotoodehnia, et al. Mitochondrial DNA copy number and incident atrial fibrillation. BMC Med 2020; 18(1): 246.
15. Xue Liu, Ryan J. Longchamps, Kerri L. Wiggins, et al. Association of mitochondrial DNA copy number with cardiometabolic diseases. Cell Genomics 2021; 1(1): 100006.
16. D. A. Lawlor, K. Tilling, G. Davey Smith. Triangulation in aetiological epidemiology. Int J Epidemiol 2016; 45(6): 1866-86.
17. C. Sudlow, J. Gallacher, N. Allen, 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(3): e1001779.

18. S. Hägg, J. Jylhävä, Y. Wang, K. Czene, F. Grassmann. Deciphering the genetic and epidemiological landscape of mitochondrial DNA abundance. Hum Genet 2020.
19. I. R. White, P. Royston, A. M. Wood. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med 2011; 30(4): 377-99.
20. R. J. Longchamps, S. Y. Yang, C. A. Castellani, 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(1): 127-46.
21. T. M. Teslovich, K. Musunuru, A. V. Smith, et al. Biological, clinical and population relevance of 95 loci for blood lipids. Nature 2010; 466(7307): 707-13.
22. M. Nikpay, A. Goel, H. H. Won, et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. Nat Genet 2015; 47(10): 1121-30.
23. F. P. Hartwig, N. M. Davies, G. Hemani, G. Davey Smith. Two-sample Mendelian randomization: avoiding the downsides of a powerful, widely applicable but potentially fallible technique. Int J Epidemiol 2016; 45(6): 1717-26.
24. S. Burgess, A. Butterworth, S. G. Thompson. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol 2013; 37(7): 658-65.
25. J. Bowden, G. Davey Smith, P. C. Haycock, S. Burgess. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. Genet Epidemiol 2016; 40(4): 304-14.
26. J. Bowden, G. Davey Smith, S. Burgess. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol 2015; 44(2): 512-25.
27. M. Verbanck, C. Y. Chen, B. Neale, R. Do. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. Nat Genet 2018; 50(5): 693-8.
28. C. Minelli, M. F. Del Greco, D. A. van der Plaat, J. Bowden, N. A. Sheehan, J. Thompson. The use of two-sample methods for Mendelian randomization analyses on single large datasets. Int J Epidemiol 2021.
29. N. R. Madamanchi, M. S. Runge. Mitochondrial dysfunction in atherosclerosis. Circ Res 2007; 100(4): 460-73.
30. F. Bursi, S. A. Weston, M. M. Redfield, et al. Systolic and diastolic heart failure in the community. Jama 2006; 296(18): 2209-16.
31. T. Ahmad, M. J. Pencina, P. J. Schulte, et al. Clinical Implications of Chronic Heart Failure Phenotypes Defined by Cluster Analysis. Journal of the American College of Cardiology 2014; 64(17): 1765-74.
32. G. S. Francis, R. Cogswell, T. Thenappan. The heterogeneity of heart failure: will enhanced phenotyping be necessary for future clinical trial success? J Am Coll Cardiol 2014; 64(17): 1775-6.
33. J. Huang, L. Tan, R. Shen, L. Zhang, H. Zuo, D. W. Wang. Decreased Peripheral Mitochondrial DNA Copy Number is Associated with the Risk of Heart Failure and Long-term Outcomes. Medicine (Baltimore) 2016; 95(15): e3323.
34. A. Volobueva, A. Grechko, S. F. Yet, I. Sobenin, A. Orekhov. Changes in Mitochondrial Genome Associated with Predisposition to Atherosclerosis and Related Disease. Biomolecules 2019; 9(8).

Supplementary materials

Table S1 GWAS data sources for Mendelian Randomization

Table S2 Hazard ratio of incident CAD and HF by quintiles of mtDNA-CN

Table S3 Hazard ratio of incident CAD and HF by levels of mtDN-CN, stratified by sex

Table S4 Hazard ratio of incident CAD and HF by levels of mtDN-CN, stratified by age

Table S5 Number and frequencies of missingness in each variable (N = 273,619)

Table S6 Differences between participants with complete data and participants with missing data

Table S7 Cox proportional hazard regression in complete cases

Table S8 Genetic instruments at genome-wide significant level for mtDNA-CN in the main MR analyses (Instruments retrieved from Longchamps et al.)

Table S9 Mendelian Randomization results of mtDNA-CN on the risk of CAD (Instruments retrieved from Longchamps et al.)

Table S10 Mendelian Randomization results mtDNA-CN on the risk of HF (Instruments retrieved retrieved from Longchamps et al.)

Table S11 Genetic instruments at genome-wide significant level for mtDNA-CN in the sensitivity MR analysis (Instruments retrieved from Hägg et al.)

Table S12 Mendelian Randomization results of mtDNA-CN on the risk of CAD (Instruments retrieved from Hägg et al.)

Table S13 Mendelian Randomization results mtDNA-CN on the risk of HF (Instruments retrieved from Hägg et al.)

Figure S1 Flowchart of participants inclusion in UK Biobank

Figure S2 Cumulative Incidence of MI and IHD by quintiles of mtDNA-CN

Figure S3 Mendelian Randomization study of mtDNA-CN on the risk of CAD and HF (Instruments retrieved from Hägg et al.)

Figure S4 Statistical power of Mendelian Randomization analyses

Figure S5 Mendelian Randomization study of mtDNA-CN on the risk of CAD and HF (Instruments retrieved from Longchamp et al. excluding 22 SNPs from chromosome 19)

Supplementary methods

Supplementary Discussion

The Supplementary materials for this article can be found online at: *https://drive.google.com/drive/folders/1d46G5jf6fIZUUIp6aHpjnz74SnL-jzQPd?usp=sharing*

CHAPTER 4

Mitochondrial DNA copy number and circulating metabolomic profiling: multivariable-adjusted and Mendelian randomization analyses

Jiao Luo, Saskia le Cessie, Ko Willems van Dijk, Sara Hägg, Felix Grassmann, Diana van Heemst, Raymond Noordam

Under submission

Abstract

Background: Low leukocyte mitochondrial DNA copy number (mtDNA-CN) has been associated with a high risk of atherosclerotic cardiovascular disease, but the mechanisms remain unclear. Changes in metabolomic profiles have frequently been observed during the transition from health to disease and are considered important intermediate phenotypes. Therefore, we aimed to investigate whether low mtDNA-CN is associated with worse metabolomic profiling using cross-sectional and genetic studies.
Methods: Among 61,186 unrelated European participants from the UK Biobank, we performed multivariable-adjusted linear regression analyses to examine the associations between mtDNA-CN and 168 NMR-based circulating metabolic biomarkers and 9 metabolic principal components (PCs) that collectively covered 91.5% of the total variation of individual metabolic biomarkers. Subsequently, we conducted Mendelian randomization (MR) to estimate the causal effects of mtDNA-CN on individual biomarkers and the metabolic PCs. Genetic variants associated with mtDNA-CN, used as instrumental variables, were derived from a publicly available genome-wide association study of 465,809 individuals of White European ancestry, and genetic-biomarker associations were derived from the UK Biobank.
Results: After correction for multiple testing, low mtDNA-CN was associated with 130 metabolic biomarkers, predominantly lower concentrations of some amino acids and higher concentrations of lipids, lipoproteins and fatty acids; moreover, mtDNA-CN was associated with 7 out of the 9 metabolic PCs. Using MR, genetically-influenced low mtDNA-CN was associated with lower lactate, higher acetate, and a higher degree of unsaturation, and corresponding standardized betas (95% CIs) from the inverse variance weighted method were -0.17 (-0.26, -0.09), 0.15 (0.07, 0.23), and 0.14 (0.08, 0.20), respectively. In addition, per-SD decrease of mtDNA-CN was associated with lower metabolic PC2 (related to lower concentrations of lipids and fatty acids), and higher metabolic PC9 (related to lower concentrations of glycolysis-related metabolites).
Conclusion: Low mtDNA-CN was associated with metabolomic perturbations, particularly reflecting a pro-atherogenic metabolic profile. These findings support the potential of metabolomics to provide new insights into the mitochondrial induced changes preceding chronic diseases.

Introduction

Mitochondria play a central role in the human body, as they are responsible for most of the cellular energy production through oxidative phosphorylation, biosynthetic processes of macromolecules such as heme and steroid hormone, and maintenance of calcium and redox homeostasis for signaling pathways[1-3]. Mitochondria have their own circular genome, the mitochondrial DNA (mtDNA), consisting of 37 genes, 13 of which encode proteins on the electron transport chain. The individual mitochondrion contains several copies of the mitochondrial genome, known as mtDNA copy number (mtDNA-CN), and therefore large variations of mtDNA-CN may exist between cells, tissues, and individuals. Altered mtDNA levels contribute to defects of mtDNA-coded proteins and RNAs, and disrupted cellular signals for apoptosis and autophagy, resulting in mitochondrial dysfunction[4,5]. Consequently, this will further aggravate oxidative stress and inflammation, both of which are key components in many pathophysiological processes, thus promoting chronic diseases and accelerating ageing[6,7].

Mitochondrial dysfunction has been widely considered as a hallmark of the ageing process and compromised health[6,7]. Indeed, mtDNA-CN decreases with increasing age, especially among individuals older than 65 years[8,9]. In prospective cohort studies, low mtDNA-CN measured in the blood leukocyte has been linked to increased risk of several cardiovascular outcomes among the general population[10-13] and mortality among patients with chronic kidney diseases[14], as well as poor poststroke prognosis[15]. Similarly, in our previous studies, we observed associations between low mtDNA-CN and higher risks of coronary artery disease (CAD), heart failure, and stroke[16,17], and for CAD we provided evidence favoring possible causality. However, the underlying mechanisms are not clear yet.

Importantly, all these diseases experience metabolic changes prior to and during their onset and progression. Metabolites are small molecules that result from biological processes, representing effective intermediate phenotypes to a given disease[18]. Metabolomics has offered the opportunity to comprehensively measure a broad range of metabolites in biological fluids, and thus to provide insight into disease mechanisms and to discover potential biomarkers. Therefore, we investigated the associations between mtDNA-CN in leukocytes and blood-derived metabolomic profiles with 168 metabolic biomarkers of predominantly lipids and lipoproteins (sub)particles measured using the Nightingale NMR-based platform in the UK Biobank (UKB). We triangulated our study by performing cross-sectional multivariable-adjusted analyses, followed by Mendelian randomization (MR) studies that are free of most confounding factors and reverse causation.

Methods

Cross-sectional study

Study population

The UKB is a prospective cohort with 502,628 participants between the age of 40 and 69 years recruited from the general population at multiple assessment centers across the UK between 2006 and 2010. 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 written informed consent. 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 biochemical analyses and genotyping. All participants provided electronically written informed consent for the study. Detailed information about the study design, investigation methods, as well as limitations has been reported previously[19]. Further details of the array design, genotyping, and imputation procedures have been described elsewhere[20]. All analyses were conducted under UKB application 56340.

mtDNA-CN

Full genotyped data were available from 488,377 individuals. MtDNA-CN was computed for unrelated European ancestry participants. A detailed step-wise exclusion of the study population was described elsewhere[16]. Somatic blood-derived mtDNA-CN were assessed from the intensities of genotyping probes on the mitochondrial chromosome on the Affymetrix Array. The method for computing mtDNA-CN has been described in detail before (*https://github.com/GrassmannLab/MT_UKB*)[8]. Briefly, the relative amount of mtDNA hybridized to the array at each probe was 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. The median L2R values across all 265 variants passing quality control on the MT chromosome were used as an initial raw measure of mtDNA-CN. To correct for the confounding induced by poorly performing probes, L2R values of each probe were multiplied by the weight of the probe that was generated from a multivariate linear regression model in which those intensities statistically significantly predicted normalized mitochondrial coverage from exome sequencing data. This resulted in a single mtDNA-CN estimate for each individual. To eliminate the plate effect, we subsequently standardized the CN to a mean of zero and standard deviation (SD) of one within each genotyping plate comprising 96 wells. An additional quality control step was performed by eliminating individuals with high standard deviation (SD) (two SD from the mean) of autosomal probes log2 ratio (L2R). Consequently, 293,245 individuals with eligible mtDNA-CN remained.

Metabolomic profiling

A subset of ~120,000 UKB participants were randomly selected for plasma metabolomic profiling using a high throughput NMR-metabolomics platform (Nightingale Health, Helsinki, Finland). Details of the experimentation and applications of the platform have been described previously[21], as well as representative coefficients of variations for the metabolic biomarkers.[22] This platform provides simultaneous quantification of 168 metabolomic measures and 81 ratios from the plasma samples. We included only the direct measures in the current analyses and thus excluded the ratios, including lipoprotein subclasses (size-specific, n = 98), lipoprotein particle sizes (n = 7), cholesterol (n = 15), glycerides and phospholipids (n = 12), total lipids (n = 4), fatty acids (n= 9), amino acids (n= 10), glycolysis-related metabolites (n= 4), fluid balance (n= 2), inflammation (n= 1), ketone bodies (n= 4), apolipoproteins (n= 2). A full list of the measured metabolic biomarkers and their concentrations are presented in **Table S1**.

Statistical analyses

Individuals with any prevalent cardiovascular disease at baseline assessment were excluded. All metabolic biomarkers were log-transformed and subsequently scaled to standard deviation (SD) units for comparison before analyses.
Confounding factors were selected based on previous studies. Confounders included demographic parameters (age at recruitment, sex, deprivation index); the first ten genetic principal components to correct for possible remaining population stratification; genotyping batch; cell numbers (white blood cell counts and platelet counts); an anthropometric measure of body mass index (BMI) in kg/m^2; self-reported lifestyle factors (smoking status [never, past and current], alcohol consumption frequency [twice or less per week/ more than three times per week], physical activity [MET hours per week for moderate-vigorous activity], sleep duration in hours and insomnia symptoms [yes/no]); lipid-lowering medication, blood pressure (mmHg, average of the two measurements taken a few moments apart when applicable).

Multivariable-adjusted linear regression models were performed to estimate the coefficient, representing changes in SD unit of individual metabolic biomarker corresponding to per one-SD lower mtDNA-CN, adjusted for confounding factors. Due to the high correlation of most metabolomic traits, particularly lipid subclasses, conventional correction for multiple testing is too stringent. Therefore, we estimated the number of independent tests from the correlation matrix of the metabolomic measures using the “effective number” (Meff) procedure as described previously to correct for the significance level[23], with an adjusted p-value of 0.05/37 (0.00134), where 37 represents the number of independent metabolomic measures.

Furthermore, given the intercorrelation of metabolic biomarkers, especially among lipid and lipoprotein (sub)particles, we subsequently performed principal component analysis (PCA), which aims at explaining the maximum variation in a multivariate dataset without a priori defined group, on all biomarkers to reduce the large number of correlated biomarkers to a much smaller number of uncorrelated metabolic principal components (PCs). PCs explained at least 1% variations and had an eigenvalue above 1 were selected for further analyses.

Consequently, 9 PCs met the criteria and were selected. The same analyses were performed for each PC as those for individual biomarkers.

We additionally conducted sensitivity analyses for individual biomarkers by excluding lipid-lowering medication users (N = 7770). Missing data in the covariates (all less than 20%) were imputed using multiple imputations by chain equations (MICE)[24], setting the number of imputed datasets to 10. The imputation model included all covariates. A detailed description and comparisons between individuals with and without missing data in the UKB were available in our previous study[16].

Mendelian randomization

Data source

133 independent (linkage disequilibrium < 0.05) nuclear single-nucleotide polymorphisms (SNPs) on autosomal chromosomes were used as genetic instruments that were previously identified in relation to mtDNA-CN at a genome-wide significance threshold ($p < 5e\text{-}08$)[25]. This genome-wide association study (GWAS) was performed in 465,809 individuals of White European ancestry combining the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium and UKB. Genetic associations were adjusted for age, sex, and covariates that were specific in each cohort, such as genetic PCs, blood collection sites, family structure, and cell composition. F statistics [$(\beta/se)^2$] were computed to evaluate instrumental strength. Furthermore, we calculated the proportion of total variance in the exposure explained by each instrument (R^2) separately[26].

For gene-outcome associations, summary statistics (standardized regression coefficients for the SNP-metabolic biomarker association) for individual metabolic biomarkers were obtained from the MR-Base database by the MRC Integrative Epidemiology Unit at the University of Bristol (*https://gwas.mrcieu.ac.uk/*). In addition, we performed new GWAS on 9 identified metabolic PCs. We used the BOLT-LMM software with adjustment for age, sex, genetic principal components. Standardized regression coefficients for the SNP-metabolic PCs were used in the MR analyses.

Mendelian randomization analysis

The primary MR analysis was performed using the inverse-variance weighted (IVW) method to combine the SNP-specific Wald ratios for both individual metabolic biomarkers and metabolic PCs, assuming all instrumental variables are valid[27]. Results were expressed as changes in SD of each biomarker or metabolic PC per one-SD decrease in genetically-predicted mtDNA-CN. When the MR assumptions were met, this estimate approximated the causal effect. Sensitivity analyses accounting for pleiotropy were conducted, including Weighted-Median Estimator and MR-Egger regression[28,29], both of which assume the instrumental variable assumptions are valis for at least half of the SNPs. The intercept from MR-Egger represents the average pleiotropic effect; when the intercept deviates from zero, estimates from IVW might be biased. MR-PRESSO (MR Pleiotropy RESidual Sum and Outlier) was applied to detect and correct for horizontal plei-

Table 1 Baseline charactertistics of the study participants

Variables	N = 61,186
Age (years)	56.5 (8.0)
Sex (female)	33,417 (54.6%)
White blood cell count (10^9/L)	6.9 (1.9)
Platelet count (10^9/L)	253.9 (59.1)
Deprivation index	-1.7 (2.9)
Body mass index (kg/m^2)	27.2 (4.6)
Diastolic blood pressure (mmHg)	82.4 (10.1)
Systolic blood pressure (mmHg)	138.1 (18.6)
Blood pressure-lowering medication	
Yes	10,664 (17.4)
No	50,522 (82.6)
Blood chemistry	
Total cholesterol (mmol/L)	5.8 (1.1)
HDL (mmol/L)	1.5 (0.4)
LDL (mmol/L)	3.6 (0.8)
Triglycerides (mmol/L)	1.7 (1.0)
Cholesterol lowering medication	
Yes	7,770 (12.7)
No	53,416 (87.3)
Physical activity (moderate-vigorous MET hours/week)	27.0 (34.2)
Alcohol consumption frequency	
At least three times per week	28,246 (46.2)
Twice or less per week	32,899 (53.8)
Data missing	41 (0.1)
Smoking status	
Current	5,928 (9.7)
Previous	20,810 (34.0)
Never	34,242 (56.0)
Data missing	206 (0.3)
Sleep duration (hours)	7.1 (1.2)
Insomnia	
Usually	16,716 (27.3)
Sometimes	29,575 (48.3)
Never/rarely	14,855 (24.3)
Data missing	40 (0.1)

Data are mean (SD) for continuous variables and frequency (percentage) for categorical variables. Some percentages do not add up to 100 because of rounding. BMI: Body mass index; HDL: high-density lipoprotein cholesterol; LDL: low-density lipoprotein cholesterol; CVD: cardiovascular disease.

Table 2 Cross-sectional associations of mtDNA-CN with the first 9 metabolomic principal components (PCs)

PCs	Main loadings	Main model1
1	Positive associations with most lipids and lipoproteins (except for extra-large and large HDL particles) and fatty acids.	0.262 (0.194, 0.330)
2	Negative associations with most lipids and lipoproteins (except for LDL subclass chylomicrons and triglycerides within each subparticle), and fatty acids.	-0.194 (-0.235, -0.153)
3	Negatively associated with IDL, LDL, and HDL (but not extra-large HDL);	-0.027 (-0.060, 0.006)
4	Negative associations with HDL	0.022 (-0.001, 0.045)
5	Positive associations with amino acids	-0.049 (-0.066, -0.032)
6	Negative associations with ketone bodies	-0.02 (-0.035, -0.006)
7	Negative associations with ketone bodies (associations with lipids and lipoproteins are opposite to the associations of PC6 with lipids and lipoproteins, despite very weak)	-0.031 (-0.043, -0.018)
8	Positive associations with fatty acids	-0.096 (-0.107, -0.085)
9	Negative associations with glycolysis-related metabolites	-0.038 (-0.049, -0.027)

[1]Betas (95%CI) were adjusted for age at recruitment, sex, deprivation index, the first ten genetic principal components, genotyping batch, white blood cell counts and platelet counts, body mass index, smoking status, alcohol consumption, physical activity, sleep duration and insomnia, lipid-lowering medication, blood pressure and hypertensive medication.

otropy through removing outliers[30]. Moreover, we examined the heterogeneity using Cochran's Q statistic among all SNPs within each outcome database.

All the analyses were performed using R (v4.0.2) statistical software (The R Foundation for Statistical Computing, Vienna, Austria). Packages used in the analyses included "mice" for multiple imputations, "TwoSampleMR" for MR analyses. All results were reported as betas with accompanied 95% confidence intervals.

Results

Characteristics of the study population are presented in **Table 1**. In total, 61,186 unrelated individuals with no history of cardiovascular diseases were included in the analyses with available mtDNA-CN and metabolomic measures. The mean (SD) age was 56.5 (8.0) years, and 54.6% (N= 33,417) were women.

Individual metabolic biomarker

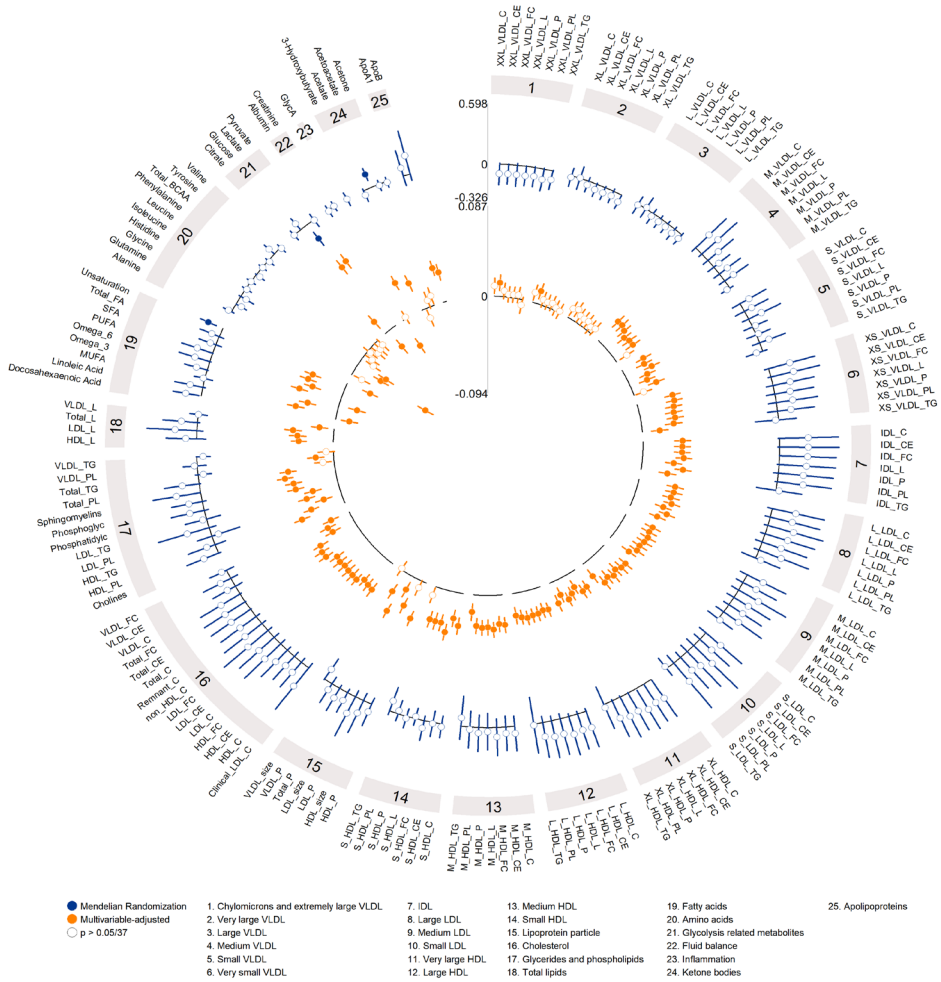


Figure 1 Associations between low leukocytes mtDNA-CN and 168 NMR-based circulating metabolic biomarkers

The inner circle represents the associations derived from multivariable linear regression models among 61,168 unrelated individuals from the UK Biobank, adjusted for age, sex, deprivation index, the first ten genetic principal components, genotyping batch, white blood cell counts and platelet counts, body mass index, smoking status, alcohol consumption, physical activity, sleep duration in hours and insomnia, lipid-lowering medication, blood pressure and hypertensive medication. The outer circle represents the estimates obtained from Mendelian randomization inverse-variance weighted method. All point estimates represent the standardized change in metabolic biomarker concentration per standard deviation change in mtDNA-CN. Hollow points indicate the associations were not statistically significant after correction for multiple testing ($p > 0.05/37$, where 37 represents independent metabolomic measures).

After adjustment for potential confounding factors, mtDNA-CN was significantly associated with 130 out of 168 metabolic biomarkers after correcting for mul-

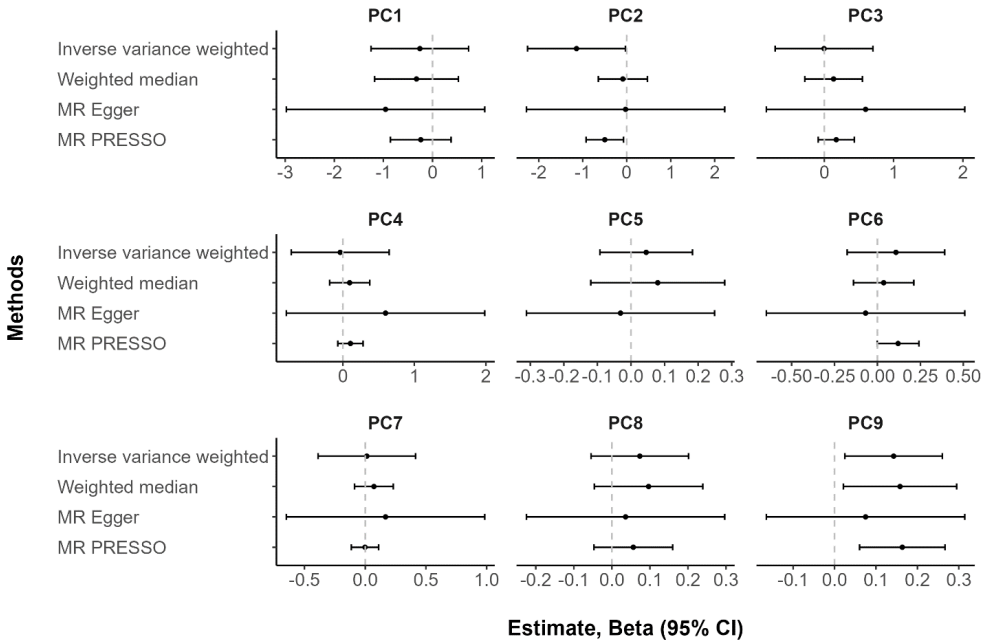


Figure 2 Associations between leukocyte mtDNA-CN with 9 metabolic principal components (PCs) derived from 168 NMR-based metabolic biomarkers in Mendelian randomization (MR)

X-axis represents the estimates derived from Mendelian randomization (MR) analyses and are expressed as changes in SD of each metabolic PC per one-SD decrease in genetically-predicted mtDNA-CN. Y-axis represent different MR methods. No outlier was detected for PC5, and therefore no estimate was available in MR-PRESSO for outlier correction.

tiple testing ($p < 0.05/37$), as shown in **Figure 1** (inner circle). A one-SD lower mtDNA-CN was associated with higher concentrations of most lipoprotein (sub) particles, except for chylomicrons, and the associations with triglycerides within each subclass were generally smaller than other (sub)particles. Similarly, low mtDNA-CN was related to higher concentrations of several biomarkers, including most fatty acids except for omega-3 fatty acids and docosahexaenoic acid, amino acids of alanine, glutamine, glycine, and valine, glycolysis related metabolites of citrate, lactate, and pyruvate, and two apolipoproteins. Results from sensitivity analyses after excluding individuals with lipid-lowering medication remained similar to those obtained from the main analyses (**Table S2**).

We included 112 genetic variants available in individual metabolic biomarker GWAS in the MR analyses. F-statistics for each SNP ranged from 15.6 to 634.4, and all the included genetic variants explained approximately 2.0% variation in mtDNA-CN. Detailed information on genetic variants is available in **Table S3**. Genetically-influenced mtDNA-CN were not associated with any lipid (sub)particles or amino acids upon correction for multiple testing (0.05/37) in the IVW analyses **(Figure 1**, outer circle). Notwithstanding, the associations of mtDNA-CN with lipids and lipoproteins in MR analyses generally showed similar trends

To the best of our knowledge, we provide the first evidence for potential causal associations of mtDNA-CN on metabolomic profiling in a large cohort. Previously, the association between mtDNA-CN with lipid dysregulation defined according to biochemically measured lipids levels is inconclusive. In a recent cross-sectional study including 408,361 participants of multiple ancestries (including UKB), lower mtDNA-CN was associated with higher odds of hyperlipidemia as high total cholesterol or triglycerides or use of lipid-lowering medication[9]. Conversely, mtDNA-CN was not associated with any cardiometabolic phenotypes including lipids among two cohorts of UK young and old women[31]. Nevertheless, in case-control settings, mtDNA-CN was lower among individuals with either hyperlipidemia[32] or metabolic syndrome[33-35]. In only one study with merely 310 participants and 44 measured metabolites, it was observed that mtDNA-CN was associated with circulating metabolites indicative of perturbed lipid metabolism[36]. In the present study, we specifically identify the relationship between low mtDNA-CN with several lipids and lipoproteins reflecting a pro-atherogenic metabolic profile through cross-sectional analyses and MR analyses on metabolic PC2. Although MR results from individual metabolic biomarker did not yield any significant associations on lipids and lipoproteins, this might be due to insufficient statistical power and the trend of these associations were similar to those from multivariable-adjusted linear models. Collectively, altered mitochondrial respiration, reflected by low mtDNA-CN may contribute to lipid dysregulation that leads to further cardiovascular diseases.

Glycolysis is the most important pathway for glucose metabolism in cellular metabolism converting glucose to pyruvate for supplying energy via a series of intermediate metabolites. Pyruvate lies at the crossroads of glycolysis and mitochondrial oxidation, and lactate elevates when the flux through glycolysis overwhelms the utilization of pyruvate in the mitochondria. Often, patients with mitochondrial disease suffer from lactic acidosis, a condition in which lactic acids are increased in the body. In experimental studies, pyruvate has been shown to protect against stroke possibly by mitigating inflammation[37,38]. Nevertheless, most of the currently available epidemiological studies have shown positive associations between high levels of either lactate or pyruvate with a higher risk of several cardiovascular endpoints and poor prognosis[39-42], whereas other studies failed to observe similar relationships[42-45]. In our cross-sectional study, we observed an association between low mtDNA-CN and high concentrations of both lactate and pyruvate. However, in MR analyses, genetically-determined low mtDNA-CN was associated with low lactate and high metabolic PC9 (inversely related to glycolysis-related metabolites). These discrepancies are likely owing to that either lactate or pyruvate solely is not a prominent biomarker of mitochondrial function. Interestingly, the lactate-to-pyruvate ratio in conjunction with high lactate has been considered as the most reliable biomarker to differentiate inherited mitochondrial disorders or disturbances in the tricarboxylic acid cycle[46], whether it is also a stable biomarker for compromised mitochondrial function needs to be validated in future research.

There is no evidence favoring a causal link between mtDNA-CN with other metabolomic measures, including amino acids and fatty acids. For amino acids, results from both cross-sectional and MR consistently showed null associations. Although amino acids have been associated to several metabolic-related dis-

eases from both observational and MR studies[46,47], this may indicate that the effect of low mtDNA-CN on chronic diseases is independent of dysregulation of amino acids. For fatty acids, in spite of the solid associations in cross-sectional analyses, MR analyses that are free of reverse causation and most of the confounding factors from either individual biomarker or metabolic PC that are closely related to fatty acids (PC6) failed to validate these associations. Taken together, fatty acids are unlikely the causal intermediate phenotypes from mitochondrial function to chronic diseases.

The main strength of our study is that we used two different epidemiological methods that have different and largely unrelated sources of potential bias, and two different dimensions, i.e. individual metabolic biomarkers and metabolic PCs, to infer causality. Similar effects converged from different approaches largely strengthens the associations and favors potential causality between mtDNA-CN and metabolomic profiling. Our study had several limitations. First, mtDNA-CN were derived from the blood, which might not be representative for those in different tissues and organs. Tissues with higher energy demands harbor more mitochondria and consequently have a higher mtDNA-CN, such as the heart, brain, and skeletal muscles. However, mtDNA-CN measured in the blood and different tissues are not necessarily intercorrelated[48], and therefore the interpretation of the findings should cautiously restrict to the quantitative mtDNA information specifically about the blood. In addition, the calculation of mtDNA-CN from chip arrays might have introduced noise due to the small number of variants. Consequently, we exploited a validated weighted mtDNA-CN approximating what would be estimated from exome sequencing[8]. Furthermore, mtDNA mutations and deletions have been implicated in the pathogenesis of several diseases, particularly atherosclerosis. Nevertheless, due to the lack of data, we could not perform additional analyses and therefore further deep sequencing of the entire mitochondrial genome will facilitate the understanding of mitochondrial function in metabolomic changes.

Conclusion

Low mtDNA-CN was associated with metabolomic perturbations, particularly a more pro-atherogenic metabolic profile. These findings support the potential of metabolomics to provide new insights into the mitochondrial-induced changes preceding chronic diseases.

References

1. H. Vakifahmetoglu-Norberg, A. T. Ouchida, E. Norberg. The role of mitochondria in metabolism and cell death. Biochem Biophys Res Commun 2017; 482(3): 426-31.
2. D. E. Handy, J. Loscalzo. Redox regulation of mitochondrial function. Antioxid Redox Signal 2012; 16(11): 1323-67.
3. L. Contreras, I. Drago, E. Zampese, T. Pozzan. Mitochondria: the calcium connection. Biochim Biophys Acta 2010; 1797(6-7): 607-18.
4. D. C. Wallace. A mitochondrial bioenergetic etiology of disease. J Clin Invest 2013; 123(4): 1405-12.
5. R. W. Taylor, D. M. Turnbull. Mitochondrial DNA mutations in human disease. Nat Rev Genet 2005; 6(5): 389-402.
6. C. López-Otín, M. A. Blasco, L. Partridge, M. Serrano, G. Kroemer. The hallmarks of aging. Cell 2013; 153(6): 1194-217.
7. C. López-Otín, G. Kroemer. Hallmarks of Health. Cell 2021; 184(1): 33-63.
8. S. Hägg, J. Jylhävä, Y. Wang, K. Czene, F. Grassmann. Deciphering the genetic and epidemiological landscape of mitochondrial DNA abundance. Hum Genet 2020.
9. X. Liu, R. J. Longchamps, K. L. Wiggins, et al. Association of mitochondrial DNA copy number with cardiometabolic diseases. Cell Genom 2021; 1(1).
10. Y. Zhang, E. Guallar, F. N. Ashar, et al. Association between mitochondrial DNA copy number and sudden cardiac death: findings from the Atherosclerosis Risk in Communities study (ARIC). Eur Heart J 2017; 38(46): 3443-8.
11. F. N. Ashar, Y. Zhang, R. J. Longchamps, et al. Association of Mitochondrial DNA Copy Number With Cardiovascular Disease. JAMA Cardiol 2017; 2(11): 1247-55.
12. Y. S. Hong, R. J. Longchamps, D. Zhao, et al. Mitochondrial DNA Copy Number and Incident Heart Failure: The Atherosclerosis Risk in Communities (ARIC) Study. Circulation 2020; 141(22): 1823-5.
13. D. Zhao, T. M. Bartz, N. Sotoodehnia, et al. Mitochondrial DNA copy number and incident atrial fibrillation. BMC Med 2020; 18(1): 246.
14. F. Fazzini, C. Lamina, L. Fendt, et al. Mitochondrial DNA copy number is associated with mortality and infections in a large cohort of patients with chronic kidney disease. Kidney Int 2019; 96(2): 480-8.
15. M. R. Chong, S. Narula, R. Morton, et al. Mitochondrial DNA Copy Number as a Marker and Mediator of Stroke Prognosis: Observational and Mendelian Randomization Analyses. Neurology 2022; 98(5): e470-e82.
16. J. Luo, R. Noordam, J. W. Jukema, et al. Low mitochondrial copy number drives atherogenic cardiovascular disease: evidence from prospective cohort analyses in the UK Biobank combined with Mendelian Randomization. medRxiv; 2021: 2021.07.01.21259854.
17. L. G. Martens, J. Luo, M. J. H. Wermer, et al. The association between lymphocyte mitochondrial DNA abundance and Stroke: a combination of multivariable-adjusted survival and mendelian randomization analyses. medRxiv; 2021: 2021.10.27.21265463.
18. H. Dharuri, A. Demirkan, J. B. van Klinken, et al. Genetics of the human metabolome, what is next? Biochim Biophys Acta 2014; 1842(10): 1923-31.

19. C. Sudlow, J. Gallacher, N. Allen, 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(3): e1001779.
20. C. Bycroft, C. Freeman, D. Petkova, et al. The UK Biobank resource with deep phenotyping and genomic data. Nature 2018; 562(7726): 203-9.
21. P. Soininen, A. J. Kangas, P. Würtz, et al. Quantitative Serum Nuclear Magnetic Resonance Metabolomics in Cardiovascular Epidemiology and Genetics. Circ Cardiovasc Genet. 2015; 8(1): 192-206.
22. J. Kettunen, A. Demirkan, P. Wurtz, et al. Genome-wide study for circulating metabolites identifies 62 loci and reveals novel systemic effects of LPA. Nat Commun 2016; 7: 11122.
23. J. Li, L. Ji. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. Heredity (Edinb) 2005; 95(3): 221-7.
24. I. R. White, P. Royston, A. M. Wood. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med 2011; 30(4): 377-99.
25. R. J. Longchamps, S. Y. Yang, C. A. Castellani, 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(1): 127-46.
26. T. M. Teslovich, K. Musunuru, A. V. Smith, et al. Biological, clinical and population relevance of 95 loci for blood lipids. Nature 2010; 466(7307): 707-13.
27. S. Burgess, A. Butterworth, S. G. Thompson. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol 2013; 37(7): 658-65.
28. J. Bowden, G. Davey Smith, P. C. Haycock, S. Burgess. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. Genet Epidemiol 2016; 40(4): 304-14.
29. J. Bowden, G. Davey Smith, S. Burgess. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol 2015; 44(2): 512-25.
30. M. Verbanck, C. Y. Chen, B. Neale, R. Do. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. Nat Genet 2018; 50(5): 693-8.
31. A. L. Guyatt, K. Burrows, P. A. I. Guthrie, et al. Cardiometabolic phenotypes and mitochondrial DNA copy number in two cohorts of UK women. Mitochondrion 2018; 39: 9-19.
32. C. S. Liu, C. L. Kuo, W. L. Cheng, C. S. Huang, C. F. Lee, Y. H. Wei. Alteration of the copy number of mitochondrial DNA in leukocytes of patients with hyperlipidemia. Ann N Y Acad Sci 2005; 1042: 70-5.
33. D. Révész, J. E. Verhoeven, M. Picard, et al. Associations Between Cellular Aging Markers and Metabolic Syndrome: Findings From the CARDIA Study. J Clin Endocrinol Metab 2018; 103(1): 148-57.
34. J. H. Kim, J. A. Im, D. C. Lee. The relationship between leukocyte mitochondrial DNA contents and metabolic syndrome in postmenopausal women. Menopause 2012; 19(5): 582-7.
35. C. H. Huang, S. L. Su, M. C. Hsieh, et al. Depleted leukocyte mitochondrial DNA copy number in metabolic syndrome. J Atheroscler Thromb 2011; 18(10): 867-73.

36. J. Knez, V. G. Marrachelli, N. Cauwenberghs, et al. Peripheral blood mitochondrial DNA content in relation to circulating metabolites and inflammatory markers: A population study. PLoS One 2017; 12(7): e0181036.
37. Q. Wang, M. van Hoecke, X. N. Tang, et al. Pyruvate protects against experimental stroke via an anti-inflammatory mechanism. Neurobiol Dis 2009; 36(1): 223-31.
38. M. G. Ryou, R. Liu, M. Ren, J. Sun, R. T. Mallet, S. H. Yang. Pyruvate protects the brain against ischemia-reperfusion injury by activating the erythropoietin signaling pathway. Stroke 2012; 43(4): 1101-7.
39. R. Zymliński, J. Biegus, M. Sokolski, et al. Increased blood lactate is prevalent and identifies poor prognosis in patients with acute heart failure without overt peripheral hypoperfusion. Eur J Heart Fail 2018; 20(6): 1011-8.
40. T. Kawase, M. Toyofuku, T. Higashihara, et al. Validation of lactate level as a predictor of early mortality in acute decompensated heart failure patients who entered intensive care unit. J Cardiol 2015; 65(2): 164-70.
41. D. Vojinovic, M. Kalaoja, S. Trompet, et al. Association of circulating metabolites in plasma or serum and risk of stroke: Meta-analysis from seven prospective cohorts. Neurology 2020; 96(8): e1110-23.
42. K. Matsushita, E. K. Williams, M. L. Mongraw-Chaffin, et al. The association of plasma lactate with incident cardiovascular outcomes: the ARIC Study. Am J Epidemiol 2013; 178(3): 401-9.
43. P. Würtz, A. J. Kangas, P. Soininen, D. A. Lawlor, G. Davey Smith, M. Ala-Korpela. Quantitative Serum Nuclear Magnetic Resonance Metabolomics in Large-Scale Epidemiology: A Primer on -Omic Technologies. Am J Epidemiol 2017; 186(9): 1084-96.
44. D. Ko, E. M. Riles, E. G. Marcos, et al. Metabolomic Profiling in Relation to New-Onset Atrial Fibrillation (from the Framingham Heart Study). Am J Cardiol 2016; 118(10): 1493-6.
45. N. Becerra-Tomás, M. Ruiz-Canela, P. Hernández-Alonso, et al. Glycolysis Metabolites and Risk of Atrial Fibrillation and Heart Failure in the PREDIMED Trial. Metabolites 2021; 11(5).
46. F. G. Debray, G. A. Mitchell, P. Allard, B. H. Robinson, J. A. Hanley, M. Lambert. Diagnostic accuracy of blood lactate-to-pyruvate molar ratio in the differential diagnosis of congenital lactic acidosis. Clin Chem 2007; 53(5): 916-21.
47. L. A. Lotta, R. A. Scott, S. J. Sharp, et al. Genetic Predisposition to an Impaired Metabolism of the Branched-Chain Amino Acids and Risk of Type 2 Diabetes: A Mendelian Randomisation Analysis. PLoS Med 2016; 13(11): e1002179.
48. M. Picard. Blood mitochondrial DNA copy number: What are we counting? Mitochondrion 2021; 60: 1-11.

Supplementary materials

Table S1 Information of 168 NMR-based metabolomic biomarkers in the UK Biobank (N= 61,186)

Table S2 Multivariable regression analysis of mtDNA-CN on different metabolomic biomarker excluding individuals with lipid-lowering medication from the UK Biobank

Table S3 Information of genetic variants for mtDNA-CN included in Mendelian randomization analyses

Table S4 Mendelian randomization of mtDNA-CN on different metabolomic biomarker

Table S5 Mendelian randomization of mtDNA-CN on metabolomic PCs from the UK Biobank

Figure S1 The percentage of variance explained by each metabolic principal components in the UK Biobank

Figure S2 Individual metabolic biomarker loading in metabolic principal components

The Supplementary materials for this article can be found online at:
https://drive.google.com/drive/folders/1d46G5jf6fIZUUlp6aHpjnz74SnLjzQPd?usp=sharing

CHAPTER 5

Dietary-derived circulating antioxidants and risk of coronary heart disease: a Mendelian randomization study

Jiao Luo, Saskia le Cessie, Diana van Heemst, Raymond Noordam

J Am Coll Cardiol. 2021 Jan 5;77(1):45-54

Abstract

Objectives: To investigate possible causal associations between dietary-derived circulating antioxidants and primary CHD risk using two-sample Mendelian Randomization (MR).

Background: Previously, observational studies have identified associations between higher levels of dietary-derived antioxidants and lower risk of coronary heart disease (CHD), while randomized clinical trials showed no reduction in CHD risk following antioxidant supplementation.

Methods: Single-nucleotide polymorphisms (SNPs) for circulating antioxidants (vitamin E, C, retinol, β-carotene, and lycopene), assessed as absolute levels and metabolites, were retrieved from literature and were used as genetic instrumental variables. Summary statistics for gene-CHD associations were obtained from three databases (cases/controls): the CARDIoGRAMplusC4D consortium (60 801/123 504), UK Biobank (25 306/46 2011) and FinnGen study (7123/89 376), respectively. For each exposure, MR analyses were performed per outcome database, and subsequently meta-analyzed.

Results: Among the analytic sample of 768 121 individuals (93 230 cases), genetically predicted circulating antioxidants were not causally associated with CHD risk. For absolute antioxidants, the odds ratio (95% CI) for CHD ranged between 0.94 (0.63, 1.41) for retinol and 1.03 (0.97, 1.10) for β-carotene per unit increase in ln-transformed antioxidant values. For metabolites, the odds ratio ranged between 0.93 (0.82, 1.06) for γ-tocopherol and 1.01 (0.95, 1.08) for ascorbate per 10-fold increase in metabolite levels.

Conclusions: Evidence from our study did not support a protective effect of genetic predisposition to high dietary-derived antioxidant levels on CHD risk. Therefore, it is unlikely that taking antioxidants to increase blood antioxidants levels, will have a clinical benefit for the prevention of primary CHD.

Introduction

Coronary heart disease (CHD) is one of the foremost causes of mortality worldwide and is responsible for approximately 0.36 million of all deaths in the US and 1.78 million in Europe each year [1-3]. Well-established risk factors for CHD include smoking, physical inactivity, inappropriate nutrition, overweight and obesity, high blood cholesterol and other lipids, high blood pressure, diabetes mellitus, and insufficient/long sleep [2], and interventions targeted to ameliorate these risk factors showed a significant reduction in CHD risk. Apart from conventional risk factors, oxidative stress has also been hypothesized as a vital component in the development and progression of CHD by promoting macromolecular damage and endothelium dysfunction [4]. Consequently, antioxidants, the scavengers of free radicals to diminish oxidative stress-induced damage, would be of interest as targets for primary CHD prevention [5]. Specifically, dietary-derived antioxidants, better known as vitamin E, C, and carotenoids, are the most easily accessible and modifiable approach for consideration.

Based on this hypothesis, a large amount of studies have been conducted to explore the association between antioxidants and primary CHD. In multiple observational studies, dietary intake, either as dietary components or supplements, or blood concentration of vitamin E, C, and carotenoids were associated with a lower risk of primary CHD [6-12]. Similarly, adherence to a diet containing high amounts of antioxidants, irrespective of the type of antioxidants, was associated with a lower risk of cardiovascular diseases [13]. However, associations as these in observational studies are prone to biases including reverse causality and unmeasured confounding. While randomized clinical trials (RCTs) generally failed to demonstrate a causal benefit of antioxidants supplement on primary CHD [14-20], with the exception of lycopene supplement on cardiovascular risk factors [21], there are notable limitations. For example, timing, dosage, duration, use of natural or synthetic antioxidants, as well as the uncertain time of onset and long-term progression of CHD pathogenesis might explain the observed null effect [22,23]. Therefore, the conflicting results from observational studies and RCTs should be interpreted with caution, and the causality between dietary-derived antioxidants and CHD is still unclear.

Mendelian Randomization (MR) is an alternative approach to infer causality of life-long risk factors (exposure) on diseases (outcome) using genetic variants as instrumental variables [24]. In the present study, we used MR analyses to assess the associations between genetically determined dietary-derived circulating antioxidants and their metabolites with primary CHD risk, in the absence of reverse causality and residual confounding factors.

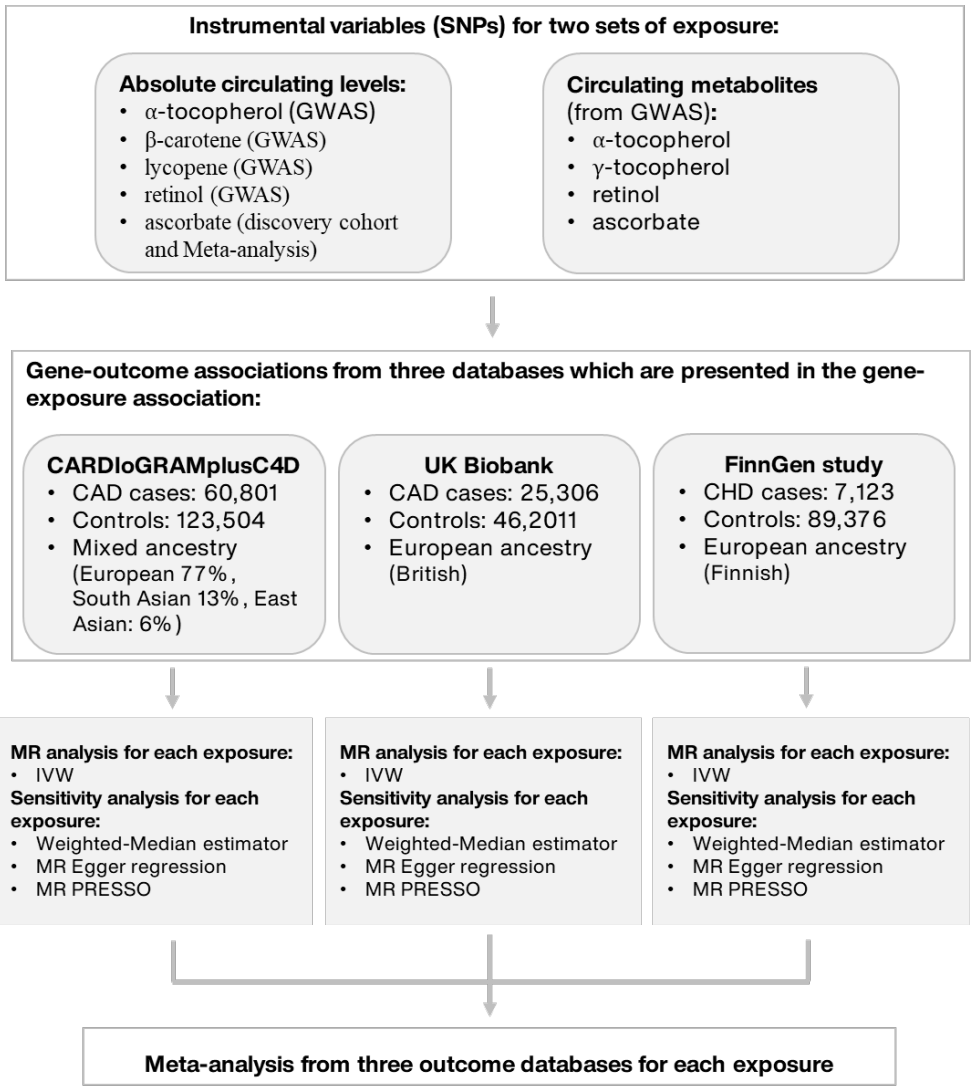


Figure 1 Schematic overview of the present study design

CAD: Coronary Artery Disease; CARDIoGRAMplusC4D: Coronary Artery Disease Genome-Wide Replication and Meta-analysis plus the Coronary Artery Disease Genetics; CHD: Coronary Heart Disease; GWAS, genome-wide association study; IVW: Inverse-Variance weighted; MR PRESSO: MR Pleiotropy RESidual Sum and Outlier.

Methods

Study design

For the current study, we conducted a two-sample MR, which tests the association between genetic instrumental variable(s), as a proxy for the exposure, and outcome from three separate data sources, and estimates the effect of an

exposure on an outcome [25]. MR is based on 3 principle assumptions, notably that the genetic variant(s) should be 1) associated with the exposure, 2) associated with the outcome exclusively through the exposure, and 3) independent of any measured and unmeasured confounders. Data used in the present study are publicly available, and ethical approval and informed consent were obtained in each original study. A schematic overview of the present study design is presented in **Figure 1**.

Selection of genetic instrumental variables

Five main dietary-derived antioxidants were considered in the present study, which included vitamin E (α-tocopherol and γ-tocopherol), β-carotene, lycopene, vitamin C (L-ascorbic acid or ascorbate), and retinol. We considered both antioxidants that were measured as authentic absolute levels in the blood, and their corresponding circulating metabolites that were quantified as relative concentrations in plasma or serum, respectively. For absolute antioxidants levels, α-tocopherol, β-carotene, lycopene, ascorbate, and retinol were identified, while for antioxidants metabolites, α-tocopherol, γ-tocopherol, ascorbate, and retinol were used.

Genome-Wide Association Studies (GWAS) were searched to extract leading SNPs as genetical instrumental variables. When we identified multiple GWAS for a single trait, only the largest study with replication was used [26-31]. Though GWAS was not available for absolute ascorbate levels, a study with a two-stage design which used a discovery cohort
and five replication cohorts and consequently meta-analysis assessed the relationship between genetic variants located in vitamin C active transporter locus of SLC23A1 (Solute Carrier Family 23 Member 1) and circulating levels of ascorbic acid and therefore was considered as qualified for genetic instrument extraction [32]. A summary table of instruments is presented in **Supplemental Table 1**.

Absolute circulating antioxidants

Three SNPs for α-tocopherol levels were identified in a GWAS with 7781 European individuals [26]. However, those 3 loci were previously reported to be associated with lipid metabolism and/or regulation in GWAS on lipid levels [33,34] or coronary artery disease [35], and therefore were not considered for MR analysis due to likely pleiotropic bias. Three genetic variants (linkage disequilibrium, LD < 0.2 as indicated in the study, $p < 5\times10^{-8}$) associated with plasma β-carotene levels were identified in a GWAS within 2344 participants in the Nurses' Health Study [27]. Five variants (LD < 0.001, $p < 5\times10^{-6}$) associated with circulating lycopene level were identified in a GWAS performed in 441 older Amish adults [28]. Two SNPs (LD < 0.001, $p < 5\times10^{-8}$) associated with circulating retinol levels were identified in a GWAS of 5006 Caucasian individuals from two cohorts [29]. As for ascorbate, one genetic variant ($p = 2.0\times10^{-7}$) was identified with over 15 000 participants [32]. Summary of the demographic characteristics of the cohort that were used to generate genetic instrumental variables is presented in **Supplemental Table 2**.

Circulating antioxidants metabolites
Genetic variants for each metabolite at suggestive genome-wide significance level ($p < 1\times10^{-5}$) were extracted from published GWAS [30,31], notably 11 instruments for α-tocopherol (n = 7276), 13 for γ-tocopherol (n = 5822) and 14 for ascorbate (n = 2063) derived from 7824 adult individuals from 2 European population studies, 24 for retinol (n = 1957) from 1960 subjects of European descent. Linkage disequilibrium between all SNPs for the same exposure was assessed, and when LD was present (LD > 0.001), the variant with the smallest P-value was selected.

Explained variance and instrument strength
Variance (R^2) in the MR study refers to the proportion of total variation in the exposure which is explained by the genetic instruments. R^2 for each trait were either derived from the original study or calculated based on the derived summary statistics in line with what has been described previously [34], and ranged from 0.9% to 30.1% for absolute antioxidant levels, and from 3.3% to 18.6% for antioxidants metabolites, separately (**Supplemental Table 1**).

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.

Data source for instrument-outcome associations
Summary statistics for the associations of the identified exposure-related SNPs with primary CHD were extracted from three large databases, namely Coronary Artery Disease Genome-Wide Replication and Meta-analysis plus the Coronary Artery Disease Genetics (CARDIoGRAMplusC4D) consortium, UK Biobank, and FinnGen study.

CARDIoGRAMplusC4D assembled 60 801 cases and 123 504 controls for 48 studies, of which, 77% of the participants were of European ancestry, 19% were of south and east Asian ancestry, and a small proportion of Hispanic and African Americans. CAD cases were identified as an inclusive diagnosis of myocardial infarction, acute coronary syndrome, chronic stable angina, or coronary stenosis > 50% [36,37]. The summary statistics of the instruments-CAD associations were provided in the database.

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 [38] (*https://www.ukbiobank.ac.uk*). We restricted the analyses to the participants who reported their ancestry as European, and who were in the full released imputed genomics databases (UK10K + HRC). CAD diagnoses were coded according to the International Classification of Diseases (ICD) [38], and CAD cases were retrieved from linkage with NHS database and were defined as angina pectoris (I20), myocardial infarction (I21 and I22), and acute and chronic ischemic heart disease (I24 and I25). In total, 25 306 cases and 46 2011 controls were identified. We performed logistic regression analyses to assess the associations between genetic instruments and CAD, adjusted for

age, sex, and 10 principal components and corrected for a familial relationship using BOLD_LMM (v2.3.2).

The FinnGen study is an ongoing nation-wide cohort study launched in 2017, which combines genetic data from Finnish biobanks and health record data from Finnish health registries (*https://www.finngen.fi/en*). Major CHD was defined as angina pectoris (I20), myocardial infarction (I21-I23), ischemic heart diseases (I24 and I25), cardiac arrest (I46), and other unattended or cause unknown sudden death (R96 and R98). The analyses were based on the FinnGen data freeze 2, which consists of 7123 cases of major CHD and 89 376 controls with complete instruments-CHD associations.

Statistical analysis

All the analyses were undertaken using R (v3.6.1) statistical software (The R Foundation for Statistical Computing, Vienna, Austria).

Mendelian Randomization

The primary MR analysis was conducted by using Inverse-Variance weighted (IVW) regression analysis, which assumes the absence of invalid genetic instruments (e.g., no directional pleiotropy) [39]. The mean effect estimate was obtained from each outcome database separately by a fixed-effect IVW meta-analysis of the Wald ratios (gene-outcome [log odds ratio] divided by gene-exposure associations) estimated for each instrumental variable [40]. Results are expressed as odds ratios (OR) on CHD risk for a corresponding unit change in absolute circulating levels of antioxidants on natural log-transformed levels (β-carotene and retinol), μg/dL (lycopene) or μmol/L (ascorbate), or a 10-fold change in metabolites concentrations. When the MR assumptions are met, this odds ratio is an estimate of the causal effect of the exposure on the outcome. MR analyses were performed using the R-based package "TwoSampleMR" (*https://mrcieu.github.io/TwoSampleMR/*).

Sensitivity analysis

In order to examine whether there was a violation of the main MR assumptions due to directional pleiotropy, we performed MR-Egger regression analysis and Weighted-Median Estimator [40-42]. In MR-Egger, the intercept estimates the average pleiotropic effect across the genetic variants; a value that differs from zero indicates that the IVW estimate is biased [41]. A Weighted-Median estimator analysis can provide a consistent valid estimate if at least half of the instrumental variables are valid [42]. In addition, MR-PRESSO (MR Pleiotropy RESidual Sum and Outlier) was applied to detect and correct for horizontal pleiotropy through removing outliers [43], as implemented in the R-based package MRPRESSO (*https://github.com/rondolab/MR-PRESSO*). Furthermore, Cochran's Q statistic was used to test the heterogeneity among the estimated Wald ratios from different genetic variants [44]. Additional sensitivity analyses were performed for β-carotene by using a stringent LD threshold of $r^2 < 0.001$, and for lycopene by restricting the analyses to only GWAS significant-level ($p < 5\times10^{-8}$) SNPs.

Meta-analysis of the estimates from three outcome databases
All exposure-specific MR analyses were performed in each outcome database of CARDIoGRAMplusC4D consortium, UK Biobank, and FinnGen study, separately, and then were meta-analyzed to generate the pooled estimates for each exposure on CHD risk. We calculated I^2 statistics to quantify heterogeneity between estimates from three studies and the corresponding p-value derived from Cochran's Q test. Given no heterogeneity was present across three databases, fixed-effect model meta-analyses were used to pool instrumental variable estimates across the three outcome databases for each exposure. All meta-analyses were performed in the R-based "meta" package (*https://cran.r-project.org/web/packages/meta/index.html*).

Results

Summary information of instruments identified for dietary-derived antioxidants and their metabolites are presented in **Supplemental Table 1**, and summary information of the cohorts contributing to the GWAS of absolute levels is given in **Supplemental Table 2**. Retinol and ascorbate are available both as absolute circulating antioxidants and metabolites. Detailed information on the genetic variants, their associations with antioxidants ($\beta_{gene\text{-}exposure}$), and with CHD ($\beta_{gene\text{-}outcome}$) across databases is given in **Supplemental Tables 3 and 4**. F-statistics for all genetic instruments used in the present study were above 10.

Absolute circulating antioxidants and CHD

Overall, in the primary analyses using IVW, genetically determined absolute dietary-derived antioxidants levels were not associated with the risk of CHD in any of the three databases (**Figure 2 and Supplemental Table 5**). Pooled OR for CHD per unit increase of antioxidants were 1.03 (95% confidence interval, CI: 0.97, 1.10) and 0.94 (95%CI: 0.63, 1.41) for natural log-transformed β-carotene and retinol, 1.02 (95%CI: 0.99, 1.06) for 1 μg/dL lycopene, and 1.00 (95%CI: 0.99, 1.00) for 1 μmol/L ascorbate, respectively.

For β-carotene and lycopene with three or more genetic instruments, Weighted-Median estimator and MR-Egger regression were conducted. The estimates did not change substantially compared with IVW regression (**Supplemental Figure 1**). In addition, MR-Egger regression analysis suggested no evidence of overall pleiotropy, and there was no evidence of heterogeneity between individual genetic instrument estimation (**Supplemental Table 5**). In addition, no outlier SNP was identified in MR-PRESSO test for these two antioxidants in any of the databases.

In the sensitivity analysis for β-carotene with LD < 0.001 using 2 SNPs (rs6564851 and rs7501331), and for lycopene with only GWAS-level significant variant (rs7680948, $p < 5\times10^{-8}$), similar results were observed (**Supplemental Figure 2**).

Circulating antioxidants metabolites and CHD

Consistent with the findings from absolute circulating antioxidants, no association between genetically predicted circulating antioxidants metabolites concentration

and CHD risk was observed, as shown in **Figure 3**. The combined ORs for CHD per 10-fold increase in metabolites concentration were 1.00 (95% CI: 0.75, 1.35) for α-tocopherol, 0.93 (0.82, 1.06) for γ-tocopherol, 1.00 (0.98, 1.02) for retinol and 1.01 (0.95, 1.08) for ascorbate.

Sensitivity analyses for metabolites on CHD risk were provided in **Supplemental Table 6**. Estimates using the Weighted-Median estimator were consistently comparable to those from IVW regression. No horizontal pleiotropy was detected in MR-Egger regression, with the exception of retinol in UK Biobank database (intercept: -0.028, se: 0.010, p-value: 0.01). Although MR PRESSO detected outliers for α-tocopherol, retinol, and ascorbate in the UK Biobank database, the estimate did not change materially after correction.

Heterogeneity was detected using Cochran's Q statistics for all metabolites in different databases, especially with outliers as identified using MR PRESSO. However, in the leave-one-out analyses, we found that the risk estimates of genetically predicted antioxidants' metabolites and risk of CHD did not change substantially after excluding one SNP at each time, indicating that it is unlikely that potential driving SNPs could bias the causal association (data not shown).

Discussion

In the present study, we investigated the relationship between dietary-derived antioxidants and CHD risk using Mendelian Randomization. Instrumental variables were used as proxies for circulating antioxidants assessed both as absolute levels and metabolites, and comparable results were obtained. Our findings indicate that dietary-derived antioxidants are unlikely to be causal determinants of primary CHD risk.

Two previous studies using an MR approach found that a genetic predisposition to high a vitamin E level was associated with an increased risk of CAD [45,46]. However, instruments selected in these studies play clear roles in lipid metabolism, which violated the InSIDE assumption in MR design and introduced bias in the effect estimates [47]. To provide insights into the magnitude of the effects of genetic instruments on circulating antioxidant levels, we compared the effects from the genetic instruments and dietary supplementation (**Supplemental Tables 7 and 8**). The effects on circulating antioxidant levels achieved by genetic instruments, with the exception of ascorbate, are within the range of the effects observed by antioxidant supplementation in RCTs, for which we prioritized the trials on cardiovascular outcomes that have been included in previous meta-analyses [14-20]. However, direct comparisons between these two effects should be interpreted with caution, given that the effect of genetic predisposition is assumed to be lifelong whereas the effect of supplementation only lasts for the duration of the trial. The exposure during the whole life course with a slightly minor effect could have a potential biological effect that exceeds the temporarily larger effect of supplements given the long period needed to develop coronary heart disease. The robust null results in our studies, however, suggest that lifelong exposure to somewhat higher antioxidant levels did not decrease the risk of CHD, in line with earlier findings from the trials and meta-analyses on trials.

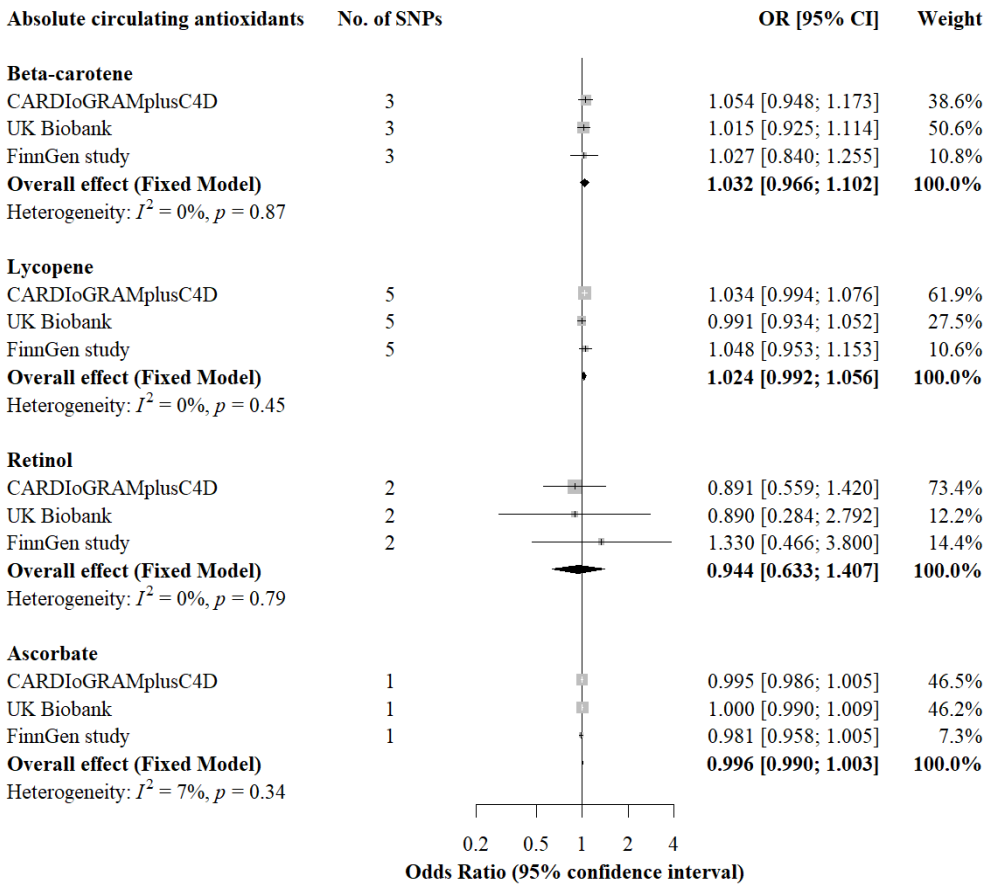


Figure 2 Causal association between absolute circulating antioxidants with Coronary Heart Disease

Estimated odds ratios for the effect of per unit increase in ln-transformed β-carotene and retinol values, 1 μg/dL lycopene and 1 μmol/L ascorbate on coronary heart disease, obtained from an Inverse-Variance Weighted (IVW) analysis, per outcome database separately and combined over the three databases using fixed-effect meta-analyses.

Study strengths and limitations

There are two main strengths in the present study. First, we used two separate sets of instrumental variables, for both absolute circulating antioxidants and their metabolites. Specifically, for retinol and ascorbate that is presented in both sets, similar results were generated, which is supportive of the robustness of our findings. Second, three large databases comprising 768 121 participants with 93 230 CHD cases for gene-outcome associations were meta-analyzed in the present study. The results from these 3 databases are generally consistent with no evidence of heterogeneity. Therefore, the precision for final MR estimates and the reliability of the results were significantly improved despite the limited number of strong genetic instruments.

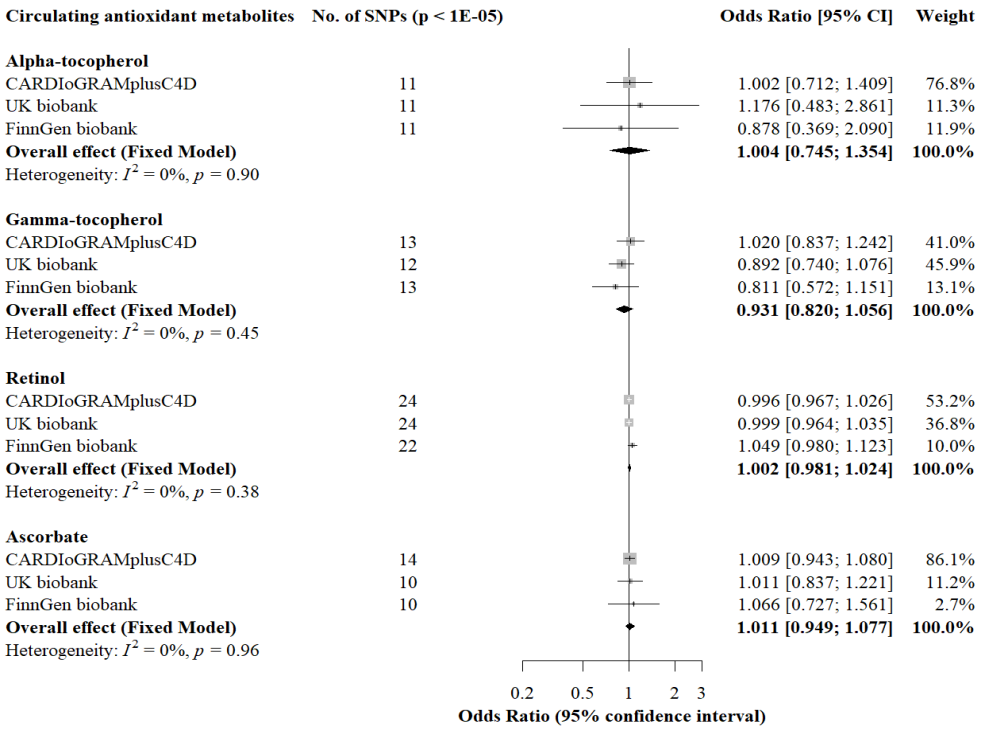


Figure 3 Causal association between circulating antioxidants metabolites with Coronary Heart Disease

Estimated odds ratios for the effect of per 10-fold increase in antioxidants metabolites' concentrations on coronary heart disease, obtained from an Inverse-Variance Weighted (IVW) analysis, per outcome database separately and combined over the three databases using fixed-effect meta-analyses.

This study has some limitations to address. First, we are unable to test for a non-linear causal association between the antioxidant levels and CHD that has previously been suggested, especially for α-tocopherol and β-carotene [6]. Despite analytic methods having been developed, these require individual-level data of the exposure [40,48], and the published data we used are summary-level statistics, therefore we were not able to perform such analyses. Second, no sensitivity analysis could be performed for some absolute antioxidants (retinol and ascorbate) with limited genetic variants. Notwithstanding, these instruments are mapped in the genes which are crucial in the metabolism of antioxidants and are not associated with any other CHD risk factors in the GWAS catalog or PhennoScanner databases, suggesting no directional pleiotropy as also confirmed in the analyses. Third, only one SNP with a small R^2 of 0.9% for absolute ascorbate was used. However, with the considerable instrumental strength and large sample size and cases in which the analyses were conducted, we had more than sufficient statistical power to estimate a possible causal effect [49]. In addition, results from ascorbate metabolites with larger R^2 (18.6%) gave very similar estimates which further reinforces the validity of the findings. Forth, protective effects of antioxidants might still exist in discriminatingly selected subgroups who have

elevated oxidative stress levels, for example, vitamin E supplement provided cardiovascular-protective effects only in individuals with both diabetes and haptoglobin2-2 genotype. Besides, multiple treatments simultaneously might be more effective for multifactorial diseases since there could be a synergistic benefit from two agents with acceptable safety and efficacy, for example, traditional treatment plus antioxidants, to achieve better effects than antioxidants only [50]. However, we could not explore these associations in the population with high risk or a known nutritional deficiency that might be more promising for antioxidant supplements or test the effect of antioxidants in combination with other treatments. Lastly, although no causal associations between circulating antioxidants and CHD risk were detected, we could not completely rule out the possibility that the effect size is too small to be identified even within our large sample size. Nevertheless, such a potential effect, if it exists at all by incorporating additional databases, will be extremely small and is unlikely to result in a clinically relevant reduction of CHD risk as obtained by other strategies, for instance, a 25% to 45% reduction of cardiovascular events of statins, physical activity or weight loss for primary CHD prevention [51-54].

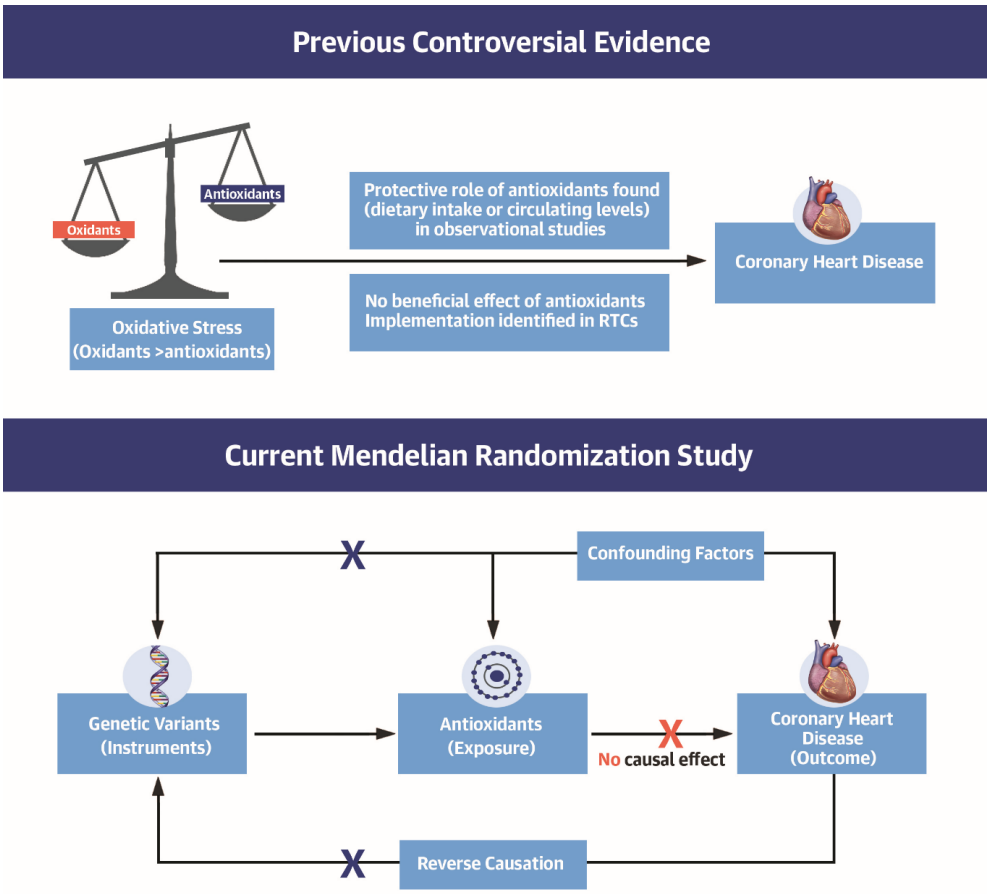
Conclusion

In conclusion, evidence from the present study did not support a beneficial role of circulating dietary-derived vitamin E, C, β-carotene, lycopene, or retinol on CHD risk in the general population. This signifies the absence of a substantial role of antioxidants supplements on CHD risk identified in RCTs and is in accordance with the recommendations from the U.S. Preventive Services Task Force [19]. Therefore, for healthy adults without nutritional deficiency, dietary-derived antioxidant supplement use that improves circulating antioxidant levels to prevent primary CHD is of limited clinical benefit.

Clinical Perspectives

Competency in Medical Knowledge: Dietary-derived circulating antioxidants (vitamin E, vitamin C, β-carotene, lycopene and retinol) are not causally associated with coronary heart disease in healthy adults without nutritional deficiency.

Translational Outlook: Future studies are needed to confirm the effect of dietary-derived antioxidants in discriminatingly selected subgroups who have elevated oxidative stress or known nutritional deficiency and to investigate the preventive and therapeutic effect of antioxidants in combination with other agents or traditional treatments that may generate synergistic benefit than antioxidants only.



Central Illustration Antioxidants and Coronary Heart Disease

Theoretically, antioxidants can act as scavengers of oxidants to mitigate oxidative stress-induced damage, thus preventing coronary heart disease. However, findings from our study demonstrated that there is no casual association between dietary-derived antioxidants with primary coronary heart disease risk.

References

1. M. Nichols, N. Townsend, P. Scarborough, M. Rayner. Cardiovascular disease in Europe 2014: epidemiological update. Eur Heart J 2014; 35(42): 2950-9.
2. E. J. Benjamin, P. Muntner, A. Alonso, et al. Heart Disease and Stroke Statistics-2019 Update: A Report From the American Heart Association. Circulation 2019; 139(10): e56-e528.
3. E. S. Ford, U. A. Ajani, J. B. Croft, et al. Explaining the decrease in U.S. deaths from coronary disease, 1980-2000. N Engl J Med 2007; 356(23): 2388-98.
4. A. Daiber, F. D. Lisa, M. Oelze, et al. Crosstalk of mitochondria with NADPH oxidase via reactive oxygen and nitrogen species signalling and its role for vascular function. Br J Pharmacol 2017; 174(12): 1670-89.
5. K. Sugamura, J. F. Keaney. Reactive oxygen species in cardiovascular disease. Free Radic Biol Med 2011; 51(5): 978-92.
6. D. Aune, N. Keum, E. Giovannucci, et al. Dietary intake and blood concentrations of antioxidants and the risk of cardiovascular disease, total cancer, and all-cause mortality: a systematic review and dose-response meta-analysis of prospective studies. Am J Clin Nutr 2018; 108(5): 1069-91.
7. P. Knekt, J. Ritz, M. A. Pereira, et al. Antioxidant vitamins and coronary heart disease risk: a pooled analysis of 9 cohorts. Am J Clin Nutr 2004; 80(6): 1508-20.
8. G. Li, Y. Li, X. Chen, H. Sun, X. Hou, J. Shi. Circulating tocopherols and risk of coronary artery disease: A systematic review and meta-analysis. Eur J Prev Cardiol 2016; 23(7): 748-57.
9. B. Song, K. Liu, Y. Gao, et al. Lycopene and risk of cardiovascular diseases: A meta-analysis of observational studies. Mol Nutr Food Res 2017; 61(9).
10. J. Kim, J. Choi, S. Y. Kwon, et al. Association of Multivitamin and Mineral Supplementation and Risk of Cardiovascular Disease: A Systematic Review and Meta-Analysis. Circ Cardiovasc Qual Outcomes 2018; 11(7): e004224.
11. Z. Ye, H. Song. Antioxidant vitamins intake and the risk of coronary heart disease: meta-analysis of cohort studies. Eur J Cardiovasc Prev Rehabil 2008; 15(1): 26-34.
12. H. M. Cheng, G. Koutsidis, J. K. Lodge, A. W. Ashor, M. Siervo, J. Lara. Lycopene and tomato and risk of cardiovascular diseases: A systematic review and meta-analysis of epidemiological evidence. Crit Rev Food Sci Nutr 2019; 59(1): 141-58.
13. M. Parohan, J. Anjom-Shoae, M. Nasiri, et al. 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(6): 2175-89.
14. L. Al-Khudairy, N. Flowers, R. Wheelhouse, et al. Vitamin C supplementation for the primary prevention of cardiovascular disease. Cochrane Database Syst Rev 2017; (3).
15. L. Schwingshackl, H. Boeing, M. Stelmach-Mardas, et al. Dietary Supplements and Risk of Cause-Specific Death, Cardiovascular Disease, and Cancer: A Systematic Review and Meta-Analysis of Primary Prevention Trials. Adv Nutr 2017; 8(1): 27-39.
16. S. K. Myung, W. Ju, B. Cho, 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: f10.

17. S. P. Fortmann, B. U. Burda, C. A. Senger, J. S. Lin, E. P. Whitlock. Vitamin and mineral supplements in the primary prevention of cardiovascular disease and cancer: An updated systematic evidence review for the U.S. Preventive Services Task Force. Ann Intern Med 2013; 159(12): 824-34.
18. D. J. A. Jenkins, J. D. Spence, E. L. Giovannucci, et al. Supplemental Vitamins and Minerals for CVD Prevention and Treatment. J Am Coll Cardiol 2018; 71(22): 2570-84.
19. V. A. Moyer, on behalf of the U.S. Preventive Services Task Force. Vitamin, Mineral, and Multivitamin Supplements for the Primary Prevention of Cardiovascular Disease and Cancer: U.S. Preventive Services Task Force Recommendation Statement. Ann Intern Med 2014; 160(8): 558-64.
20. A. W. Ashor, R. Brown, P. D. Keenan, et al. Limited evidence for a beneficial effect of vitamin C supplementation on biomarkers of cardiovascular diseases: an umbrella review of systematic reviews and meta-analyses. Nutr Res 2019; 61: 1-12.
21. H. M. Cheng, G. Koutsidis, J. K. Lodge, A. Ashor, M. Siervo, J. Lara. Tomato and lycopene supplementation and cardiovascular risk factors: A systematic review and meta-analysis. Atherosclerosis 2017; 257: 100-8.
22. S. R. Steinhubl. Why Have Antioxidants Failed in Clinical Trials? Am J Cardiol 2008; 101(10, Supplement): S14-S9.
23. M. P. Murphy. Antioxidants as therapies: can we improve on nature? Free Radic Biol Med 2014; 66: 20-3.
24. D. A. Lawlor, R. M. Harbord, J. A. C. Sterne, et al. Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology. Stat Med 2008; 27(8): 1133-63.
25. S. Burgess, R. A. Scott, N. J. Timpson, et al, Epic- InterAct Consortium. Using published data in Mendelian randomization: a blueprint for efficient identification of causal risk factors. Eur J Epidemiol 2015; 30(7): 543-52.
26. J. M. Major, K. Yu, W. Wheeler, et al. Genome-wide association study identifies common variants associated with circulating vitamin E levels. Hum Mol Genet 2011; 20(19): 3876-83.
27. S. J. Hendrickson, A. Hazra, C. Chen, et al. 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(6): 1379-89.
28. C. R. D'Adamo, A. D'Urso, K. A. Ryan, et al. A Common Variant in the SETD7 Gene Predicts Serum Lycopene Concentrations. Nutrients 2016; 8(2): 82.
29. A. M. Mondul, K. Yu, W. Wheeler, et al. Genome-wide association study of circulating retinol levels. Hum Mol Genet 2011; 20(23): 4724-31.
30. S. Y. Shin, E. B. Fauman, A. K. Petersen, et al. An atlas of genetic influences on human blood metabolites. Nat Genet 2014; 46(6): 543-50.
31. T. Long, M. Hicks, H. C. Yu, et al. Whole-genome sequencing identifies common-to-rare variants associated with human blood metabolites. Nat Genet 2017; 49(4): 568-78.
32. N. J. Timpson, N. G. Forouhi, M. J. Brion, et al. Genetic variation at the SLC23A1 locus is associated with circulating concentrations of L-ascorbic acid (vitamin C): evidence from 5 independent studies with >15,000 participants. Am J Clin Nutr 2010; 92(2): 375-82.
33. C. J. Willer, E. M. Schmidt, S. Sengupta, et al. Discovery and refinement of loci associated with lipid levels. Nat Genet 2013; 45(11): 1274-83.

34. T. M. Teslovich, K. Musunuru, A. V. Smith, et al. Biological, clinical and population relevance of 95 loci for blood lipids. Nature 2010; 466(7307): 707-13.
35. J. M. M. Howson, W. Zhao, D. R. Barnes, et al. Fifteen new risk loci for coronary artery disease highlight arterial-wall-specific mechanisms. Nat Genet 2017; 49(7): 1113-9.
36. M. Nikpay, A. Goel, H. H. Won, et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. Nat Genet 2015; 47(10): 1121-30.
37. P. Deloukas, S. Kanoni, C. Willenborg, et al. Large-scale association analysis identifies new risk loci for coronary artery disease. Nat Genet 2013; 45(1): 25-33.
38. C. Sudlow, J. Gallacher, N. Allen, 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(3): e1001779.
39. S. Burgess, A. Butterworth, S. G. Thompson. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol 2013; 37(7): 658-65.
40. J. R. Staley, S. Burgess. Semiparametric methods for estimation of a nonlinear exposure-outcome relationship using instrumental variables with application to Mendelian randomization. Genet Epidemiol 2017; 41(4): 341-52.
41. J. Bowden, G. Davey Smith, S. Burgess. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol 2015; 44(2): 512-25.
42. J. Bowden, G. Davey Smith, P. C. Haycock, S. Burgess. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. Genet Epidemiol 2016; 40(4): 304-14.
43. M. Verbanck, C. Y. Chen, B. Neale, R. Do. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. Nat Genet 2018; 50(5): 693-8.
44. M. Fd Greco, C. Minelli, N. A. Sheehan, J. R. Thompson. Detecting pleiotropy in Mendelian randomisation studies with summary data and a continuous outcome. Stat Med 2015; 34(21): 2926-40.
45. T. Wang, L. Xu. Circulating Vitamin E Levels and Risk of Coronary Artery Disease and Myocardial Infarction: A Mendelian Randomization Study. Nutrients 2019; 11(9).
46. C. Fan, T. Huang, X. Kong, X. Zhang, Z. Zou, J. Xiao. Circulating vitamin E and cardiometabolic measures: a Mendelian randomization analysis. J Clin Biochem Nutr 2019; 65(2): 160-9.
47. P. C. Haycock, S. Burgess, K. H. Wade, J. Bowden, C. Relton, G. Davey Smith. Best (but oft-forgotten) practices: the design, analysis, and interpretation of Mendelian randomization studies. Am J Clin Nutr 2016; 103(4): 965-78.
48. S. Burgess, N. M. Davies, S. G. Thompson, on behalf of EPIC-InterAct Consortium. Instrumental Variable Analysis with a Nonlinear Exposure–Outcome Relationship. Epidemiology 2014; 25(6): 877-85.
49. S. Burgess. Sample size and power calculations in Mendelian randomization with a single instrumental variable and a binary outcome. Int J Epidemiol 2014; 43(3): 922-9.
50. A. A. El-Aal, E. A. Abd El-Ghffar, A. A. Ghali, et al. The effect of vitamin C and/or E supplementations on type 2 diabetic adult males under metformin treatment: A single-blinded randomized controlled clinical trial. Diabetes Metab Syndr 2018; 12(4): 483-9.

51. F. Taylor, M. D. Huffman, A. F. Macedo, et al. Statins for the primary prevention of cardiovascular disease. Cochrane Database Syst Rev 2013; (1): Cd004816.
52. T. Varghese, W. M. Schultz, A. A. McCue, et al. Physical activity in the prevention of coronary heart disease: implications for the clinician. Heart 2016; 102(12): 904-9.
53. C. J. Lavie, R. V. Milani, H. O. Ventura. Obesity and cardiovascular disease: risk factor, paradox, and impact of weight loss. J Am Coll Cardiol 2009; 53(21): 1925-32.
54. J. W. Jukema, C. P. Cannon, A. J. de Craen, R. G. Westendorp, S. Trompet. The controversies of statin therapy: weighing the evidence. J Am Coll Cardiol 2012; 60(10): 875-81.

Supplementary materials

Supplemental Table 1 Summary of instruments for circulating dietary-derived antioxidants as both absolute levels and metabolites

Supplemental Table 2 Summary information of the cohorts contributing to the GWAS used for genetic instrumental variables extraction of absolute circulating antioxidants

Supplemental Table 3 Effect estimates of the associations between genetic instrumental variables for absolute circulating dietary derived antioxidants and risk of coronary heart diseases across three databases

Supplemental Table 4 Effect estimates of the associations between genetic instrumental variables for circulating dietary derived antioxidant metabolites and risk of coronary heart diseases across three databases

Supplemental Table 5 Associations between genetically predicted per unit increase in dietary-derived antioxidants and coronary heart disease in Mendelian Randomization analyses

Supplemental Table 6 Associations between genetically predicted per unit (log10) increase in dietary-derived antioxidants metabolites and coronary heart disease in Mendelian Randomization analyses

Supplemental Table 7 The effects of genetic instruments and dietary supplementation on circulating antioxidant concentrations

Supplemental Table 8 Circulating antioxidant concentration from antioxidants supplementation trials

Supplemental Figure 1 Estimated odds ratios for the effect of per unit increase in ln-transformed β-carotene and 1 μg/dL lycopene absolute levels on coronary heart disease, obtained from Weighted-Median estimator (upper panel) and MR Egger regression (lower panel), per outcome database separately and combined over the three databases using fixed-effect model meta-analysis.

Supplemental Figure 2 Sensitivity analysis for β-carotene by using a stringent LD threshold of < 0.001, and for lycopene by restricting the analyses to only GWAS significant-level ($p < 5\times10^{-8}$) SNP only. Analyses were performed per outcome database separately and combined over the three databases using fixed-effect model meta-analysis

The Supplementary materials for this article can be found online at:
https://www.jacc.org/doi/10.1016/j.jacc.2020.10.048

CHAPTER 6

Urinary oxidized, but not enzymatic vitamin E metabolites are inversely associated with measures of glucose homeostasis in middle-aged healthy individuals

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

Clin Nutr. 2021 Jun;40(6):4192-4200

Abstract

Background & aims: Damage induced by lipid peroxidation has been associated with impaired glucose homeostasis. Vitamin E (α-tocopherol, α-TOH) competitively reacts with lipid peroxyl radicals to mitigate oxidative damage, and forms oxidized vitamin E metabolites. Accordingly, we aimed to investigate the associations between α-TOH metabolites (oxidized and enzymatic) in both circulation and urine and measures of glucose homeostasis in the general middle-aged population.

Methods: This cross-sectional study was embedded in the population-based Netherlands Epidemiology of Obesity (NEO) Study. α-TOH metabolites in plasma (α-TOH and α-CEHC-SO_3) and urine [sulfate (SO_3) and glucuronide (GLU) of both α-TLHQ (oxidized) and α-CEHC (enzymatic)] were quantified by liquid chromatography coupled with tandem mass spectrometry (LC/MS-MS). Measures of glucose homeostasis (HOMA-B, HOMA-IR, Insulinogenic index, and Matsuda index) were obtained from fasting and postprandial blood samples. Multivariable linear regression analyses were performed to assess the associations of α-TOH metabolites and measures of glucose homeostasis.

Results: We included 498 participants (45% men) with mean (SD) age of 55.8 (6.1) years who did not use glucose-lowering medication. While plasma α-TOH was not associated with measures of glucose homeostasis, urinary oxidized α-TOH metabolites (α-TLHQ-SO_3/GLU) were associated with HOMA-IR and Matsuda index. For example, a one-SD higher α-TLHQ-SO_3 was associated with 0.92 (95% CI: 0.87, 0.97) fold lower HOMA-IR and 1.06 (1.01, 1.11) fold higher Matsuda index, respectively. Similar results were obtained for the urinary α-TLHQ to α-CEHC ratio as a measure of oxidized-over-enzymatic conversion of α-TOH.

Conclusion: Higher urinary levels of oxidized α-TOH metabolites as well as higher oxidized-to-enzymatic α-TOH metabolite ratio, but not circulating α-TOH or enzymatic metabolites, were associated with lower insulin resistance. Rather than circulating α-TOH, estimates of the conversion of α-TOH might be informative in relation to health and disease.

Introduction

Impaired glucose homeostasis is one of the first steps in the pathogenesis of type 2 diabetes mellitus (T2D). In addition to classical risk factors such as (central) obesity, lipid peroxidation, which is an autocatalytic chain reaction induced by free radicals reacting with lipids, has been implicated to play an important role in the impairment of glucose homeostasis [1].

Vitamin E (α-tocopherol, α-TOH) has been unequivocally demonstrated as an effective lipophilic radical scavenger to prevent the chain propagation in lipid peroxidation [2]. In oxidative catabolism, α-TOH is initially oxidized to α-tocopheroxyl radical by one-electron oxidation, and further reacts with lipid peroxides, opening the chromanol ring and consecutively forming α-tocopheryl quinone (α-TQ), α-tocopheryl hydroquinone (α-THQ), and α-tocopheronic acid. Subsequently, followed by β-oxidation and cyclization of the phytyl side chain, α-tocopherono lactone (α-TLHQ) will be generated (illustrated in our previous review [3]). Alternatively, in the hepatic enzymatic catabolism, α-TOH starts with ω-hydroxylation, and successively shortens the phytyl chain via multiple β-oxidation, generating α-carboxymethyl-hydroxychroman (α-CEHC). Thereafter, α-TLHQ and α-CEHC are both conjugated with glucuronic acid or sulfate to form water-soluble polar metabolites and are excreted mainly via urine. The group of oxidized metabolites α-TLHQ, therefore, reflects the extent of lipid peroxidation and α-TOH bioactivity.

Despite multiple studies having investigated the associations between vitamin E and glucose homeostasis in both observational studies and clinical trials yet with inconsistent results [4-19], studies scarcely addressed the most important aspect of functional α-TOH, i.e. to what extent α-TOH acts as antioxidants. Interestingly, oxidized α-TOH metabolites have been linked to increased risk of multiple diseases with elevated levels of lipid peroxidation [20-22]. Theoretically, impaired glucose homeostasis related to excessive lipid peroxidation is likely to have increased demand for antioxidants protection via α-TOH, and consequently, lead to higher levels of urinary oxidized metabolites. Of note, the only study that specifically explored the association between oxidized metabolites and diabetics identified increased levels of α-TLHQ in diabetic children compared with age- and sex-matched healthy controls [23]. No data so far are available about the association of α-TOH oxidized metabolites with measures of glucose homeostasis in the general population.

We hypothesized that oxidized urinary vitamin E metabolites, but not circulating α-TOH or enzymatic metabolites, would positively associate with worse glucose homeostasis measures. In the present study, we aimed to investigate the associations between α-TOH metabolites and measures of glucose homeostasis in a cross-sectional study of middle-aged healthy individuals embedded in the Netherlands Epidemiology of Obesity study (NEO).

Methods

Study design and Study population

This study was embedded in the population-based prospective Netherlands Epidemiology of Obesity (NEO) study, which is designed to investigate the pathways that are responsible for obesity-related disorders. The NEO study started in 2008 and includes 6,671 individuals aged 45-65 years, with an oversampling of individuals with a self-reported body mass index (BMI) of 27 kg/m^2 or higher. Besides, all inhabitants aged between 45 to 65 years from the municipality of Leiderdorp were invited irrespective of their BMI. The study was approved by the medical ethical committee of the Leiden University Medical Center (LUMC), and all participants gave written informed consent. Detailed information on the study design and data collection have been described previously [24].

Participants were invited to come to the NEO study center of the LUMC for one baseline study visit after an overnight fast. Prior to this study visit, participants collected their urine over 24h and completed a general questionnaire at home in terms of their demographic, lifestyle, and clinical data in addition to specific questionnaires on diet and physical activity. Medication use within one month prior to the visit was asked to bring with participants and was recorded by research nurses. Fasting blood samples were drawn, and within five minutes later, a 400 mL, 600 kcal mixed meal (energy derived from protein, carbohydrate and fat were 16%, 50%, and 34%, respectively) was consumed. Postprandial blood samples were then drawn at 30 and 150 minutes after the meal.

The present study cross-sectionally analyzed the baseline measurements. We included a random subset of 35% of the Leiderdorp participants from Western European ancestry with imaging and genomics information collected (N = 599). In total, 536 participants were eligible with urine collected for at least 20 hours. We excluded participants for the main analyses with: 1) glucose-lowering medication use (n = 12) or no glucose homeostasis measures (n = 1); 2) sample failed to measure urinary metabolites (n = 1); 3) biologically implausible urinary metabolites measures due to sample problem (no TLHQ metabolites, n = 2); outliers (see 2.5. Statistical analysis) on plasma (α-TOH) or urinary metabolites (n = 9); 3) missing data (n = 1) or outliers on HOMA-B or HOMA-IR (n = 1); 4) missing data on potential confounders including physical activity, Dutch health diet index, education level and body fat percentage (n = 11, 1 overlap with vitamin E metabolites outliers). Consequently, a total of 498 individuals were used in the fasting analyses. In addition, we excluded individuals with (1) uncompleted mixed meal challenge (n = 33); (2) missing data on Matsuda index or Insulinogenic Index (n = 20, 3 overlapped with uncompleted mix meal), leaving a total of 448 participants in the postprandial analyses. Summary of the participant exclusions is presented in **Supplementary Figure 1**.

Alpha-tocopherol metabolites measures

Alpha-tocopherol (metabolite) measurements in plasma

In fasting plasma samples, α-TOH and α-CEHC sulfate conjugates (α-CEHC-SO_3) were detected and quantified by untargeted metabolomics provider Metabolon, Inc. (Durham, NC, USA) on a platform encompassing four liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS) methods (LC-MS/MS negative, LC-MS/MS positive early, LC-MS/MS positive late and LC-MS/MS polar). More detailed descriptions have been described previously [25,26].

Alpha-tocopherol metabolite measurements in urine

Urinary oxidized α-TOH metabolites (α-TLHQ) and enzymatic metabolites (α-CEHC), presented as their sulfate or glucuronide conjugates (α-TLHQ-SO_3, α-TLHQ-GLU, α-CEHC-SO_3, α-CEHC-GLU), were measured by LC-MS/MS at the University College London, UK.

Prior to the measurement, urine samples were thawed, and 100µl fresh urine was then centrifuged in Eppendorf tubes at 14 000 g for 10 min 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. 10 µl 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 carryover, 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 a blank sample (H_2O: 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. The different elution time (min) for internal standards (LA 4.33, AD4 2.7) and each metabolite (2.39, 2.12 and 2.29 for α-TLHQ sulfate, glucuronide minor and major, 2.64, 2.50, 2.56 for α-CEHC sulfate, glucuronide minor and major) guaranteed that all metabolites could be separated in a single chromatographic run.

After separation, the metabolites were then analyzed by MS using a Waters ACQUITY UPLC coupled to a triple-quadrupole Xevo TQ-S fitted with electrospray ionization in negative ion mode. The gas temperatures persisted at 600°C for desolvation. In addition, nitrogen was used as the nebulizing gas with a 7.0 bar. α-CEHC and α-TLHQ are isobaric because of the same molecular mass ($C_{15}H_{22}O_4$). 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 the following collision activated dissociation (CAD) with argon. These ions were then used to quantify each α-TOH metabolite from transitions previously established by Sharma et al. [27] (glucuronide conjugates 453.3>113.0 m/z and sulfate conjugates, 357.1>79.9 m/z) that corresponded to their theoretical molecular masses.

Urinary creatinine concentrations (mmol/L) were measured to correct 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 [23,28].

The final concentrations 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 as the α-TLHQ-to-α-CEHC ratio, α-TLHQ-GLU-to-α-CEHC-GLU ratio, and α-TLHQ-SO_3-to-α-CEHC-SO_3 ratio.

Glucose homeostasis measures

For blood samples, plasma glucose concentrations were obtained by enzymatic and colorimetric methods (Roche Modular Analytics P800, Roche Diagnostics, Mannheim, Germany; CV < 5%), while serum insulin concentrations were detected via an immunometric method (Siemens Immulite 2500, Siemens Healthcare Diagnostics, Breda, The Netherlands; CV < 5%). All these measurements were performed in the central clinical chemistry laboratory of the LUMC. Homeostatic model of insulin resistance (HOMA-IR), a marker of hepatic insulin resistance, homeostatic model of and β-cell function (HOMA-B), a measure of β-cell to glucose-stimulated insulin secretion were then calculated by using the formula (fasting glucose × fasting insulin)/22.5 and (20 × fasting insulin)/(fasting glucose - 3.5) respectively [29,30]. Matsuda index, which represents both hepatic and peripheral tissue sensitivity to insulin, was generated as $10000/\sqrt{}$ (fasting glucose × fasting insulin) ($glucose_{mean(0\text{-}150min)}$ × $insulin_{mean(0\text{-}150min)}$) [31]. Moreover, Insulinogenic Index also reflects β-cell function but specifically, the first-phase insulin response to glucose challenge was calculated with the formula ($insulin_{30min}$ - fasting insulin)/($glucose_{30min}$ - fasting glucose) [32].

Covariates

Education level was grouped into high (including higher vocational school, university, and postgraduate education) and low based on the Dutch Education system. Smoking habits were reported in three categories: current smoker, for-

mer smoker and never smoke. Familial diabetes history (yes/no) was collected through a general questionnaire and defined as having a father, mother, or brother or sister with a diagnosis of diabetes mellitus. The frequency and duration of leisure physical activity over the past 4 weeks were reported by participants on the Short Questionnaire to Assess Health-enhancing physical activity (SQUASH), which is expressed as metabolic equivalent (MET-hours per week). A semiquantitative food frequency questionnaire was used to assess food and beverage intake, and total energy intake. Dutch Healthy Diet Index (DHD-index) was then calculated based on dietary intake [33]. Percentage body fat was measured by Bio Impedance Balance (TBF-310, Tanita International Division, UK). Total cholesterol levels (mmol/L) were obtained in fasting blood samples in our central clinical chemistry laboratory using standard assays. Lipid-lowering medication defined as the use of statins fibrates and other lipid-lowering medication was recorded by medicine inventory by the research nurses. Information on the use of vitamin E supplements was collected through questionnaires as the combination of vitamin E supplements only or the use of multivitamin supplements.

Statistical analysis

Main analysis

Descriptive characteristics at the NEO baseline were presented as mean (standard deviation, SD) for normally distributed variables, median (interquartile range) for skewed variables, and frequency (proportions) for categorical variables. As the distribution of both plasma and urinary vitamin E metabolites and the measures of glucose homeostasis (HOMA-B, HOMA-IR, Insulinogenic Index, and Matsuda index) were skewed, these were natural log-transformed. Outliers (defined as located more than 4 SD from the mean) were removed after natural log transformation, comprising approximately 2% of the total observations.

In order to explore the α-TOH metabolism process, pairwise Pearson correlations were performed between circulatory and urinary metabolites. Multivariable linear regression analyses were performed to examine the association between α-TOH metabolites (determinant) and measures of glucose homeostasis (outcome). To facilitate interpretation and comparison between measures of glucose homeostasis, determinants were then z-transformed (mean = 0, SD = 1), so that the regression coefficient with its corresponding 95% confidence interval (CI) can be interpreted as the fold difference in the outcome with respect to a one-SD change in the determinant. For each outcome and determinant, four models were fitted. The basic regression model was adjusted for age and sex (Model 1). Model 2 was additionally adjusted for potential confounders based on biological knowledge and previous studies including educational level (high or low), familial history of diabetes (yes or no), physical activity (MET-hours per week), smoking habits (never smoke, current smoker or former smoker), Dutch Healthy Diet Index and total energy intake (kJ/day). Total cholesterol level (mmol/L) and lipid-lowering medication use (yes/no) were also included in Model 2 in models on the fasting levels for plasma α-TOH only since the transportation and uptake of α-TOH are largely dependent on lipoproteins [34]. Of the lipoproteins, it was shown before that specifically (but not exclusively) LDL cholesterol is acting as α-TOH transporter [35]. Therefore, we performed additional analyses with adjustment of LDL-cholesterol only in Model 2(A), LDL, and HDL-cholesterol in Model

2(B) instead of the adjustment of total cholesterol levels. Moreover, obesity has been shown to directly have an influence on glucose homeostasis, especially insulin resistance [36], and meanwhile, obesity also increases lipid peroxidation levels [37] and might thus affect the conversion of α-TOH in the body. Therefore, we included total body fat (Model 3) as an indicator of fat content to explore the underlying mechanisms.

Sensitivity analyses

Urine sample collection missingness and other exclusions

To evaluate whether the current study population is representative of the total study population, the study characteristics among individuals of urine sample collection missingness and other exclusion reasons were compared with the included participants separately. Chi-square test was used for categorical variables, while t-test was used for normally distributed numeric variables and Mann-Whitely U test was used for non-normally distributed numeric variables.

Vitamin E supplement use

Vitamin E supplement use might have an influence on measures of glucose homeostasis as well as vitamin E conversion in the body. Vitamin E supplement use in the current study was defined as either the use of vitamin E supplements only or the use of multiple vitamin supplements. In order to fully rule out the supplement effect, we further performed sensitivity analyses in participants who did not use vitamin E and/or multivitamin supplements.

Censored normal regression

In the main analysis, we excluded all participants with glucose-lowering medication use. However, exclusion might underestimate the effect size of the determinants, reduce power, and could introduce (collider-stratification) bias in the analyses. A reliable approach to correct this dilution is to use censored normal regression model analysis which has been established previously [38]. Therefore, we additionally performed censored normal regression model analysis for the final adjusted model (Model 3) as in the main analysis.

All the analyses were undertaken using R (v3.6.1) statistical software (The R Foundation for Statistical Computing, Vienna, Austria) (https://www.r-project.org/). Pairwise correlation plots were performed using the R package “GGally” (https://CRAN.R-project.org/package=GGally) 39, and the censored normal linear regression models were performed using the R package “survival” (https://CRAN.R-project.org/package=survival) 40.

Results

Characteristics of the study population

A total of 498 participants (45% men) were analyzed, of whom 448 were available for postprandial analyses. The baseline characteristics of the study population are presented in **Table 1**.

Table 1 Characteristics the study population[1]

	N = 498
Demographics	
Age (years)	55.8 (6.1)
Sex (male)	225 (45.2%)
BMI (kg/m^2)	25.4 (23.1, 27.8)
Total body fat (%)	30.9 (8.3)
Education level (high)	253 (50.8%)
Family history of diabetes (yes)	136 (27.3%)
Lifestyle factors	
Dutch healthy diet index	59.8 (8.4)
Energy intake (KJ/day)	9106 (7304, 11053)
Physical activity a (MET-h/week)	29.4 (16.5, 48.8)
Smoking	
Current	53 (10.6%)
Former	235 (47.2%)
Never	210 (42.2%)
Vitamin E supplement use (yes)[2]	128 (25.7%)
Lipid-lowering medication use (yes)	28 (5.6%)
Vitamin E metabolites measurements	
Blood (log-transformed, no unit)[3]	
α-tocopherol	19.7 (0.2)
α-CEHC-SO$_3$[4]	12.1 (0.6)
Urinary	
α-TLHQ-SO$_3$ (nmol/mmol creatinine)	2.6 (1.6, 4.1)
α-TLHQ-GLU (nmol/mmol creatinine)	1822.5 (1339.3, 2744.8)
α-CEHC-SO$_3$ (nmol/mmol creatinine)	165.4 (97.4, 298.3)
α-CEHC-GLU (nmol/mmol creatinine)	91.0 (62.3,139.4)
α-TLHQ/α-CEHC	7.1 (5.1, 9.9)
α-TLHQ-SO$_3$/α-CEHC-SO$_3$	0.016 (0.011, 0.024)
α-TLHQ-GLU/α-CEHC-GLU	20.8 (15.4, 28.5)
Urinary (log-transformed)	
α-TLHQ-SO$_3$	1.0 (0.7)
α-TLHQ-GLU	7.6 (0.6)
α-CEHC-SO$_3$	5.1 (0.8)
α-CEHC-GLU	4.5 (0.6)
α-TLHQ/α-CEHC	1.9 (0.5)
α-TLHQ-SO$_3$/α-CEHC-SO$_3$	-4.1 (0.7)
α-TLHQ-GLU/α-CEHC-GLU	3.0 (0.4)
Glucose homeostasis	
Fasting levels	
HOMA-IR	1.9 (1.3, 2.9)
HOMA-B (%)	89.7 (63.8, 129.6)
Postprandial response[5]	
Matsuda Index	5.6 (3.8, 7.9)
Insulinogenic index	7.2 (5.7, 9.4)

[1]Data are presented as median (interquartile range) for numeric variables, and number (proportions) for categorical variables. [2] Vitamin E supplement use was defined as either vitamin E supplement use or multiple vitamin supplement use. [3] The concentrations were measured by Metabolon platform, and no units were presented. [4] Only complete cases in metabolon platform (n = 263). [5] Participants with completed meal challenge and no missing data on postprandial response were used (N = 448).

Table 2 Associations between circulatory α-tocopherol and measures of glucose homeostasis in the general population[1]

	Fasting measures		Postprandial responses	
	HOMA-B	HOMA-IR	Insulinogenic index	Matsuda Index
Model 1	0.95 (0.91, 1.00)	0.97 (0.92, 1.03)	1.00 (0.96, 1.03)	0.99 (0.94, 1.04)
Model 2	0.98 (0.92, 1.04)	0.98 (0.91, 1.05)	1.01 (0.96, 1.05)	1.00 (0.94, 1.07)
Model 2 (A)	0.96 (0.91, 1.01)	0.96 (0.90, 1.03)	1.00 (0.96, 1.04)	1.01 (0.95, 1.07)
Model 2 (B)	0.97 (0.92, 1.02)	0.98 (0.92, 1.04)	1.00 (0.96, 1.05)	1.00 (0.94, 1.06)
Model 3	0.97 (0.92, 1.03)	0.97 (0.91, 1.03)	1.01 (0.96, 1.05)	1.01 (0.96, 1.06)
Model 3 (A)	0.96 (0.91, 1.01)	0.97 (0.91, 1.02)	1.00 (0.96, 1.04)	1.00 (0.95, 1.05)
Model 3 (B)	0.97 (0.92, 1.02)	0.97 (0.92, 1.03)	1.00 (0.96, 1.05)	1.00 (0.95, 1.05)

[1] In total, 498 participants with available α-tocopherol were used for fasting HOMA-B and HOMA-IR analysis while 448 for postprandial Insulinogenic index and Matsuda Index analysis. Results are derived from linear regression coefficients with 95% confidence interval (CI) and were expressed as a one-SD change in blood α-tocopherol with corresponding fold difference in log-transformed glucose homeostasis traits. Model 1: age and sex. Model 2: Model 1 + education level (high or low), familial history of diabetes (yes or no), physical activity (MET hours per week), smoking habits (never smoke, current smoker or former smoker), Dutch Healthy Diet index (DHD-index), energy intake (KJ/day), total cholesterol level (mmol/L) and lipid lowering medication (yes or no). Model 2 (A): Model 1 + education level (high or low), familial history of diabetes (yes or no), physical activity (MET hours per week), smoking habits (never smoke, current smoker or former smoker), Dutch Healthy Diet index (DHD-index), energy intake (KJ/day), LDL cholesterol level (mmol/L) and lipid lowering medication (yes or no). Model 2 (B): Model 1 + education level (high or low), familial history of diabetes (yes or no), physical activity (MET hours per week), smoking habits (never smoke, current smoker or former smoker), Dutch Healthy Diet index (DHD-index), energy intake (KJ/day), LDL and HDL cholesterol level (mmol/L) and lipid lowering medication (yes or no). Model 3: Model 2 + body fat percentage (%). Model 3 (A): Model 2 (A) + body fat percentage (%). Model 3 (B): Model 2 (B) + body fat percentage (%).

Main analysis

Pairwise correlation between plasma and urinary metabolites

Correlations between circulatory and urinary α-TOH metabolites are shown in **Figure 1**. In all participants, plasma α-TOH was weakly associated with any urinary metabolite (r: 0.18 - 0.23), while plasma α-CEHC-SO_3 was moderately correlated with urinary metabolites (r: 0.31 - 0.49,) with the highest correlation with urinary α-CEHC-SO_3 (r = 0.49). Urinary metabolites were highly correlated with each other with r ranging from 0.44 to 0.73.

Circulatory α-TOH and glucose homeostasis measures

In plasma, we did not find evidence that α-TOH was associated with measures of glucose homeostasis after fully adjusting for potential confounders, as shown in **Table 2**.

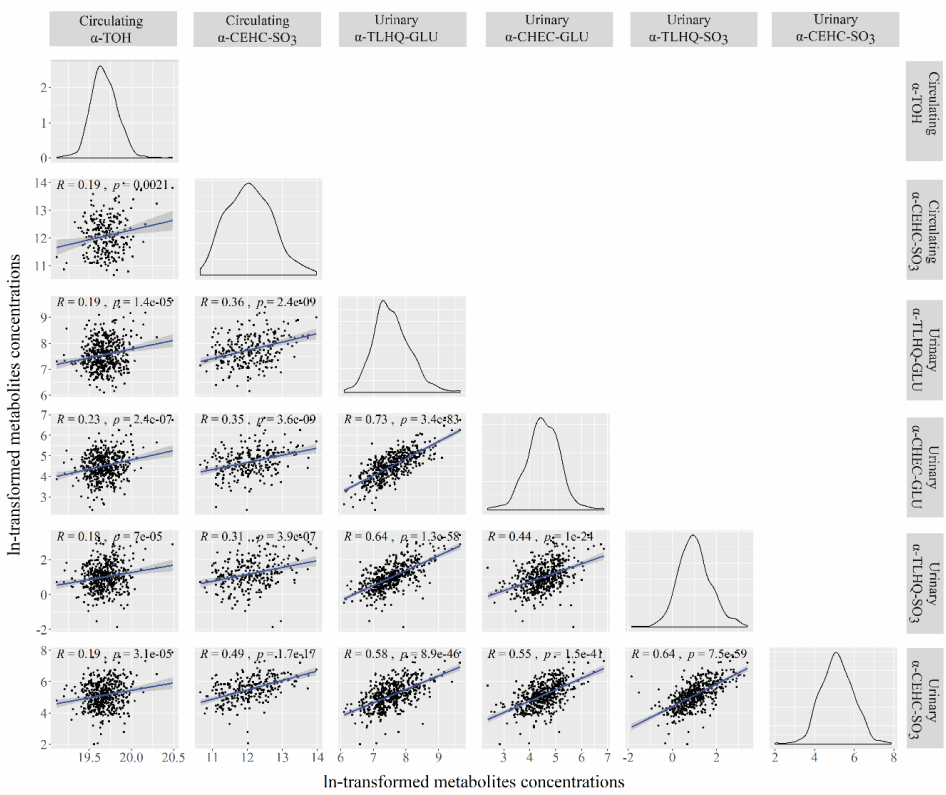


Figure 1 Pairwise correlation

Pairwise correlation of α-tocopherol metabolites in blood and urine (N = 498). Pearson correlations were calculated after natural log-transformation of metabolites. For blood α-CEHC-SO_3, only the complete cases in the metabolon platform were used (n = 263).

Urinary metabolites and glucose homeostasis measures

The urinary oxidized metabolites α-TLHQ-SO_3 and α-TLHQ-GLU were associated with HOMA-IR and Matsuda index as visualized in **Figure 2** (summary statistics in **Supplementary Table 2**). For fasting measures (**Figure 2A** and **2B**), in the age- and sex- adjusted basic model (Model 1), a one-SD higher α-TLHQ-SO_3 was associated with 0.93 [95% confidence interval (CI): 0.89, 0.98] fold lower HOMA-B and 0.89 (95% CI: 0.83, 0.94) fold lower HOMA-IR, respectively. However, when we additionally adjusted for other potential confounders including lifestyle factors and adiposity measures (Model 3), the association observed for HOMA-B diminished towards null and α-TLHQ-SO_3 was only associated with HOMA-IR (β: 0.92, 95% CI: 0.87, 0.97). Similarly, α-TLHQ-GLU followed the same pattern, and a one-SD higher α-TLHQ-GLU was associated with 0.93 (95% CI: 0.88, 0.98) fold lower HOMA-IR in the fully adjusted Model 3. This pattern was also observed in postprandial analyses, and α-TLHQ-SO_3 and α-TLHQ-GLU were only associated with Matsuda index (**Figure 2D**), with a one-SD higher α-TLHQ-SO_3 corresponding to 1.06 (95% CI: 1.01, 1.11) fold higher Matsuda

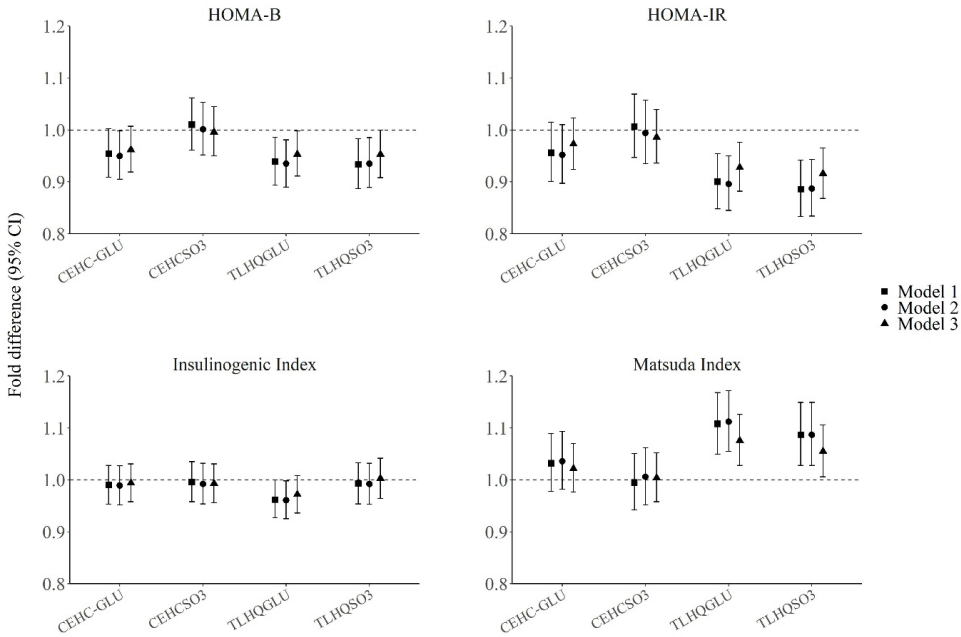


Figure 2 Association between urinary metabolites and glucose homeostasis measurements

Association between urinary metabolites and (A) HOMA-B, (B) HOMA-IR, (C) Insulinogenic Index, and (D) Matsuda Index in general population. Results were based on general population (N = 498 for HOMA-B and HOMA-IR; N = 448 for Insulinogenic Index and Matsuda Index), and were derived from multivariable linear regression coefficient with corresponding 95% confidence interval (CI) and were expressed as fold difference in log-transformed outcomes with a one-standard deviation change in determinants. Model 1: adjust for age and sex. Model 2: Model 1 + educational level (high or low), familial history of diabetes (yes or no), physical activity (MET hours per week), smoking habits (never smoke, current smoker or former smoker), Dutch Healthy Diet Index and total energy intake (KJ/day). Model 3: Model 2 + total body fat percentage (%).

index, and a one-SD higher α-TLHQ-GLU corresponding to 1.08 (95% CI: 1.03, 1.13) fold higher Matsuda index after adjustment for all potential confounders (Model 3). However, no association was found between enzymatic metabolites and any measures of glucose homeostasis.

Higher oxidized-to-enzymatic metabolite ratios showed similar associations as oxidized metabolites with measures of glucose homeostasis (**Figure 3** and summary statistics in **Supplementary Table 3**). For example, in the basic

model (Model 1), a one-SD higher α-TLHQ-to-α-CEHC ratio was associated with 0.90 (95% CI: 0.85, 0.95) fold lower HOMA-IR (**Figure 3B**) and 1.11 (1.05, 1.17) fold higher Matsuda index (**Figure 3D**). After adjustment for potential confounding factors (Model 3A), the association was attenuated, and a one-SD higher α-TLHQ-to-α-CEHC ratio was associated with 0.94 (95% CI: 0.89, 0.99) fold lower HOMA-IR and 1.07 (95% CI: 1.02, 1.12) fold higher Matsuda index. Simi-

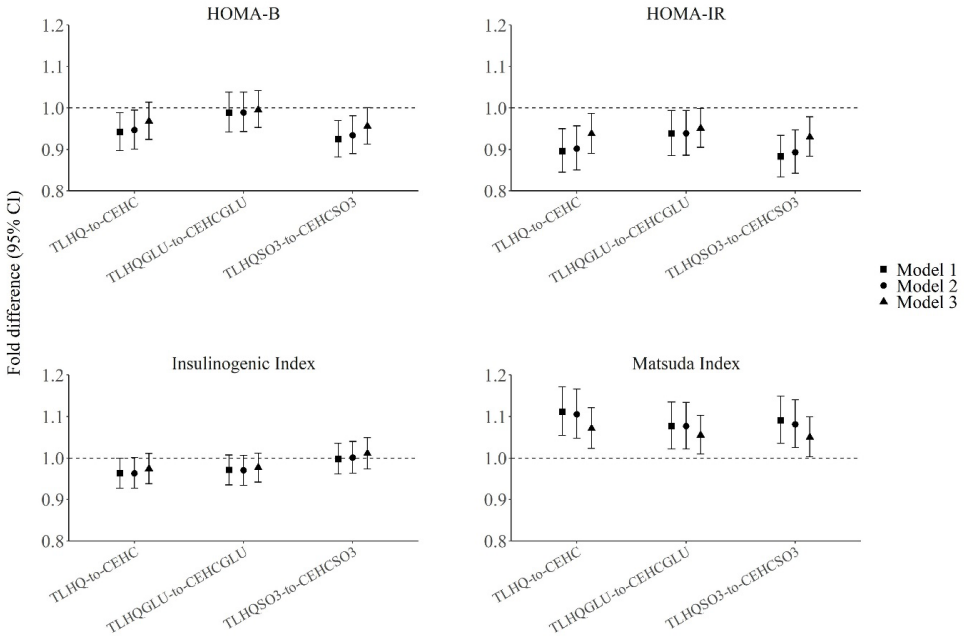


Figure 3 Association between urinary enzymatic-to-oxidized metabolite ratios and glucose homeostasis measurements

Association between urinary metabolite ratios and (A) HOMA-B, (B) HOMA-IR, (C) Insulinogenic Index, and (D) Matsuda Index in general population. Results were based on general population (N = 498 for HOMA-B and HOMA-IR; N = 448 for Insulinogenic Index and Matsuda Index), and were derived from multivariable linear regression coefficient with corresponding 95% confidence interval and were expressed as fold difference in log-transformed outcomes with a one-standard deviation change in determinants. Model 1: adjust for age and sex. Model 2: Model 1 + educational level (high or low), familial history of diabetes (yes or no), physical activity (MET hours per week), smoking habits (never smoke, current smoker or former smoker), Dutch Healthy Diet Index and total energy intake (KJ/day). Model 3A: Model 2 + total body fat percentage (%).

lar associations were observed for the sulfate conjugate ratio and glucuronide conjugate ratio.

Sensitivity analysis

Urine sample collection missingness and other exclusions

Basic characteristics were compared between included participants (N = 498), individuals with less than 20 hours or no urine collection (N = 63) and excluded participants due to other reasons (N = 38), as presented in **Supplementary Table 1**. The participants excluded due to other reasons had higher percentages of family history of diabetes (27.3%, 30.2% versus 47.4%) and lipid-lowering medication use (5.6%, 11.1% versus 29.0%). Moreover, this group had slightly higher age and more male participants.

Vitamin E supplementation
Participants with vitamin E supplementation (n = 128) were further excluded in this sensitivity analysis, leaving 370 individuals for fasting and 337 for postprandial analyses, respectively. The correlation between α-TOH and other metabolites remained low. Plasma α-CEHC-SO_3 was moderately correlated to urinary α-CEHC-SO_3 (r = 0.41) only (**Supplementary Figure 2**). In addition, effect sizes did not change materially in the multivariable regressions after adjusting for confounders compared with the full population analyses (**Supplementary Table 4, Figure 3 and 4**).

Censored normal regression
In censored normal regression models with 12 individuals taking glucose-lowering medication included, 510 participants with complete data were included for fasting analysis, and 460 participants for postprandial analysis. We did not observe substantial differences in Model 3 (Supplementary Table 5) compared with the main analysis.

Discussion
In the present cross-sectional study, we aimed to investigate the association between metabolites of vitamin E (α-TOH) and measures of glucose homeostasis in both fasting and postprandial state. When adjusted for potential confounders, we did not find evidence supporting an association between plasma α-TOH with measures of glucose homeostasis. However, we observed that higher urinary oxidized metabolites of α-TOH, as well as oxidized-to-enzymatic metabolite α-TOH ratios, were associated with reduced insulin resistance. Additional adjustment for total body fat only minimally explained these observations. Sensitivity analysis including restriction to individuals not taking vitamin E or multivitamin supplements and censored normal regression model provided similar results emphasizing the robustness of the observations.

The lack of evidence supporting an association between plasma α-TOH with measures of glucose homeostasis observed in the present study is in accordance with some [8-12], but not all previous observational studies [6,7]. The discrepancies might due to several factors such as study design, sample size as well as used confounders. In line with our study, in a more than 20 years follow-up study of middle-aged Swedish men, plasma α-TOH was not associated with future insulin response or T2D incidence [8]. Interestingly, in some clinical trials, α-TOH supplement only, which is mostly corresponding to an increase of circulating α-TOH [41-43], was not beneficial for the improvement of glucose levels, lipid levels, or insulin sensitivity [17-19]. However, this may be explained by factors such as dosage, timing, duration, and type of vitamin E.

With respect to oxidized metabolites, contrary to our hypotheses, we found associations between higher oxidized metabolites and better measures of glucose homeostasis. Previously, a study particularly identified that urinary α-TLHQ was higher in children with type 1 diabetes mellitus than in healthy controls [23]. Participants with metabolic syndrome also had decreased concentration of uri-

nary α-CEHC due to increased oxidative stress levels and inflammation in spite of rather similar plasma α-TOH concentration compared with healthy individuals, thereby indicating a higher antioxidative demand for α-TOH [44]. Moreover, metabolic syndrome patients had approximately 12% greater static oxidation reduction potential, i.e. oxidants, and 59% lower readily available antioxidant reserves compared with healthy adults [45]. An explanation for the inconsistency identified in our study may lie in the bioavailability and elimination of α-TOH. The bioavailability was shown to be reduced and the elimination delayed in metabolic syndrome patients who have increased lipid peroxidation, observed as lower plasma and urinary α-CEHC, independent of the co-ingested dairy fat amount [46]. In disease conditions where oxidants outweigh antioxidants, more antioxidants are required to diminish the damage caused by oxidative stress. Together with the decreased hepatic turnover, the conversion of α-TOH may shift to a preference for non-enzymatic oxidation. However, in our relatively healthy population, lipid oxidation level is assumed to be relatively low and antioxidants outweigh oxidants with no delayed enzymatic conversion. In addition, other antioxidative systems still have the potential to neutralize oxidants, therefore, the scavenging function of α-TOH might be compensated. Furthermore, the excretion of these metabolites may also alter in different health conditions, and a higher/lower excretion via bile or feces may occur.

In the pairwise correlation analysis, the moderate correlation of α-CEHC-SO_3 between plasma and urine is a validation of the measurement. However, a very weak correlation was found between plasma α-TOH and oxidized metabolites indicating that the excretion of oxidized metabolites does not increase with the increase of circulating α-TOH. This raises the argument that despite of an increased level of circulating α-TOH, the body does not fully make use of this α-TOH as antioxidant. Taken together, the circulating level of α-TOH may not reflect the antioxidative capacity and may not affect glucose homeostasis.

One of the strengths of the present study is the general population-based setting and the various confounding factors considered. Former studies of oxidized α-TOH metabolites and health outcomes have been conducted in relatively small patient cohorts, limiting the generalization of the results. Another strength is the measurement of oxidized metabolites. Previous chromatography-mass spectrometry (GC-MS) based method required long sample preparation of deconjugation, extraction, and derivatization, and only detected free unconjugated metabolites [47,48]. These may result in artefactual oxidation products of α-CEHC during sample preparation, α-TOH acid, and α-tocopherono lactone, better known as Simon metabolites [49,50]. The LC-MS/MS based method used in our study was developed and validated previously to have solid reliability and reproducibility [27,28], and the intact conjugate with minimal preparation ensures that the metabolites are unlikely due to artifact formation [27].

There are several limitations in this study. Firstly, a considerable number of individuals had no urine collection for at least 20 hours, however, study characteristics were comparable with our study population. Secondly, since multiple urinary metabolites and study outcomes were used in the analyses, there might be multiple testing and a chance of false-positive results. However, our exposures and outcomes were highly intercorrelated, and conventional corrections

for multiple testing (e.g, Bonferonni) are too conservative. Thirdly, the habitual dietary intake of all participants was estimated by using a self-administered, semiquantitative 125-item food frequency questionnaire (FFQ). As this FFQ did not contain complete information on certain food items (micronutrients), we had very limited information to calculate dietary intake of alpha-tocopherol, therefore, we are not able to perform any analysis based on intake of alpha-tocopherol. Lastly, because of the cross-sectional design, we are not able to rule out residual confounding or reverse causation.

In conclusion, the present study suggests that, in the middle-aged population, circulating α-TOH is not associated with glucose homeostasis measures. However, higher amounts of urinary oxidized α-TOH metabolites and oxidized-to-enzymatic metabolite ratios are associated with lower insulin resistance. This finding supports the hypothesis that, rather than circulating α-TOH, its conversion might be more informative in relation to health and diseases. These findings also highlight the importance of disentangling the conversion preference of α-TOH in different health states in future studies.

References

1. S. Sasson. Nutrient overload, lipid peroxidation and pancreatic beta cell function. Free Radic Biol Med 2017; 111: 102-9.
2. E. Niki. Role of vitamin E as a lipid-soluble peroxyl radical scavenger: in vitro and in vivo evidence. Free Radic Biol Med 2014; 66: 3-12.
3. J. Luo, K. Mills, S. le Cessie, R. Noordam, D. van Heemst. Ageing, age-related diseases and oxidative stress: What to do next? Ageing Res Rev 2020; 57: 100982.
4. E. S. Eshak, H. Iso, I. Muraki, A. Tamakoshi. Fat-soluble vitamins from diet in relation to risk of type 2 diabetes mellitus in Japanese population. Br J Nutr 2018: 1-18.
5. C. Zhou, L. Na, R. Shan, et al. Dietary Vitamin C Intake Reduces the Risk of Type 2 Diabetes in Chinese Adults: HOMA-IR and T-AOC as Potential Mediators. PLoS One 2016; 11(9): e0163571.
6. E. J. Mayer-Davis, T. Costacou, I. King, D. J. Zaccaro, R. A. Bell. Plasma and dietary vitamin E in relation to incidence of type 2 diabetes: The Insulin Resistance and Atherosclerosis Study (IRAS). Diabetes Care 2002; 25(12): 2172-7.
7. T. Costacou, B. Ma, I. B. King, E. J. Mayer-Davis. Plasma and dietary vitamin E in relation to insulin secretion and sensitivity. Diabetes Obes Metab 2008; 10(3): 223-8.
8. J. Arnlov, B. Zethelius, U. Riserus, et al. Serum and dietary beta-carotene and alpha-tocopherol and incidence of type 2 diabetes mellitus in a community-based study of Swedish men: report from the Uppsala Longitudinal Study of Adult Men (ULSAM) study. Diabetologia 2009; 52(1): 97-105.
9. A. Reunanen, P. Knekt, R. K. Aaran, A. Aromaa. Serum antioxidants and risk of non-insulin dependent diabetes mellitus. Eur J Clin Nutr 1998; 52(2): 89-93.
10. H. Z. Pan, L. Zhang, M. Y. Guo, et al. The oxidative stress status in diabetes mellitus and diabetic nephropathy. Acta Diabetol 2010; 47 Suppl 1: 71-6.
11. L. A. Da Costa, B. Garcia-Bailo, C. H. Borchers, A. Badawi, A. El-Sohemy. Association between the plasma proteome and plasma alpha-tocopherol concentrations in humans. J Nutr Biochem 2013; 24(1): 396-400.
12. I. Mehmetoglu, F. H. Yerlikaya, S. Kurban. Correlation between vitamin A, E, coenzyme Q(10) and degree of insulin resistance in obese and non-obese subjects. J Clin Biochem Nutr 2011; 49(3): 159-63.
13. G. Paolisso, A. D'Amore, D. Galzerano, et al. Daily vitamin E supplements improve metabolic control but not insulin secretion in elderly type II diabetic patients. Diabetes Care 1993; 16(11): 1433-7.
14. G. Paolisso, A. D'Amore, D. Giugliano, A. Ceriello, M. Varricchio, F. D'Onofrio. Pharmacologic doses of vitamin E improve insulin action in healthy subjects and non-insulin-dependent diabetic patients. Am J Clin Nutr 1993; 57(5): 650-6.
15. A. Sharma, S. Kharb, S. N. Chugh, R. Kakkar, G. P. Singh. Evaluation of oxidative stress before and after control of glycemia and after vitamin E supplementation in diabetic patients. Metabolism 2000; 49(2): 160-2.

16. M. Boshtam, M. Rafiei, I. D. Golshadi, M. Ani, Z. Shirani, M. Rostamshirazi. Long term effects of oral vitamin E supplement in type II diabetic patients. Int J Vitam Nutr Res 2005; 75(5): 341-6.
17. A. M. de Oliveira, P. H. Rondo, L. A. Luzia, F. H. D'Abronzo, V. K. Illison. The effects of lipoic acid and alpha-tocopherol supplementation on the lipid profile and insulin sensitivity of patients with type 2 diabetes mellitus: a randomized, double-blind, placebo-controlled trial. Diabetes Res Clin Pract 2011; 92(2): 253-60.
18. N. Suksomboon, N. Poolsup, S. Sinprasert. Effects of vitamin E supplementation on glycaemic control in type 2 diabetes: systematic review of randomized controlled trials. J Clin Pharm Ther 2011; 36(1): 53-63.
19. S. Shab-Bidar, Z. Mazloum, Z. Mousavi-Shirazifard. Daily vitamin E supplementation does not improve metabolic and glycemic control in type 2 diabetic patients: a double blinded randomized controlled trial. J Diabetes 2013; 5(1): 57-8.
20. P. Torquato, D. Bartolini, D. Giusepponi, 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: 115-25.
21. G. Ravaglia, P. Forti, A. Lucicesare, et al. Plasma tocopherols and risk of cognitive impairment in an elderly Italian cohort. Am J Clin Nutr 2008; 87(5): 1306-13.
22. M. Casati, V. Boccardi, E. Ferri, et al. Vitamin E and Alzheimer's disease: the mediating role of cellular aging. Aging Clin Exp Res 2019.
23. G. Sharma, D. P. Muller, S. M. O'Riordan, et al. Urinary conjugated alpha-tocopheronolactone--a biomarker of oxidative stress in children with type 1 diabetes. Free Radic Biol Med 2013; 55: 54-62.
24. R. de Mutsert, M. den Heijer, T. J. Rabelink, et al. The Netherlands Epidemiology of Obesity (NEO) study: study design and data collection. Eur J Epidemiol 2013; 28(6): 513-23.
25. E. T. Cirulli, L. Guo, C. Leon Swisher, et al. Profound Perturbation of the Metabolome in Obesity Is Associated with Health Risk. Cell Metab 2019; 29(2): 488-500.e2.
26. T. Long, M. Hicks, H. C. Yu, et al. Whole-genome sequencing identifies common-to-rare variants associated with human blood metabolites. Nat Genet 2017; 49(4): 568-78.
27. G. Sharma, D. P. Muller, S. M. O'Riordan, et al. Urinary conjugated α-tocopheronolactone--a biomarker of oxidative stress in children with type 1 diabetes. Free Radic Biol Med 2013; 55: 54-62.
28. G. Sharma, D. Muller, S. O'Riordan, 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): 599-600.
29. D. R. Matthews, J. P. Hosker, A. S. Rudenski, B. A. Naylor, D. F. Treacher, R. C. Turner. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. Diabetologia 1985; 28(7): 412-9.
30. T. M. Wallace, J. C. Levy, D. R. Matthews. Use and abuse of HOMA modeling. Diabetes Care 2004; 27(6): 1487-95.
31. M. Matsuda, R. A. DeFronzo. Insulin sensitivity indices obtained from oral glucose tolerance testing: comparison with the euglycemic insulin clamp. Diabetes Care 1999; 22(9): 1462-70.
32. B. Singh, A. Saxena. Surrogate markers of insulin resistance: A review. World J Diabetes 2010; 1(2): 36-47.

33. M. Looman, E. J. Feskens, M. de Rijk, et al. Development and evaluation of the Dutch Healthy Diet index 2015. Public Health Nutr 2017; 20(13): 2289-99.
34. D. I. Thurnham, J. A. Davies, B. J. Crump, R. D. Situnayake, M. Davis. The use of different lipids to express serum tocopherol: lipid ratios for the measurement of vitamin E status. Ann Clin Biochem 1986; 23 (Pt 5): 514-20.
35. L. Schmölz, M. Birringer, S. Lorkowski, M. Wallert. Complexity of vitamin E metabolism. World J Biol Chem 2016; 7(1): 14-43.
36. S. E. Kahn, R. L. Hull, K. M. Utzschneider. Mechanisms linking obesity to insulin resistance and type 2 diabetes. Nature 2006; 444(7121): 840-6.
37. L. Marseglia, S. Manti, G. D'Angelo, et al. Oxidative stress in obesity: a critical component in human diseases. Int J Mol Sci 2014; 16(1): 378-400.
38. M. D. Tobin, N. A. Sheehan, K. J. Scurrah, P. R. Burton. Adjusting for treatment effects in studies of quantitative traits: antihypertensive therapy and systolic blood pressure. Stat Med 2005; 24(19): 2911-35.
39. B. Schloerke, J. Crowley, D. Cook, et al. Ggally: extension to ggplot2. (R package version 1.4). 2018.
40. T. M. Therneau, T. Lumley. A Package for Survival Analysis in R. 2015.
41. N. V. Dimitrov, C. Meyer, D. Gilliland, M. Ruppenthal, W. Chenoweth, W. Malone. Plasma tocopherol concentrations in response to supplemental vitamin E. Am J Clin Nutr 1991; 53(3): 723-9.
42. K. M. Lebold, A. Ang, M. G. Traber, L. Arab. Urinary α-carboxyethyl hydroxychroman can be used as a predictor of α-tocopherol adequacy, as demonstrated in the Energetics Study. Am J Clin Nutr 2012; 96(4): 801-9.
43. A. Ascherio, M. J. Stampfer, G. A. Colditz, E. B. Rimm, L. Litin, W. C. Willett. Correlations of vitamin A and E intakes with the plasma concentrations of carotenoids and tocopherols among American men and women. J Nutr 1992; 122(9): 1792-801.
44. M. G. Traber, E. Mah, S. W. Leonard, G. Bobe, R. S. Bruno. Metabolic syndrome increases dietary alpha-tocopherol requirements as assessed using urinary and plasma vitamin E catabolites: a double-blind, crossover clinical trial. Am J Clin Nutr 2017; 105(3): 571-9.
45. G. Bobe, T. J. Cobb, S. W. Leonard, et al. Increased static and decreased capacity oxidation-reduction potentials in plasma are predictive of metabolic syndrome. Redox Biol 2017; 12: 121-8.
46. E. Mah, T. N. Sapper, C. Chitchumroonchokchai, et al. alpha-Tocopherol bioavailability is lower in adults with metabolic syndrome regardless of dairy fat co-ingestion: a randomized, double-blind, crossover trial. Am J Clin Nutr 2015; 102(5): 1070-80.
47. S. A. Pope, P. T. Clayton, D. P. Muller. A new method for the analysis of urinary vitamin E metabolites and the tentative identification of a novel group of compounds. Arch Biochem Biophys 2000; 381(1): 8-15.
48. J. K. Lodge, M. G. Traber, A. Elsner, R. Brigelius-Flohé. A rapid method for the extraction and determination of vitamin E metabolites in human urine. J Lipid Res 2000; 41(1): 148-54.
49. C. S. Gross, A. T. Milhorat, E. J. Simon. The metabolism of vitamin E. I. The absorption and excretion of d-alpha-tocopheryl-5-methyl-C14-succinate. J Biol Chem 1956; 221(2): 797-805.

50. A. Eisengart, A. T. Milhorat, E. J. Simon, L. Sundheim. The metabolism of vitamin E. II. Purification and characterization of urinary metabolites of alpha-tocopherol. J Biol Chem 1956; 221(2): 807-17.

Supplementary materials

Supplementary Table 1 Comparison of main characteristics between included and excluded participants

Supplementary Table 2 Associations between urinary α-tocopherol metabolites and measures of glucose homeostasis in general population

Supplementary Table 3 Associations between urinary α-tocopherol enzymatic-to-oxidized metabolite ratios and measures of glucose homeostasis in general population

Supplementary Table 4 Associations between circulatory α-tocopherol metabolites and measures of glucose homeostasis in participants without vitamin E supplement use

Supplementary Table 5 Associations between circulatory and urinary metabolites and measures of glucose homeostasis in all participants in censored normal regression (Model 3)

Supplementary Figure 1 Flowchart of participants exclusion

Supplementary Figure 2 Intercorrelation of α-tocopherol metabolites in blood and urine in participants without vitamin E supplement use

Supplementary Figure 3 Association between urinary metabolites and (A) HOMA-B, (B) HOMA-IR, (C) Insulinogenic Index, and (D) Matsuda Index in participants without vitamin E supplement use

Supplementary Figure 4 Association between urinary metabolite ratios and (A) HOMA-B, (B) HOMA-IR, (C) Insulinogenic Index, and (D) Matsuda Index in participants without vitamin E supplement use

The Supplementary materials for this article can be found online at:
https://doi.org/10.1016/j.clnu.2021.01.039

CHAPTER 7

Associations of metabolomic profiles with circulating vitamin E and urinary vitamin E metabolites in middle-aged individuals

Jiao Luo, Yasufumi Hashimoto, Leon G Martens, Fleur L Meulmeester, Nadia Ashrafi, Dennis O Mook-Kanamori, Frits R Rosendaal, J Wouter Jukema, Ko Willems van Dijk, Kevin Mills, Saskia le Cessie, Raymond Noordam, Diana van Heemst

Nutrition. 2022 Jan; 93:111440

Abstract

Vitamin E (α-tocopherol, α-TOH) is transported in lipoprotein particles in the blood, but little is known about the transportation of its oxidized metabolites. In the Netherlands Epidemiology of Obesity Study, we aimed to investigate the associations of 147 circulating metabolomic measures obtained through targeted nuclear magnetic resonance (NMR) with serum α-TOH and its urinary enzymatic (α-CEHC) and oxidized (α-TLHQ) metabolites from 24-hour urine quantified by LC/MS-MS. Multivariable linear regression analyses, in which multiple testing was taken into account, were performed to assess associations between metabolomic measures (determinants; standardized to mean = 0, SD = 1) with vitamin E metabolites (outcomes), adjusted for demographic factors. We analyzed 474 individuals (45% men) with mean (SD) age of 55.7 (6.0) years. Out of 147 metabolomic measures, 106 were associated ($p < 1.34E-3$) with serum α-TOH [median beta (IQR): 0.416 (0.383, 0.466)], predominantly lipoproteins associated with higher α-TOH. The associations of metabolomic measures with urinary α-CEHC are in similar directions as those with α-TOH, but effect sizes were smaller and non-significant [median beta (IQR):0.065 (0.047, 0.084)]. However, associations of metabolomic measures with urinary α-TLHQ were markedly different from the associations of metabolomic measures with both serum α-TOH and urinary α-CEHC, with negative and small-to-null relations to most VLDL and amino acids. Therefore, our results highlight the differences of the lipoproteins involved in the transportation of circulating α-TOH and oxidized vitamin E metabolites. This indicates that circulating α-TOH may be representative of the enzymatic but not to the antioxidative function of vitamin E.

Introduction

Vitamin E (α-tocopherol, α-TOH) is a fat-soluble and essential component in the diet. Vitamin E is important in many physiological processes including fetal development and reproduction, neurodevelopment, and cognitive function; deficiency in vitamin E will induce defects in several developing organs, including primary manifestation in the central nervous system of cerebellar ataxia [1-3]. In addition, vitamin E also has chain-breaking antioxidant activities by competitively reacting with lipid peroxyl radicals to ameliorate lipid peroxidation induced damage [4].

Observational studies have shown associations between higher dietary intake or circulating levels of vitamin E with lower risk of lipid peroxidation-related diseases, such as cardiovascular and neurodegenerative diseases [5-9]. However, there is no evidence for a causal effect based on randomized clinical trials with vitamin E supplementation [10-13] or mendelian randomization [14,15]. A potential explanation lies in the bioactivity of vitamin E that can be catabolized via either hepatic enzymatic pathways or oxidized in the periphery [16]. In the hepatic pathway, vitamin E is enzymatically converted to a spectrum of metabolites of carboxymethyl-hydroxy-chroman (CEHC), with successive shortening of the phytyl side chain, and then eliminated mainly via urine. Alternatively, α-tocopherol reacts with lipid peroxyl radicals, with the opening of the chromanol ring, and generates α-tocopherol quinone (α-TQ); thereafter, α-TQ captures hydrogens converting into α-tocopherol hydroquinone (α-THQ), following the β-oxidation and cyclization of the phytyl side chain, α-tocopheronic acid and α-tocopherol lactone (α-TL) are generated; finally, α-TL are excreted as polar conjugates of α-TL hydroquinone (α-TLHQ) [17-19]. Therefore, circulating vitamin E levels might not represent the authentical antioxidative effect.

The intestinal absorption, hepatic metabolism, and cellular uptake of vitamin E largely follow similar processes as lipids. Briefly, vitamin E is emulsified by digestive enzymes to form micelles and was subsequently absorbed by the intestine via passive diffusion or receptor-mediated transport, followed by the circulation and distribution to target organs and tissues. Once internalized into the enterocytes, lipoproteins are the carriers of vitamin E vascular transportation, independent of the types of isomers. Vitamin E is absorbed and secreted in chylomicrons into the lymphatic system and then is transformed into remnants acquired by the liver parenchymal cells via LDL receptor-mediated uptake [20]. In the liver, different forms of vitamin E were sorted; the highly expressed α-TOH transport protein (α-TTP) selectively readily binds to α-TOH for secretion in VLDL and favors the discrimination of α-TOH among other isomers, protecting α-TOH from excessive degradation and excretion; the final acquisition of vitamin E by tissues is through chylomicron and VLDL catabolism, LDL uptake via LDL receptor or lipoprotein transfer to membranes. Apart from the major portion of circulating α-TOH carried by LDL particles, some of the VLDL-derived α-TOH can also be transferred to HDL during lipolysis. HDL is important for the delivery of α-TOH to extrahepatic tissues, particularly to the central nervous system, and to facilitate the transport of α-TOH from the circulation back to the liver [21]. Therefore, vitamin E utility depends on mechanisms underlying lipoprotein metabolism and relies on lipoprotein-mediated production, processing, and uptake [22]. In addition to their role as vitamin E carriers, lipoproteins, and in particular LDL, are also

susceptible to oxidative modifications that require antioxidants protection. However, it is unclear to what extent vitamin E is catabolized via hepatic enzymatic pathways or oxidized, and what lipoproteins are involved in the transportation of these two metabolism processes.

In the present study, we aim to investigate the cross-sectional associations between circulating metabolomic profiles with circulating α-TOH and urinary enzymatic and oxidized α-TOH metabolites in middle-aged individuals in the Netherlands Epidemiology of Obesity study (NEO).

Method

Study population

The present study was embedded in the population-based prospective cohort Netherlands Epidemiology of Obesity (NEO) study, which aims to study the pathways that lead to obesity-related disorders. The NEO study was initiated in 2008, comprising 6671 participants aged between 45-65 years. Detailed information on the study design and data collection has been described previously [23]. In brief, inhabitants with a self-reported body mass index (BMI) of 27 kg/m^2 or higher living in the greater area of Leiden were eligible to participate. In addition, all inhabitants from the municipality of Leiderdorp were invited irrespective of their BMI. Participants were invited to the NEO study center of the Leiden University Medical Center (LUMC) for one baseline study visit. Blood samples were drawn after an overnight fast and were separated into serum, and aliquots were stored at -80 °C for later measurements. Participants were asked to collect their urine over 24h and completed a general questionnaire at home with their demographic, lifestyle, and clinical data as well as specific questionnaires on diet and physical activity before their first visit. The urine sample was aliquoted and stored at -80 °C for later analyses of urinary vitamin E metabolites. Additionally, participants were asked to bring all medication (prescribed medication including blood pressure-lowering medication, lipid-lowering medication, and glucose-lowering medication, etc., as well as self-medication such as supplements) they were using one month preceding the study visit. The study was approved by the medical ethical committee of the LUMC, and all participants gave written informed consent.

A random subset of 35% baseline Leiderdorp participants (N = 599) was included in this cross-sectional analysis. We excluded individuals with urine collection less than 20 hours or for whom urinary vitamin E metabolites measurements failed (n = 61). Given that the platform used for measurements contained a considerable number of lipoproteins, participants who were taking statins at the time of blood sampling (n = 38) were consequently excluded. We further consecutively excluded participants with implausible metabolites measurement (concentration ≤ 0, n = 4), missingness or outliers of either urinary vitamin E metabolites (n = 7) or serum vitamin E (n = 4) or metabolomic measures (n = 2), and missing data on confounding factors (n = 9). Therefore, the final number of participants included in the present study was 474.

Metabolomic measures profiling

The lipoprotein profiles were quantified using high-throughput ^{1}H-NMR metabolomics (Nightingale Health, Helsinki, Finland). This platform provides simultaneous quantification of 229 metabolites and ratios. After excluding the calculated ratios from the dataset, we included 147 metabolites from 11 classes: lipoprotein subclasses (n = 98), lipoprotein particle sizes (n = 3), apolipoproteins (n = 2), fatty acids (n = 10), cholesterol (n = 9), and also glycerides and phospholipids (n = 9), amino acids (n = 8), ketone bodies (n = 2), inflammation (n = 1), glycolysis related metabolites (n = 3), and fluid balance (n = 2). Lipoprotein subclasses were defined according to the particle sizes as follows: chylomicrons and extremely large VLDL (XXL-VLDL) with particle diameters from 75 nm upwards, five subclasses for VLDL (XL-VLDL: 64 nm, L-VLDL: 53.6 nm, M-VLDL: 44.5 nm, S-VLDL: 36.8 nm, XS-VLDL: 31.3 nm); IDL (28.6 nm), three subclasses for LDL (X-LDL: 25.5 nm, M-LDL: 23 nm, S-LDL: 18.7 nm), and four subclasses of HDL (XL-HDL: 14.3 nm, L-HDL: 12.1 nm, M-HDL: 10.9 nm, S-HDL: 8.7 nm). Detailed information, including quality assurance measures and applications of the platform, have been described elsewhere [24].

Vitamin E and vitamin E metabolites measurements

Circulating serum α-TOH was detected and quantified by Metabolon, Inc. (Durham, NC, USA) on a platform encompassing four liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS) methods (LC-MS/MS negative, LC-MS/MS positive early, LC-MS/MS positive late and LC-MS/MS polar). Given the relative quantification, the absolute amount of α-TOH per liter was not quantified, and therefore data was not expressed as a concentration, but as peak height relative to an internal standard per sample to allow for comparisons between individuals. More information about the quantifications has been described previously [25,26].

Urinary α-CEHC and oxidized α-TOH metabolites (α-TLHQ), presented as their sulfate or glucuronide conjugates, were measured by LC-MS/MS at University College London, the UK between March and May 2019 [19]. The final concentrations of α-TLHQ and α-CEHC were the sum of their corresponding sulfate and glucuronide isoforms.

Briefly, 100µl thawed urine (from 24-hour urine) was centrifuged for 10 min at 14 000 g at room temperature and spiked with 10µl of the internal standards (100µmol/L), lithocholic acid sulfate (LA) and androsterone D4-glucuronide (AD4), and 10 µl was subsequently injected into the LC-MS/MS for detection. Metabolites separation was performed by 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. To minimize system contamination and carryover, UPLC eluents before and after the sample elution were discarded and an additional blank sample (H_2O: MeCN) was run between each detection urine sample. Two separate peaks were observed for both α-TLHQ and α-CEHC conjugated with glucuronide, corresponding to major and minor isoforms. These isoforms had been previously described thoroughly by Pope et al [27]. and Sharma et al. [18,19]. The different elution time (minutes) for internal standards (LA 4.33, AD4 2.7) and each metabolite (2.39, 2.12 and 2.29 for α-TLHQ sulfate,

Table 1 Characteristics of the study population[1]

		N = 474
Demography		
Age (years)	Mean (SD)	55.7 (6.0)
Sex (male)	Frequency (proportions)	213 (45%)
BMI (kg/m^2)	Median (IQR)	25.3 (23.1, 27.8)
Lifestyle factors		
Dutch healthy diet index	Mean (SD)	59.7 (8.4)
Energy intake (KJ/day)	Median (IQR)	9106 (7326, 11078)
Physical activity (MET-h/week)	Median (IQR)	28.9 (16.0, 48.8)
Smoking		
Current	Frequency (proportions)	50 (10%)
Former	Frequency (proportions)	222 (47%)
Never	Frequency (proportions)	202 (43%)
Vitamin E supplement use (yes)[2]	Frequency (proportions)	124 (26%)
Vitamin E metabolites measurements		
Blood (log10-transformed)		
α-tocopherol	Mean (SD)	8.5 (0.1)
Urinary		
α-TLHQ (nmol/mmol creatinine)	Median (IQR)	1864.9 (1347.1, 2770.3)
α-CEHC(nmol/mmol creatinine)	Median (IQR)	271.0 (181.4, 439.4)
α-TLHQ (log10-transformed)	Mean (SD)	3.3 (0.2)
α-CEHC (log10-transformed)	Mean (SD)	2.4 (0.3)

[1]SD: standard deviation; IQR: interquartile range.
[2]Vitamin E supplement use was defined as either vitamin E supplement use or multiple vitamin supplement use.
BMI, body mass index; CEHC, carboxymethyl-hydroxychroman.

glucuronide minor and major, 2.64, 2.50, 2.56 for α-CEHC sulfate, glucuronide minor and major) guaranteed that all metabolites could be separated in a single chromatographic run. Metabolites analyses were then performed by MS using a Waters ACQUITY UPLC coupled to a triple-quadrupole Xevo TQ-S fitted with electrospray ionization in negative ion mode. Using multiple reaction monitoring (MRM) mode, specific parent and daughter ions were determined in scan mode and following collision activated dissociation with argon. These ions were then used to quantify each α-TOH metabolite from transitions that corresponded to their molecular masses. Since creatinine concentration is frequently used as a proxy of kidney function, and in cases of severe renal dysfunction, the creatinine clearance rate will be "overestimated" because the active secretion of creatinine will account for a larger fraction of the total creatinine. Therefore, to correct the dilution effect, urinary creatinine concentrations (mmol/L) were also measured by triple-quadrupole Micro Quattro mass spectrometry (MicroMass, Waters, UK). Therefore, the final concentration of α-TOH metabolites was in nmol per mmol of creatinine. A quality control (QC) assessment was performed throughout the

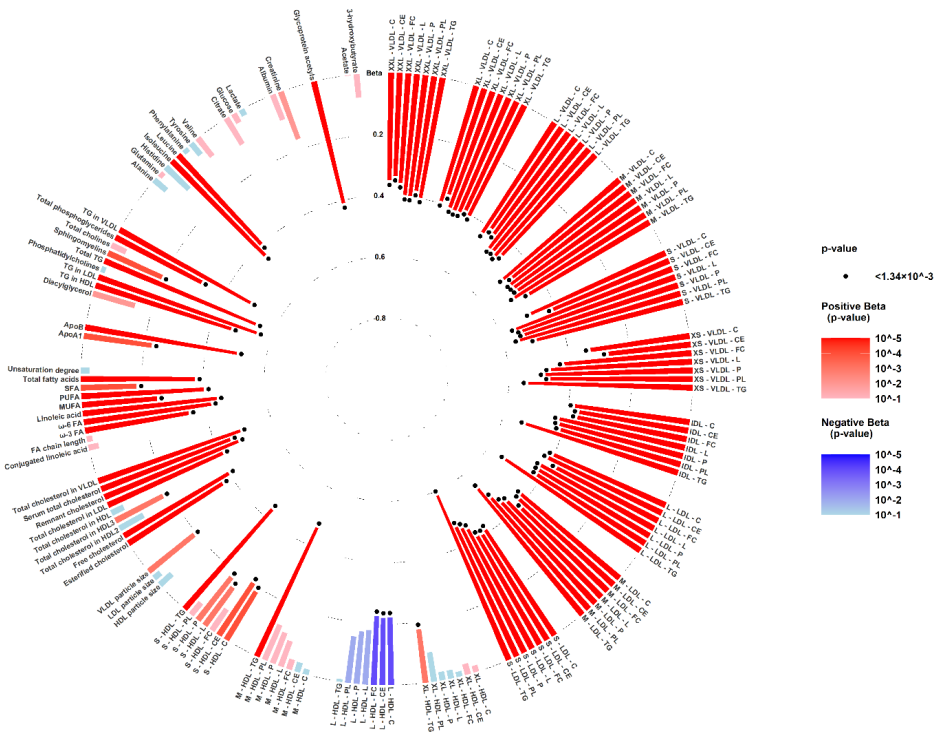


Figure 1 Associations between 147 circulating metabolomic measures with circulating α-tocopherol

Associations were derived from a multivariable linear regression model in the study population (N = 474) adjusted for age, sex, BMI, smoking status, Dutch Healthy Diet Index, energy intake and physical activity.

Effect estimates refers to the coefficients for different metabolomic measures obtained from the linear regression models. we standardized the log10-transformed metabolomic measures and the serum and urinary vitamin E measures (mean = 0, SD =1), so that the regression coefficient with its corresponding 95% confidence interval (CI) can be interpreted as the mean change in SD of the outcome with respect to a one-SD change in the determinant (standardized concentrations of the metabolomic measures). Red and blue colors indicate the directions of the estimates, namely the positive and negative coefficients respectively, and the gradients suggests different significant levels. A p-value of below 1.34×10^{-3} (0.05/37, 37 is the number of independent metabolomic measures) is considered significant, presented as a black dot above each bar. Full names and descriptive information of metabolomic measures are listed in Supplementary Table.

quantification to deal with the variations in sample quality and UPLC-MS/MS performance. Detailed information on the measurements has been described previously [28].

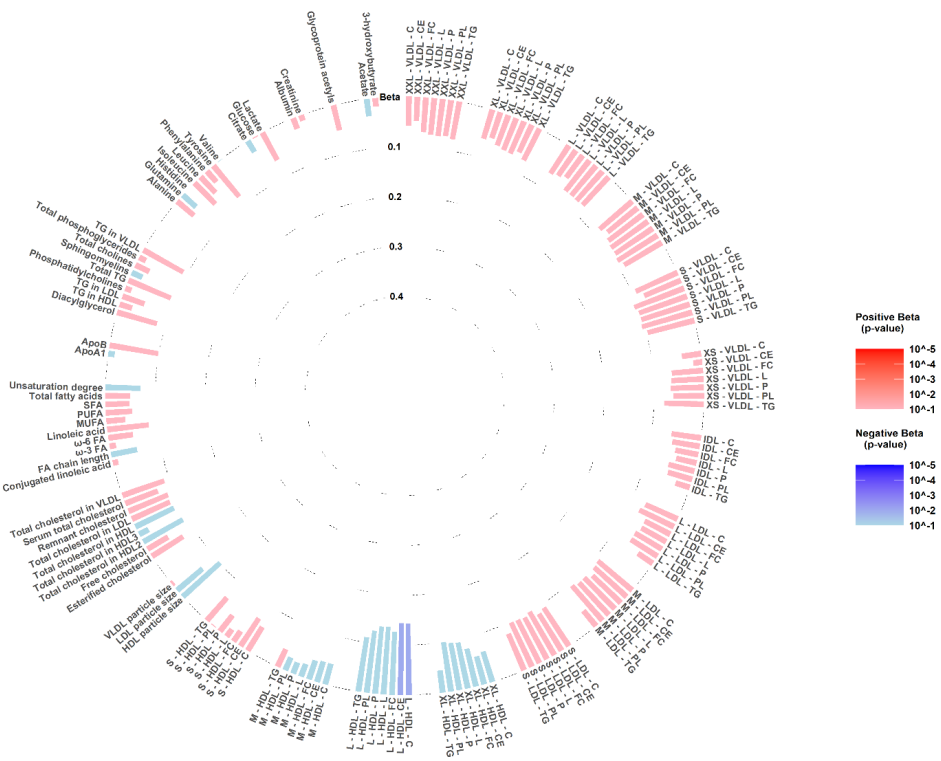


Figure 2 Associations between 147 circulating metabolomic measures with urinary α-CEHC

Associations were derived from a multivariable linear regression model in the study population (N = 474) adjusted for age, sex, BMI, smoking status, Dutch Healthy Diet Index, energy intake and physical activity. Figure legend is the same with Figure 1, except that no significant associations [$p > 1.34 \times 10^{-3}$ (0.05/37, 37 is the number of independent metabolomic measures)] were detected.

Confounding factors

To determine BMI (kg/m^2), body weight was measured without shoes, and one kilogram was subtracted for the correction of clothing weight. Smoking status was categorized into the current smoker, former smoker, and non-smoker. Physical activity levels (in MET-hours per week, MET: Metabolic equivalent of task) were estimated based on the frequency and duration of leisure physical activity over the past 4 weeks reported by participants on the Short Questionnaire to Assess Health-enhancing physical activity (SQUASH) [29]. A semiquantitative food frequency questionnaire was used to assess food and beverage intake. Total energy intake (in kJ) and Dutch Healthy Diet Index (DHD-index) were subsequently estimated based on dietary intake [30]. Vitamin E supplement use was collected via questionnaires and was defined as either vitamin E supplement use only or multivitamin supplement use (yes/no).

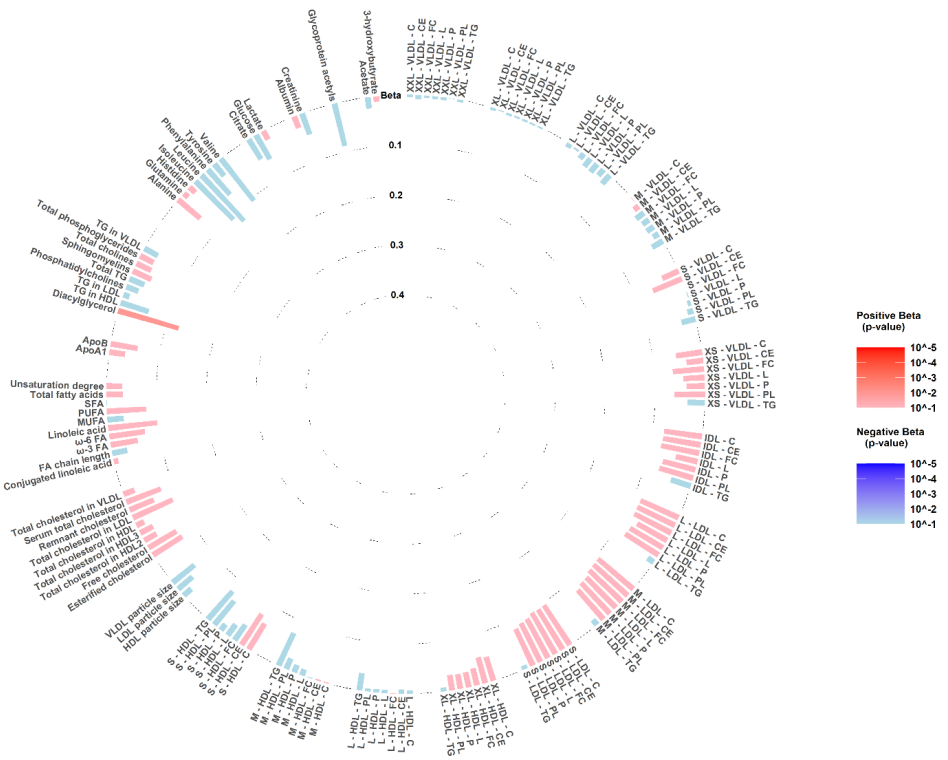


Figure 3 Associations between 147 circulating metabolomic measures with urinary α-TLHQ

Associations were derived from a multivariable linear regression model in the study population (N = 474) adjusted for age, sex, BMI, smoking status, Dutch Healthy Diet Index, energy intake and physical activity. Figure legend is the same with Figure 1, except that no significant associations [$p > 1.34 \times 10^{-3}$ (0.05/37, 37 is the number of independent metabolomic measures)] were detected.

Statistical analysis

Descriptive characteristics of the study population were presented as mean (standard deviation, SD), median (interquartile range, IQR) for normally distributed variables and skewed variables respectively, and frequency (proportions) for categorical variables.

Vitamin E metabolites and metabolomic measures were log10-transformed to approximate a normal distribution. Observed metabolites concentrations located beyond 4 standard deviations from the mean after log10-transformation were classified as outliers and further excluded. Since missing data on the metabolomic measures were most likely due to concentrations that were lower than the limit of detection, these missing values were imputed by giving them the value of half of the minimum observed value for each metabolite. In addition, we assessed

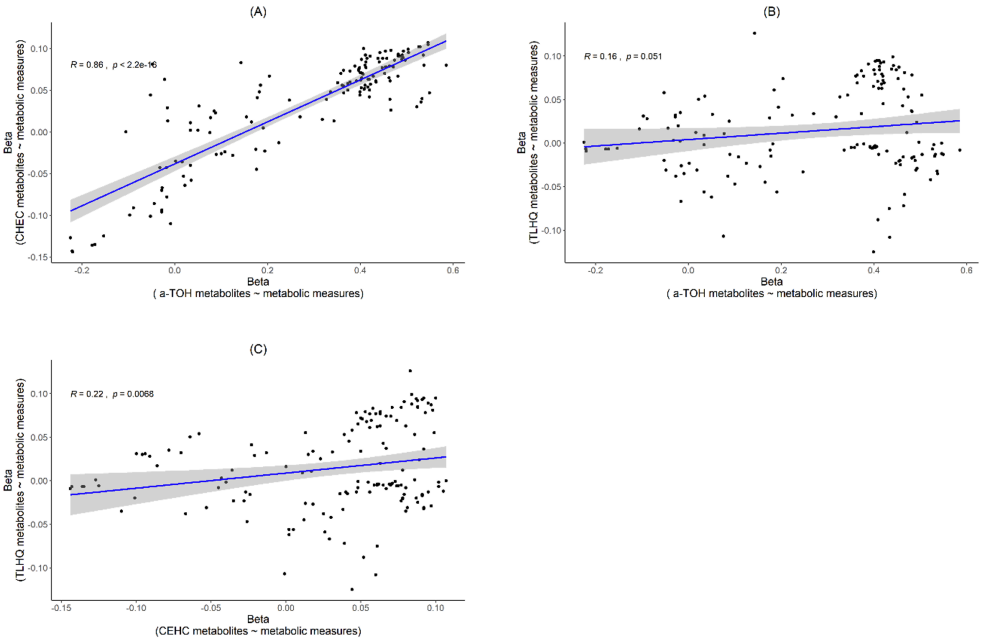


Figure 4 Correlations of effect estimates between metabolomic measures with circulating α-tocopherol, urinary α-CEHC and α-TLHQ

(A) the correlation of regression coefficients of metabolomic measures with circulating α-tocopherol and metabolomic measures with urinary α-CEHC. (B) the correlation of regression coefficients of metabolomic measures with circulating α-tocopherol and α-TLHQ. (C) the correlation of regression coefficients of metabolomic measures with urinary α-CEHC and α-TLHQ.

Axis represents the beta coefficient derived from multiple linear regression of metabolomic measures (determinants) and vitamin E metabolites (outcomes). Correlation coefficients and p value were derived from Pearson correlation.

the percentage of missingness for each metabolomic measure. To compare the effect estimates, i.e. the coefficients for different metabolomic measures obtained from the regression models, we standardized the log10-transformed metabolomic measures and the serum and urinary vitamin E measures (mean = 0, SD =1), so that the regression coefficient with its corresponding 95% confidence interval (CI) can be interpreted as the mean change in SD of the outcome with respect to a one-SD change in the determinant (standardized concentrations of the metabolomic measures).

Multivariable-adjusted linear regression models were fitted, with metabolomic measures as determinants, confounding factors as covariates, and vitamin E metabolites as outcomes. Based on prior knowledge, confounding factors included age, sex, BMI (kg/m^2), physical activity (MET-hours per week), smoking habits (non-smoker, current smoker, or former smoker), Dutch Healthy Diet Index, and total energy intake (kJ/day). Scatter plots were used to visualize the difference in both direction and effect sizes of the estimated associations among

metabolomic measures with different vitamin E metabolites, and we calculated Pearson correlations between the effect estimates derived from the regression results. Therefore, the correlation indicates the similarity of those associations.

Vitamin E supplement use may have an influence on metabolomic measures and vitamin E conversion and will potentially distort the associations. However, given the high heterogeneity of vitamin E supplement use, as either vitamin E only or multivitamin use, as well as limited information on frequency and dosage, natural or modified vitamin E acetate, we therefore additionally performed the regression analyses in participants who did not take vitamin E or multivitamin supplements (N = 350). Furthermore, in order to test the effect modification by obesity status, we stratified participants into normal weight (BMI < 25kg/m^2, N = 217) and overweight (since only 59 were obese with BMI above 30 kg/m^2, we combined individuals with BMI >= 25kg/m^2, N = 257), and all multivariable regression models in the main analyses were conducted in each stratum.

Given that most of the metabolomic biomarkers, especially lipid subclasses, were highly correlated, conventional correction for multiple testing (e.g., Bonferroni) is too stringent. Therefore, the 'effective number' (Meff) procedure was used which identified independent metabolomic traits [31]. The final significance threshold of the p-value was then defined as 0.05/37 = 1.34E-3. All the analyses were undertaken using R (v3.6.1) statistical software (The R Foundation for Statistical Computing, Vienna, Austria).

Results

Characteristics of the study population

474 participants were included in the current study after excluding participants with missing data or outliers in serum or urinary vitamin E measures; characteristics of the study population are presented in **Table 1**. The mean (SD) age was 55.7 (6.0) years with a median (IQR) BMI of 25.3 (23.1, 27.8) kg/m^2; 45% of the participants were male. Approximately 124 (26%) participants used vitamin E supplements, and 50 (10%) were current smokers. Summaries of metabolomic biomarkers are presented in **Supplementary Table 1**. The percentages of missing data of individual metabolomic measures were all below 30%, with 9 out of 147 metabolites having 20% or more missingness, of which 7 were extremely large VLDL-characteristics (**Supplementary Figure 1**).

3.2 Main analyses

In the multivariable-adjusted linear regression model, 106 out of 147 metabolomic measures were associated with serum α-TOH with p < 1.34E-03 [median effect size (IQR): 0.416 (0.383, 0.466)] (**Figure 1** and **Supplementary Table 1**). Three of the 106 associations were negative: higher levels in total cholesterol, cholesterol ester, and free cholesterol in large HDL were associated with lower mean serum α-TOH [effect estimates -0.221 (95% confidence interval, CI: -0.330, -0.112), -0.220 (95% CI: -0.329, -0.112) and -0.225 (95% CI: -0.332, -0.118), respectively, per 1 SD higher level of the metabolomic measure]. In all other cases, higher levels of the metabolic measures were associated with higher α-TOH.

These associations include VLDL, IDL, LDL, and small HDL, total cholesterol (not in HDL and HDL2) particles and its components, Apo-A and Apo-B, glycerides and phospholipids, and glycoprotein acetyls, with effect sizes ranging from 0.154 (95% CI: 0.063, 0.244) SD for total lipids in small HDL to 0.585 (95% CI: 0.509, 0.660) SD for triglycerides in small-LDL. Moreover, per-SD higher level of leucine and isoleucine were associated with a 0.399 (95% CI:0.290, 0.507) and 0.434 (95% CI:0.327, 0.540) SD higher α-TOH.

Figure 2 and **Figure 3** present the associations between metabolomic measures with urinary α-CEHC and α-TLHQ metabolites, separately. None of the analyses were statistically significant upon correction for multiple testing. Nevertheless, for α-CEHC, the direction of the associations with metabolomics measures was similar to those of α-TOH with metabolomics measures, the effect sizes, however, were much smaller [median effect size (IQR): 0.065 (0.047, 0.084)], as shown in **Figure 2**. For α-TLHQ, the strongest association was with diacylglycerol 0.126 (95%CI: 0.037, 0.215, p = 0.006). The direction of the associations was different substantially from those of α-TOH with metabolomics measures, and only very small VLDL, IDL, LDL, XL-HDL, and fatty acids showed associations in the same direction.

The effect estimates between metabolomic measures with circulating α-TOH were strongly correlated with the effect estimates of metabolomic measures with urinary α-CEHC ($r = 0.86$, $p < 0.001$). However, the correlations between the effect estimates of metabolomic measures with α-TOH and with urinary α-TLHQ, and the correlations between the effect estimates of metabolomic measures with α-CEHC and with urinary α-TLHQ were very weak, $r = 0.16$ ($p = 0.05$) and $r = 0.22$ ($p = 0.007$), respectively (**Figure 4**).

Sensitivity analyses

Excluding vitamin E supplement users

We excluded 124 participants with either vitamin E or multivitamin supplementation, leaving 350 participants for further analyses. The associations between metabolomic measures and vitamin E metabolites in this group were generally consistent with the analyses in the whole study population (**Supplementary Figure 2**). However, several effect sizes became larger. Notably, a one-SD higher level of total cholesterol and cholesterol esters in medium LDL and small LDL particles were associated, even after correction for multiple testing, with higher levels of α-CEHC with effect sizes of 0.176 (95% CI: 0.071, 0.280), 0.175 (95% CI: 0.070, 0.280), 0.180 (95% CI: 0.075, 0.285) and 0.179 (95% CI: 0.074, 0.284) SD, respectively. Furthermore, 22 out of the 147 metabolomic measures (most notably 4 IDL, 16 LDL, total cholesterol, and total cholesterol in LDL) were associated with α-TLHQ levels. Specifically, higher levels of LDL cholesterol sub-particles, except for the amount of triglycerides in LDL particles, were associated with higher levels of α-TLHQ, with effect sizes ranging from 0.186 (95% CI: 0.079, 0.293) SD for large LDL particles to 0.227 (95% CI: 0.121, 0.333) SD for cholesterol esters in small LDL. In addition, a one-SD higher level of total cholesterol, total cholesterol in LDL were associated with 0.186 (95% CI: 0.076, 0.295) and 0.207 (95% CI: 0.102, 0.313) SD higher α-TLHQ, respectively.

The effect estimates of metabolomic measures with circulating α-TOH were strongly correlated with the effect estimates of metabolomic measures with urinary α-CEHC ($r = 0.69$, $p < 0.001$), whereas no correlation was found between the effect estimates of metabolomic measures with circulating α-TOH and with urinary α-TLHQ ($r = -0.013$, $p = 0.88$). A moderate correlation was observed for the estimates between metabolomic measures with circulating α-CEHC and with urinary α-TLHQ ($r = 0.49$, $p < 0.001$), **Supplementary Figure 3**. However, this correlation was mainly due to the association with LDL ($r = 0.18$, $p = 0.053$ after excluding the LDL subclass from the list of metabolic measures).

Stratification analyses by obesity

In the normal weight subgroup analyses (N = 217), the associations of metabolomic measures and serum α-TOH were analogous to those obtained from the main analyses (**Supplementary Figure 4**). However, the associations with HDL were no longer significant, resulting in slightly fewer (95 out of 147) significant associations. The median significant effect size (IQR) was 0.375 (0.330, 0.414). Similarly, the associations between metabolomic profiles with α-CEHC did not differ materially. Interestingly, diacylglycerol was significantly positively related to α-TLHQ. However, the associations between HDL and α-TLHQ became stronger while the association between XXL- and XL-VLDL and α-TLHQ turned positive and stronger compared to the estimates from the whole population, despite being insignificant.

In the overweight subgroup analyses (N = 257), the magnitude of the estimations from regression analyses was generally larger than those derived from the main analyses in both three vitamin E metabolites analyses (**Supplementary Figure 5**). Particularly, the associations of metabolomic profiles with α-TOH remained in the same directions, and 99 out of 147 associations were significant [median effect size (IQR): 0.497 (0.423, 0.561)]. Likewise, the relationships of metabolomic measures with α-CEHC did not change substantially. However, with regards to α-TLHQ, the direction of the association with HDL, notably large and middle HDL, turned to positive, though insignificant.

Discussion

In this cross-sectional study, we investigated the association between circulating metabolomic measures and urinary vitamin E metabolites. The direction of the estimates from metabolomic measures and circulating α-TOH and urinary α-CEHC were similar but with weaker effect sizes for α-CEHC than α-TOH, whereas α-TLHQ showed distinct associations. 106 out of 147 metabolomic measures were associated with circulating α-TOH after correction for multiple testing, while no significant associations were identified for associations with urinary α-CEHC and α-TLHQ. Sensitivity analyses in participants without vitamin E supplement use generally showed consistent results with the main analyses for each vitamin E metabolite, but with significant associations between IDL, LDL, total cholesterol, and cholesterol content in LDL particle and α-TLHQ.

Our study shows similar associations of metabolomic measures with urinary α-CEHC and with circulating α-TOH. The liver parenchymal cells acquire α-TOH through taking up chylomicron remnants, which contains a major proportion of absorbed α-TOH, and the highly expressed α-TTP facilitate the preferential enrichment of α-TOH in VLDL particles secreted by the liver. These processes are responsible for the regulation and release of α-TOH into circulation and consequently for its delivery to tissues. Therefore, the vitamin E enzymatic activity in the liver is well regulated to maintain a certain level of α-TOH, i.e. higher α-TOH gives higher α-CEHC [32]. However, the different associations observed between metabolomic measures with urinary α-TLHQ and with blood α-TOH indicate that other processes that are not associated with α-TOH regulated the association for α-TLHQ. In addition, we observed that circulating α-TOH is associated with higher levels of most lipoprotein fractions and cholesterols. This observation may be somewhat counterintuitive to previous observations of higher vitamin E circulating levels with a lower risk of atherosclerosis and cardiovascular diseases [9]. Nevertheless, individuals with hyperlipidemia were found to have reduced uptake of the newly absorbed α-TOH into blood; the abnormal lipoprotein metabolism does not necessarily increase α-TOH delivery to the peripheral tissues and this uptake reduction of α-TOH may be relevant to the pathogenesis of atherosclerosis [33]. In addition, our results are in line with previous research which found that α-TOH concentrations were correlated with serum lipids levels and that the retainment of plasma α-TOH was longer with higher serum total lipids [22]. This might be attributed to the notion that higher lipid concentrations can keep vitamin E from reaching peripheral tissues, and the catabolism and uptake of lipoproteins decreases at high concentrations of lipid. However, higher levels of large HDL particles were associated with lower levels of α-TOH. This may be due to that α-TOH content in HDL particles depends not only on tocopherol levels but also on HDL concentrations, and HDL α-TOH retention was found related to a high concentration of HDL fraction [33]. Two branched-chain amino acids (leucine and isoleucine) are positively associated with higher α-TOH. Experimental studies have demonstrated that vitamin E is crucial for the maintenance of energy homeostasis, and its deficiency dysregulated energy metabolism and mitochondrial dysfunction, measured by extracellular oxygen consumption [34-36]. Ketogenic amino acids, particularly leucine, which can be utilized for ketone synthesis, were elevated in vitamin E deficient Zebrafish brains, where there are probably elevated lipid peroxidation and metabolic disruptions [35,36]. Health states that are associated with increased oxidative stress are likely to have a greater antioxidant requirement, which would result in depletion of circulating levels of vitamin E, and higher urinary concentrations of α-TLHQ. The negative though the insignificant association of leucine with α-TLHQ contrasts to the findings from previous experimental studies, understanding the underlying mechanisms, will however require additional efforts.

Results from sensitivity analyses after excluding supplement users were generally consistent with the main analyses, but with several significant associations of α-TLHQ. Notably, higher levels of substances that are susceptible to oxidative modifications, particularly IDL and LDL particles were significantly associated with higher urinary α-TLHQ. Several fatty acids are also closely related to α-TLHQ though not significantly. The inhibitory effect against lipid peroxidation of vitamin E decreases gradually from polyunsaturated fatty acids (PUFA) to cholesterol.

In accordance, LDL, as the most susceptible particle to oxidative modification compared to the other lipoproteins, was associated with α-TLHQ. However, this does not include triglycerides in LDL particles which are more resistant to lipid peroxidation. Therefore, oxidation, rather than the level of α-TOH, regulated the association of metabolomic measures and α-TLHQ. Despite the strong correlation ($r = 0.48$) between the effect estimates of metabolomic measures with α-CEHC and with α-TLHQ, this is predominantly driven by LDL particles.

In the stratification analyses by obesity, the associations of metabolomic measures with α-TOH and α-CEHC did not differ materially in general, but the effect sizes are larger in the overweight group possibly due to the higher lipids levels in overweight participants compared to the normal-weight individuals. Interestingly, the associations of lipoproteins, particularly VLDL and HDL with α-TLHQ differed. Though none of the associations is significant, VLDL is positively and HDL is negatively associated with α-TLHQ in the strata of normal weight, whereas opposite directions were observed in the strata of overweight. Excessive fat accumulation will certainly lead to elevated oxidative stress, and the supply of vitamin E by HDL might be more important under conditions of oxidative stress due to the independence of regulatory mechanisms of cholesterol metabolism [37]. Previous efforts have demonstrated the exchange of α-TOH between lipoproteins, which may depend on the ratio of HDL/LDL [38]; the discrepancies of VLDL associations might imply a different transfer in obese participants compared to non-obese provided distinct lipid profiles in obese people. However, the underlying mechanisms warrant further investigation.

One strength of the present study is that we simultaneously quantified the concentrations of circulating α-tocopherol and urinary metabolites derived from two metabolic pathways, which facilitates the exploration of the circulating level versus the functional level. In addition, the measurement of urinary α-TLHQ was performed by LC-MS/MS based method, which deliberately avoids artefactual oxidation products of α-CEHC that might result from previous chromatography-mass spectrometry (GC-MS) detection. The method we used measures the intact conjugate with minimal preparation and has been demonstrated with solid reliability and reproducibility [19]. Several limitations should also be noted. Firstly, some associations increased after excluding vitamin E supplement users, suggesting that there might be effect modification by the use of supplements or we might introduce collider stratification bias in the users of these supplements. However, the information on supplement use was very limited, as there were no data available on the dosages and frequency and natural or modified vitamin E acetate, resulting in a highly heterogeneous group. Therefore, further exploration of the effect of vitamin E supplement use on these associations was not feasible. Secondly, within our study population, we are not able to test these associations in different physiological situations that may affect the lipoproteins involved in the transportation process such as fasting or not [39-41], or in individuals with elevated lipid peroxidation that will reduce α-TOH bioavailability such as lower bioavailability identified in metabolomic syndrome patients compared with healthy controls [42]. Thirdly, we might still have insufficient power for some of these associations, particularly in the sensitivity analysis. However, we do not only perform statistical hypothesis testing, but apart from the point estimates obtained from our multivariable adjusted regression analyses illustrated in the main text,

we also calculate confidence intervals (CIs) for each estimate, as shown in the supplementary table. CIs reflect the precision of the estimation in the sample. In addition, we had a specific focus on the similarity of the directionality of the associations between metabolomic measurements with circulating α-TOH and the associations between metabolomic measurements with circulating α-CEHC/α-TLHQ. Lastly, the observational design could not rule out residual confounding.

Conclusion

Associations of metabolomic measures with circulating α-TOH and urinary oxidized vitamin E α-CEHC are very similar in direction, whereas associations of metabolomic measures with α-TLHQ were markedly different from the associations of metabolomic measures with both serum α-TOH and urinary α-CEHC. Our results highlight the differences of the lipoproteins involved in the transportation of enzymatic and oxidized vitamin E metabolites. This indicates that circulating α-TOH may be representative for the enzymatic but not to antioxidative function of vitamin E.

References

1. M. G. Traber. Vitamin E inadequacy in humans: causes and consequences. Adv Nutr 2014; 5(5): 503-14.
2. M. G. Traber. Vitamin E. Adv Nutr 2021.
3. M. G. Traber. Vitamin E: necessary nutrient for neural development and cognitive function. Proc Nutr Soc 2021: 1-8.
4. F. Khadangi, A. Azzi. Vitamin E - The Next 100 Years. IUBMB Life 2019; 71(4): 411-5.
5. S. Ashley, S. Bradburn, C. Murgatroyd. A meta-analysis of peripheral tocopherol levels in age-related cognitive decline and Alzheimer's disease. Nutr Neurosci 2019: 1-15.
6. M. C. Morris, D. A. Evans, J. L. Bienias, et al. Dietary intake of antioxidant nutrients and the risk of incident Alzheimer disease in a biracial community study. JAMA 2002; 287(24): 3230-7.
7. S. M. Zhang, M. A. Hernan, H. Chen, D. Spiegelman, W. C. Willett, A. Ascherio. Intakes of vitamins E and C, carotenoids, vitamin supplements, and PD risk. Neurology 2002; 59(8): 1161-9.
8. M. J. Engelhart, M. I. Geerlings, A. Ruitenberg, et al. Dietary intake of antioxidants and risk of Alzheimer disease. JAMA 2002; 287(24): 3223-9.
9. D. Aune, N. Keum, E. Giovannucci, et al. Dietary intake and blood concentrations of antioxidants and the risk of cardiovascular disease, total cancer, and all-cause mortality: a systematic review and dose-response meta-analysis of prospective studies. Am J Clin Nutr 2018; 108(5): 1069-91.
10. A. W. S. Rutjes, D. A. Denton, M. Di Nisio, et al. Vitamin and mineral supplementation for maintaining cognitive function in cognitively healthy people in mid and late life. Cochrane Database Syst Rev 2018; (12).
11. J. McCleery, R. P. Abraham, D. A. Denton, et al. Vitamin and mineral supplementation for preventing dementia or delaying cognitive decline in people with mild cognitive impairment. Cochrane Database Syst Rev 2018; (11).
12. N. Farina, D. Llewellyn, Mgekn Isaac, N. Tabet. Vitamin E for Alzheimer's dementia and mild cognitive impairment. Cochrane Database of Syst Rev 2017; (4).
13. H. N. Siti, Y. Kamisah, J. Kamsiah. The role of oxidative stress, antioxidants and vascular inflammation in cardiovascular disease (a review). Vascul Pharmacol 2015; 71: 40-56.
14. J. Luo, S. le Cessie, D. van Heemst, R. Noordam. Diet-Derived Circulating Antioxidants and Risk of Coronary Heart Disease. J Am Coll Cardiol 2021; 77(1): 45-54.
15. D. M. Williams, S. Hägg, N. L. Pedersen. Circulating antioxidants and Alzheimer disease prevention: a Mendelian randomization study. Am J Clin Nutr 2019; 109(1): 90-8.
16. R. Brigelius-Flohe, M. G. Traber. Vitamin E: function and metabolism. Faseb j 1999; 13(10): 1145-55.
17. L. Taylor, N. Krueger, O. Malysheva, J. Atkinson, R. S. Parker. ω-Hydroxylation of α-tocopheryl quinone reveals a dual function for cytochrome P450-4F2 in vitamin E metabolism. Bioorg Med Chem 2018; 26(20): 5555-65.

18. G. Sharma, D. P. Muller, S. M. O'Riordan, et al. Urinary conjugated α-tocopheronolactone--a biomarker of oxidative stress in children with type 1 diabetes. Free Radic Biol Med 2013; 55: 54-62.
19. G. Sharma, D. Muller, S. O'Riordan, 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): 599-600.
20. M. Hacquebard, Y. A. Carpentier. Vitamin E: absorption, plasma transport and cell uptake. Curr Opin Clin Nutr Metab Care 2005; 8(2): 133-8.
21. L. Schmölz, M. Birringer, S. Lorkowski, M. Wallert. Complexity of vitamin E metabolism. World J Biol Chem 2016; 7(1): 14-43.
22. M. G. Traber, S. W. Leonard, G. Bobe, et al. α-Tocopherol disappearance rates from plasma depend on lipid concentrations: studies using deuterium-labeled collard greens in younger and older adults. Am J Clin Nutr 2015; 101(4): 752-9.
23. R. de Mutsert, M. den Heijer, T. J. Rabelink, et al. The Netherlands Epidemiology of Obesity (NEO) study: study design and data collection. Eur J Epidemiol 2013; 28(6): 513-23.
24. P. Wurtz, A. J. Kangas, P. Soininen, D. A. Lawlor, G. Davey Smith, M. Ala-Korpela. Quantitative Serum Nuclear Magnetic Resonance Metabolomics in Large-Scale Epidemiology: A Primer on -Omic Technologies. Am J Epidemiol 2017; 186(9): 1084-96.
25. E. T. Cirulli, L. Guo, C. Leon Swisher, et al. Profound Perturbation of the Metabolome in Obesity Is Associated with Health Risk. Cell Metab 2019; 29(2): 488-500.e2.
26. T. Long, M. Hicks, H. C. Yu, et al. Whole-genome sequencing identifies common-to-rare variants associated with human blood metabolites. Nat Genet 2017; 49(4): 568-78.
27. S. A. Pope, G. E. Burtin, P. T. Clayton, D. J. Madge, D. P. Muller. Synthesis and analysis of conjugates of the major vitamin E metabolite, alpha-CEHC. Free Radic Biol Med 2002; 33(6): 807-17.
28. J. Luo, Y. Hashimoto, L. G. Martens, et al. Associations of metabolomic profiles with circulating vitamin E and urinary vitamin E metabolites in middle-aged individuals. Nutrition 2022; 93: 111440.
29. G. C. Wendel-Vos, A. J. Schuit, W. H. Saris, D. Kromhout. Reproducibility and relative validity of the short questionnaire to assess health-enhancing physical activity. J Clin Epidemiol 2003; 56(12): 1163-9.
30. M. Looman, E. J. Feskens, M. de Rijk, et al. Development and evaluation of the Dutch Healthy Diet index 2015. Public Health Nutr 2017; 20(13): 2289-99.
31. J. Li, L. Ji. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. Heredity 2005; 95(3): 221-7.
32. K. M. Lebold, A. Ang, M. G. Traber, L. Arab. Urinary α-carboxyethyl hydroxychroman can be used as a predictor of α-tocopherol adequacy, as demonstrated in the Energetics Study. Am J Clin Nutr 2012; 96(4): 801-9.
33. W. L. Hall, Y. M. Jeanes, J. K. Lodge. Hyperlipidemic subjects have reduced uptake of newly absorbed vitamin E into their plasma lipoproteins, erythrocytes, platelets, and lymphocytes, as studied by deuterium-labeled alpha-tocopherol biokinetics. J Nutr 2005; 135(1): 58-63.
34. A. A. Moazzami, S. Frank, A. Gombert, et al. Non-targeted 1H-NMR-metabolomics suggest the induction of master regulators of energy metabolism in the liver of vitamin E-deficient rats. Food Funct 2015; 6(4): 1090-7.

35. M. McDougall, J. Choi, K. Magnusson, L. Truong, R. Tanguay, M. G. Traber. Chronic vitamin E deficiency impairs cognitive function in adult zebrafish via dysregulation of brain lipids and energy metabolism. Free Radic Biol Med 2017; 112: 308-17.
36. J. Zhang, B. Head, S. W. Leonard, J. Choi, R. L. Tanguay, M. G. Traber. Vitamin E deficiency dysregulates thiols, amino acids and related molecules during zebrafish embryogenesis. Redox Biol 2021; 38: 101784.
37. I. Kolleck, M. Schlame, H. Fechner, A. C. Looman, H. Wissel, B. Rüstow. HDL is the major source of vitamin E for type II pneumocytes. Free Radic Biol Med 1999; 27(7-8): 882-90.
38. M. G. Traber, J. C. Lane, N. R. Lagmay, H. J. Kayden. Studies on the transfer of tocopherol between lipoproteins. Lipids 1992; 27(9): 657-63.
39. G. M. Kostner, K. Oettl, M. Jauhiainen, C. Ehnholm, H. Esterbauer, H. Dieplinger. Human plasma phospholipid transfer protein accelerates exchange/transfer of alpha-tocopherol between lipoproteins and cells. Biochem J 1995; 305 (Pt 2): 659-67.
40. W. A. Behrens, J. N. Thompson, R. Madere. Distribution of alpha-tocopherol in human plasma lipoproteins. Am J Clin Nutr 1982; 35(4): 691-6.
41. A. Bjørneboe, G. E. Bjørneboe, E. Bodd, B. F. Hagen, N. Kveseth, C. A. Drevon. Transport and distribution of alpha-tocopherol in lymph, serum and liver cells in rats. Biochim Biophys Acta 1986; 889(3): 310-5.
42. E. Mah, T. N. Sapper, C. Chitchumroonchokchai, et al. α-Tocopherol bioavailability is lower in adults with metabolic syndrome regardless of dairy fat co-ingestion: a randomized, double-blind, crossover trial. Am J Clin Nutr 2015; 102(5): 1070-80.

Supplementary materials

Supplementary Table 1 Summary of metabolomic biomarkers and the associations with serum α-TOH, urinary α-CEHC and α-TLHQ

Supplementary Figure 1 The percentage of observations below detection limit per metabolite in all 147 metabolomic measures.

Supplementary Figure 2 Associations between 147 metabolomic measures with (A) circulating α-tocopherol, (B) urinary α-CEHC, (C) urinary α-TLHQ in non-supplement users

Supplementary Figure 3 Correlations of effect estimates between metabolomic measures with circulating α-tocopherol, urinary α-CEHC and α-TLHQ in sensitivity analysis (N = 350)

Supplementary Figure 4 Associations between 147 metabolomic measures with (A) circulating α-tocopherol, (B) urinary α-CEHC, (C) urinary α-TLHQ in participants with normal weight (BMI < 25kg/m^2)

Supplementary Figure 5 Associations between 147 metabolomic measures with (A) circulating α-tocopherol, (B) urinary α-CEHC, (C) urinary α-TLHQ in overweight participants (BMI >= 25kg/m^2)

The Supplementary materials for this article can be found online at: *https://drive.google.com/drive/folders/1d46G5jf6fIZUUlp6aHpjnz74SnLjzQPd?usp=sharing*

CHAPTER 8

Depression and inflammatory bowel disease: a bidirectional two-sample Mendelian randomization study

Jiao Luo, Zhongwei Xu, Raymond Noordam, Diana van Heemst, Ruifang Li-Gao

J Crohns Colitis. 2022 May 10;16(4):633-642.

Abstract

Background and Aims: Observational studies have suggested a bidirectional association between depression and inflammatory bowel disease (IBD), including Crohn's disease (CD) and ulcerative colitis (UC). However, it remains unclear whether observed associations are casual due to the difficulties of determining sequential temporality. We investigated the association between depression and IBD by using bidirectional two-sample Mendelian randomization (MR).

Methods: Independent genetic variants for depression and IBD were selected as instruments from published genome-wide association studies (GWAS) among individuals of predominantly European ancestry. Summary statistics for instrument-outcome associations were retrieved from three separate databases for both depression (Psychiatric Genomics Consortium, FinnGen, and UK Biobank), and IBD (the largest GWAS meta-analysis, FinnGen, and UK Biobank), respectively. MR analyses included inverse-variance weighted method, weighted-median estimator, MR-Egger regression, and sensitivity analyses of Steiger filtering and MR PRESSO. From either direction, analyses were performed per outcome database and were subsequently meta-analyzed using fixed-effect model.

Results: Genetically predicted depression (per log-odds ratio increase) was associated with a higher risk of IBD; odds ratios (95% confidence interval) for IBD, CD and UC were 1.20 (1.05, 1.36), 1.29 (1.07, 1.56) and 1.22 (1.01, 1.47) in a combined sample size of 693,183 (36,507 IBD cases), 212,172 (13,714 CD cases) and 219,686 (15,691 UC cases) individuals, respectively. In contrast, no association was observed between genetically influenced IBD and depression in 534,635 individuals (71,466 depression cases).

Conclusions: Our findings corroborated a causal association of depression on IBD, which may impact the clinical decision on the management of depression in patients with IBD. Though our results did not support a causal effect of IBD on depression, further investigations are needed to clarify the effect of IBD activity on depression (with different symptomology).

Introduction

Inflammatory bowel disease (IBD), comprising Crohn's disease (CD) and ulcerative colitis (UC), is characterized by chronic and progressive inflammation of the gastrointestinal tract and poses a high disease burden worldwide [1]. Despite the substantial progress in the diagnosis and management of IBD to achieve long-term remission, the causes leading to IBD are yet not fully elucidated. The pathogenesis of IBD is considered as a result of an interplay between genetic susceptibility and environmental risk factors, including but not limited to smoking, unfavorable diet, and lifestyle, which subsequently lead to an inappropriate intestinal immune activation and a proinflammatory intestinal microbiome [2-4]. Over the past decade, the association between depression and IBD (and vice versa) gained considerable interest [5,6], in light of the putative pathophysiological mechanism underlying the dysregulation of the brain-gut axis [7].

The estimated prevalence of comorbid depression or depressive symptoms in IBD patients is approximately 25% in a most recent meta-analysis [8], which is significantly higher than the estimated prevalence of 3.4% in the general population [9]. In line with this evidence, the incidence of depression in IBD patients is higher than in a matched population as early as 5 years before IBD diagnosis [10]. Several longitudinal studies suggested an association between depression and an elevated risk of IBD [11-14]. However, a recent population-based nested case-control study considering prodromal gastrointestinal (GI) symptoms prior to the diagnosis of depression showed that depression alone, i.e. in the absence of prior GI symptoms, is not associated with subsequent IBD risks [15]. Since depression and IBD both involve a vague and subtle onset, it is difficult to determine the temporal order of these two conditions. Thus, the associations described previously may have been partly due to reverse causation and/or residual confounding, which are often noted in observational studies. Therefore, the directionality and causality between depression and IBD remain unclear.

Mendelian randomization (MR) utilizes genetic variants identified through genome-wide association studies (GWAS), which are randomly allocated at conception as instrumental variables to investigate whether a lifetime exposure is causally associated with an outcome [16]. The most recent bidirectional MR study using IBD GWAS summary statistics conducted in the UK Biobank (UKB) did not reveal any possible association in either direction between depression and IBD [17]. The limited number of IBD cases in the UKB likely resulted in an underpowered sample size to detect any association. Therefore, in the present study, we aimed to investigate the potential bidirectional causal relationship between depression and IBD (including both UC and CD) by implementing a bidirectional MR study design using the most up-to-date and larger GWAS on IBD and depression.

Methods

We leveraged a bidirectional two-sample MR to assess the causal association between depression and IBD. The schematic overview of the study design and data sources are detailed in **Figure 1**. All data are publicly available GWAS summary statistics, and therefore, no additional ethical approval or informed consent was required. GWAS summary statistics were searched to extract lead-

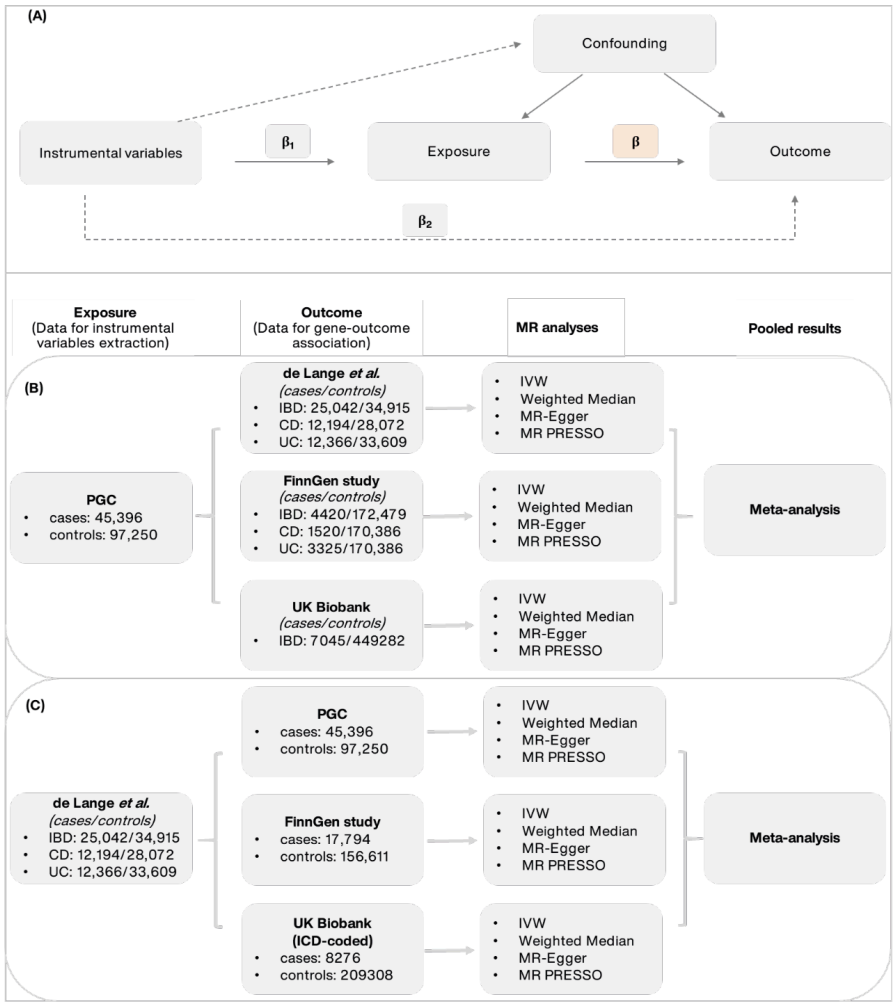


Figure 1 Schematic overview of the study design.

(A) Mendelian Randomization (MR) illustration. There are three principal assumptions in MR design, namely the genetic instrumental variables should 1) be associated with exposure; 2) be associated with outcome only via exposure; 3) not be associated with any measured or unmeasured confounding factors. β_1 and β_2 denote to the gene-exposure and gene-outcome association respectively; β represents the causal association between exposure and outcome, which can be estimated by β_1 / β_2. (B) MR study from depression to IBD: independent SNPs for depression were identified as instrumental variables, whereas summary statistics of gene-IBD associations were retrieved separately from the GWAS performed by de Lange et al., FinnGen, and UK Biobank (for IBD only). MR analyses were conducted per outcome database and were subsequently meta-analyzed to generate pooled estimates. (C) MR study from IBD to depression: SNPs for IBD were identified as instrumental variables, whereas summary statistics of gene-depression associations were retrieved separately from PGC, FinnGen study, and UK Biobank (GWAS conducted with ICD-coded depression cases only). MR analyses were performed per outcome database and were subsequently meta-analyzed. IBD: inflammatory bowel disease; PGC: Psychiatric Genomics Consortium. CD: Crohn's Disease; UC: Ulcerative colitis; IVW: Inverse-variance weighted.

ing SNPs (single nucleotide polymorphisms) associated with depression or IBD as genetic instrumental variables. Gene-outcome associations were separately retrieved from three databases for both depression and IBD: 1) large-scale GWAS meta-analysis efforts; 2) FinnGen (data freeze 4); and 3) the UKB. In the FinnGen study, GWAS were performed across a broad spectrum of phenotypes including depression and IBD; the analyses were adjusted for age, sex, principal components, and genotype batch effect. Phenotype definitions were based on the International Statistical Classification of Diseases and Related Health Problems (ICD) coded hospital discharge or death. Detailed information regarding participants, genotype platforms, and statistical analysis protocols are available at the FinnGen website (*https://www.finngen.fi/en/*).

Selection and description of the sources of the genetic instrumental variables

Depression is an exceedingly heterogeneous condition with many different measures for identification in the previous studies, ranging from subjective self-reported symptoms or help-seeking to clinical diagnosis. Genetic variants that were identified from GWAS by using less precisely defined phenotypes including help-seeking, might not be specific to depression per se [18]. Consequently, this may reduce statistical power, particularly when depression is used as an exposure. In addition, diagnosis-based depression, such as depression defined by the International Statistical Classification of Diseases and Related Health Problems (ICD) or Diagnostic and Statistical Manual of Mental Disorders (DSM) clinical guideline, showed statistically significant relation to IBD in the UKB, but not depression defined by other means such as help-seeking behavior or self-reported symptoms [17].

Two GWAS meta-analyses on depression were conducted recently by Howard et al. and Wray et al., respectively [19,20]. Despite the larger total sample size of 807,553 individuals in the study conducted by Howard et al., self-reported help-seeking behavior for mental health difficulties accounted for more than 70% of cases. We therefore used the GWAS summary statistics from Wray et al, which included participants from 29 cohorts in Psychiatric Genomics Consortium (PGC) and 6 additional cohorts including UKB and 23andMe. In each cohort, data were processed following PGC "ricopili" pipeline or using comparable procedures when applicable. Due to the data restriction, in the present study, we used the available GWAS summary statistics data that did not include UKB or 23andMe, consisting of 45,396 cases and 97,250 controls. Individuals who contributed to the data all met the international consensus criteria for a lifetime diagnosis of depression established using structured diagnostic instruments from assessments by trained interviewers, clinician-administered checklists, or medical record review. Since no SNP-depression association reached the genome-wide association threshold (p-value < 5×10^{-8}), a suggestive significant level (p-value < 1×10^{-6}) was used to extract instrumental variables, as has been adopted in the previous study in disentangling bidirectional relationships between physical activity and depression in MR analyses [21]. Linkage disequilibrium between all SNPs was based on the European 1000 Genome Project reference panel. Independent SNPs were selected by linkage disequilibrium clumping ($r^2 > 0.001$) retaining the one with the smallest P-value.

Summary statistics for IBD was obtained from the latest meta-analysis GWAS (de Lange et al.), which contained a total sample size of 59,957 participants of predominantly European ancestry (cases / controls for IBD: 25,042 / 34,915; UC: 12,366 / 33,609; CD: 12,194 / 28,072) [22]. IBD was diagnosed by accepted radiological, endoscopic, and histopathological evaluation and all included cases fulfilled clinical criteria for IBD. In every single cohort, GWAS was performed using an additive model conditioning on the first ten principal components. We selected SNPs that were significantly associated with IBD (p-value< 5×10^{-8}) and further pruned all SNPs by linkage disequilibrium clumping ($r^2 > 0.001$). To avoid any potential pleiotropic instruments, we excluded 43 SNPs that were associated with more than one phenotype, and 176 SNPs (IBD: 74; UC: 42, CD: 60) remained for the MR analyses.

SNP-outcome data sources

Gene-outcome associations for depression were obtained from three separate databases: 1) the PGC data from Wray et al. as described previously; 2) FinnGen; and 3) the UKB. Depression in FinnGen was defined as depressive episode and recurrent depressive disorder by ICD criteria consisting of 17,794 cases and 156,611 controls. In the UKB, since multiple definitions for depressive disorders were available [18], to minimize the heterogeneity of depression definition, we only considered the ICD-coded (ICD-10 primary and secondary codes for depression) depression as the outcome. GWAS analyses on ICD-coded depression in the UKB included 8,276 cases and 209,308 controls (with a prevalence of 3.80%), adjusted covariates of age, sex, genotyping array, and 8 principal components [23].

Gene-outcome associations for IBD were drawn from three separate databases: 1) latest meta-analysis GWAS by de Lange et al.; 2) FinnGen; and 3) UKB (only IBD data were available). The de Lange et al. study has been described in detail in the previous section. In the FinnGen study, CD and UC cases were defined using their corresponding ICD codes, and IBD is a term comprised of CD, UC, and colitis of indeterminate. In total, the number of cases and controls were 4420 / 172,479 for IBD, 3325 / 170,386 for UC and 1520 / 170,386 for CD, respectively. In the UKB, IBD was composed of CD (UKB data field: 131627) and UC (UKB data field: 131629), which were determined from either a death register, self-reported, hospital admission, or primary care record for the corresponding disease, resulting in a total of 7045 cases and 449,282 controls [17]; GWAS analyses were adjusted for sex, age, and 20 principal components; GWAS on UC and CD were not available in the UKB.

Instrumental strength and power calculation

F-statistics, calculated as $(beta/se)^2$, was computed to quantify the strength of instruments, and a value of above 10 was considered sufficient. The proportions of variance explained by exposures (R^2) were calculated using the Mangrove package in R [24], setting the prevalence of 3.4% for depression and 0.3%, 0.14%, 0.2% respectively for IBD, CD, and UC [1,9]. Statistical power was calculated via the tool for binary outcomes in MR studies (*https://github.com/kn3in/mRnd*) [25], where the alpha level was set to 0.05.

Mendelian Randomization (MR)

Before analysis, we first harmonized exposure and outcome data to make alignments on effect alleles to the forward strand, if it is specified or could be inferred based on the allele frequency. Palindromic genetic variants were discarded for further MR analyses [26]. We used inverse-variance weighted (IVW) meta-analysis as the main analysis to combine the SNP-specific estimates calculated using Wald ratios, assuming no directional pleiotropic effect of each SNP [27]. In addition, we performed several sensitivity analyses to assess pleiotropy and potential genetic outliers. A Weighted-Median estimator can provide a reliable estimate if more than 50% of the instrumental variables are valid [28]. In MR-Egger regression, the slope provides a causal estimate of an exposure on an outcome if the instrument strength independent of direct effect assumption is met; additionally, the intercept of a MR-Egger regression deviates from zero indicates pleiotropy. MR-Egger is statistically less efficient (i.e. with wider confidence intervals) but provides a causal estimate (i.e., the regression slope) that is corrected for directional horizontal pleiotropy [29]. Furthermore, MR Pleiotropy RESidual Sum and Outlier (MR PRESSO) test is applied to detect significant outliers and correct for horizontal pleiotropic effect through outlier removal. The global test evaluates whether horizontal pleiotropy among all instruments is present [30]. We further examined the heterogeneity among all SNPs within each database using Cochran's Q test statistic and generated scatter plots of SNP-exposure association versus SNP-outcome association to visualize MR results. Leave-one-out analysis was performed by excluding each SNP at a time sequentially and an IVW method was performed on the remaining SNPs to assess the potential influence of a particular variant on the estimates.

Additionally, when SNPs are selected from very large GWAS, these instrumental variables may have effects on the other downstream traits of the trait of interest, such as effects directly on the outcome. Thus, determining whether a SNP is primarily associated with the exposure of interest or with the outcome could be challenging. If those variants that show stronger associations to outcomes than to exposures are used in the MR analyses, the result may (erroneously) imply that the exposure and outcome are causally related due to reverse causation. Therefore, we applied MR Steiger filtering to test the direction of causality for each instrumental variable on exposure and outcome. Steiger filtering assumes that a valid instrumental variable should explain more variation in the exposure than in the outcome; if an instrumental variable meets the criterion, the direction of this instrument is "TRUE", while if it does not, it is "FALSE" [31]. After removing those SNPs with "FALSE" direction, we repeated all MR analyses.

MR results were expressed as odds ratios (OR) with the corresponding 95% confidence intervals (CIs) on outcome risk of corresponding unit changes in an exposure. Given that all exposure variables in the current analyses are binary, the final effect estimates were interpreted as OR on the risk of outcome per log-OR change in an exposure. From each direction, MR analyses were firstly performed in each outcome database separately and the individual estimates were subsequently combined using fixed-effects meta-analysis.

All the analyses were performed using R (v3.6.3) statistical software (The R Foundation for Statistical Computing, Vienna, Austria). MR analyses were per-

formed using the R-based package "TwoSampleMR" and meta-analysis were conducted using the "meta" package.

Results

Depression to IBD

Nineteen independent SNPs were identified as genetic instrumental variables for depression, which explained 0.41% of the total variation; the median (minimum, maximum) F statistics were 26.2 (23.9, 37.4) (**Supplementary Table 1**). Detailed information about the 19 genetic variants is listed in **Supplementary Table 2**.

In the meta-analyses of estimates from IVW, pooled ORs for IBD, CD, and UC of genetically predicted per log-OR increase in depression were 1.20 (95% confidence interval CI: 1.05, 1.36), 1.29 (95% CI: 1.07, 1.56), 1.22 (95% CI: 1.01, 1.47), respectively (**Figure 2**).

For each database, in sensitivity analyses (**Supplementary Table 3**), estimates obtained from Weighted-Median did not differ substantially compared to those from IVW (**Supplementary Figure 1**). No pleiotropic effect was detected by MR-Egger intercept. Potential outliers were identified by MR PRESSO for IBD in the de Lange et al. database and UKB, as well as for UC in the de Lange et al. database, which resulted in potential pleiotropy assessed by the global test. However, results remained similar after outlier correction. Heterogeneity of each instrument estimation evaluated by Cochran's Q test statistics was detected, but only in those databases with outliers indicated by MR PRESSO. Individual SNP effects and combined effects from each MR method per outcome database are visualized in scatter plots (**Figure 3**). Leave-one-out plots suggested that the associations were unlikely driven by certain extreme SNPs (**Supplementary Figure 2**). Steiger filtering detected 4 SNPs that were with a "FALSE" direction, which was likely to primarily be associated with IBD rather than depression (**Supplementary Table 4**). These included rs2060886 in the intronic region of TCF4 in all the three IBD-related traits from de Lange data; rs1936365 in the intronic region of PGBD1 in all the three IBD-related traits in FinnGen while only CD in UKB; rs1950829 in LRFN5 for CD in FinnGen; and rs1491473 in the intronic region of LINC00861 for UC in FinnGen. After removing those four SNPs, analyses were repeated for all methods, and estimates were minimally influenced by applying Steiger filtering (**Supplementary Table 5**).

IBD to Depression

A total of 70 independent genetic variants reached a genome-wide significant level with IBD, which explained about 3.4% of the total variation, whereas 55 and 41 genetic instruments were retained for CD and UC, which accounted for 4.9% and 3.0% of the total variation. The summary and detailed information about the variants for each exposure are presented in **Supplementary Table 1 and 6**.

Overall, in the primary analyses using IVW, we found no association between genetically determined IBD (including both CD and UC) and depression in any individual outcome database, except for a minimal effect of IBD on depression

in FinnGen. Combined ORs of IBD, CD, and UC on depression was 1.01 (95% CI: 1.00, 1.03), 1.00 (95% CI: 0.99, 1.01), 1.02 (95% CI: 0.99, 1.05), respectively (**Figure 4**).

In each outcome database, sensitivity analyses using the Weighted-Median estimator were consistently comparable to the estimates from IVW (**Supplementary Figure 3**). While the MR-Egger intercept test suggested no evidence of pleiotropy, MR-PRESSO global test indicated the presence of horizontal pleiotropic outliers from CD to depression in the FinnGen study (p = 0.04) and UC to depression in PGC (p = 0.03); however, after correcting for outlier by MR PRESSO, results remained similar with those estimates from IVW (**Supplementary Table 7**). No notable heterogeneity was detected by Cochran's Q statistics across single instrument effects within each database. Scatter plots (**Figure 5**) present the individual SNP effect and combined effect from each method per outcome dataset. After removing one SNP at a time, the results remained consistent in the leave-one-out analyses (**Supplementary Figure 4**). Steiger filtering indicated that all genetic instrumental variables used for IBD explained more variance in IBD than in depression in any database (**Supplementary Table 6**).

Discussion

In the present study, we evaluated the bidirectional associations between depression and IBD using MR. We found evidence that genetic liability to depression was associated with an increased risk of IBD, CD, and UC, while genetic liability to IBD, or any subtype was not associated with depression.

Previous observational studies have suggested that depression might be a risk factor for IBD. Specifically, in The Health Improvement Network cohort with 403,665 incident IBD cases, depression was associated with a higher risk of both incident CD and UC [13]. Similarly, a study using the data from 152,461 women aged 29 to 72 years in the Nurses' Health Study also found that depressive symptoms were associated with an increased risk of CD, but not UC [14]. On the contrary, depression alone, in the absence of prior GI symptoms, was not associated with subsequent development of IBD in 10,829 UC cases, 4531 CD cases, and 15,360 controls [15]. Up to now, the only MR study conducted to disentangle the bidirectional relationship between depression and IBD used IBD GWAS summary statistics (7045 cases and 449,282 controls) obtained from UKB, which revealed no association in either direction [17]. The small number of IBD cases in the UKB may account for this null association. The same study identified causal associations between depression and other gastrointestinal disorders for which a much larger number of cases were used in the GWAS. This corroborated that the absence of an association between depression and IBD was likely due to a lack of statistical power. Consistent with the previous study, we did not find an association between depression and IBD using only UKB data. However, when the estimates were combined with those from the other two databases in the meta-analysis, we did observe a causal effect of depression on IBD (**Figure 2**).

The biological connection between depression and IBD has not been fully elucidated yet. Currently, intestinal inflammation and its concomitant microbial dysbiosis have been implicated in the etiology of IBD [32,33], which may explain

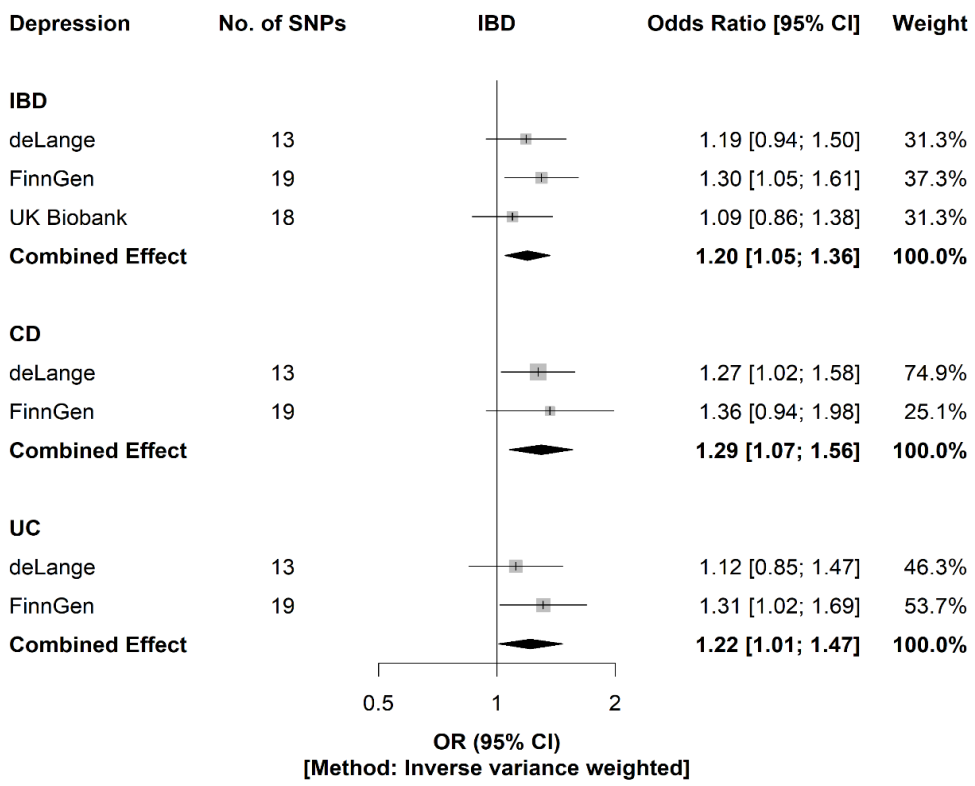


Figure 2 Association of Depression and IBD in MR analyses

Estimated ORs represent the effect of per log-OR increase in depression on IBD, obtained from an Inverse-variance weighted analysis, per outcome database separately and combined over the three databases for IBD and two databases (data are not available in the UK Biobank) for CD and UC using fixed-effect meta-analyses.

IBD: Inflammatory Bowel Disease; CD: Crohn's Disease; UC: Ulcerative Colitis; OR: Odds Ratio; CI: Confidence Interval.

the causal association between depression and IBD that we observed in the current study. Depression contributes to intestinal inflammation by modulating the psycho-neuro-endocrine-immune system in the brain-gut axis. Numerous complex mechanistic pathways underpinning brain-gut (microbiome) interaction have been extensively explored previously [34-37]. In brief, depression could activate the hypothalamic-pituitary-adrenal axis, which subsequently leads to down-regulation of the corticotropin-releasing factor system, thereby accelerating chronic inflammation and stimulating the immune response. In addition, the autonomic nervous system is also functionally involved in stress-mediated alterations. The activation of the sympathetic nervous system plays a proinflammatory role that could invoke an enhanced secretion of catecholamines. The combination of increased sympathetic outflow and adrenomedullary activity could then stimulate mast cells and macrophages to release inflammatory effectors of cytokines. In parallel, the vagus nerve has an anti-inflammatory function and could be inhibited

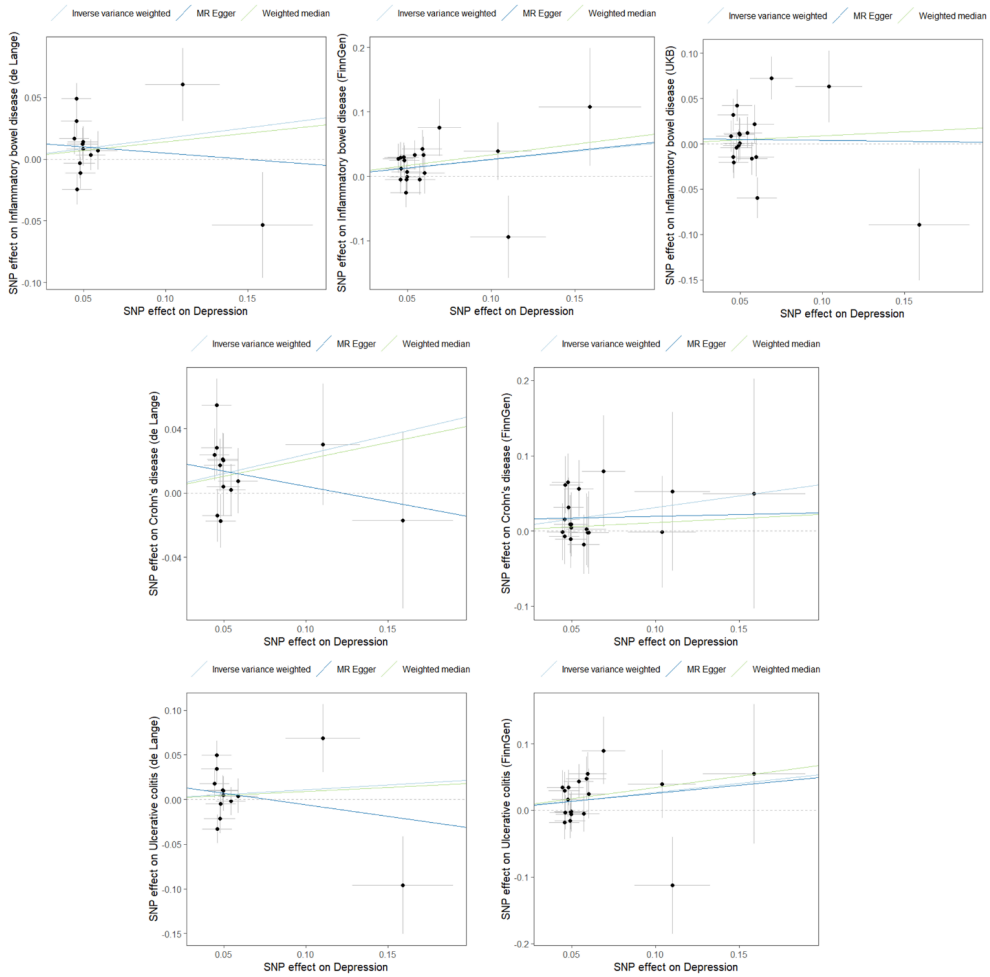


Figure 3 Scatter plot of MR analyses from Depression to Inflammatory Bowel Disease in each database

X axes represent the genetic instruments-depression associations and Y axes represent genetic instruments-IBD associations from different outcome databases. Black dots denote to the genetic instruments included in the primary MR analyses. red: Inverse-variance weighted; blue: Weighted-Median estimator; green: MR Egger. Due to the same estimate from the Inverse-variance weighted and Weighted-Median estimator methods in some analyses, those figures only contain two lines. However, the color of the overlapped lines is darker than the Weighted-Median estimator.
IBD: Inflammatory Bowel Disease; CD: Crohn's Disease; UC: Ulcerative Colitis.

by proinflammatory cytokines (such as IL-1, IL-6, and TNF-α) released from the intestinal mucosa via a decreased efferent outflow. These inflammatory profile changes are related to regional gut motility, luminal secretion, visceral hypersensitivity, and elevated intestinal permeability. As a result of the compromised epithelial barrier, intestinal permeability promotes microbiome translocation and further activates the immune response. All these changes brought about

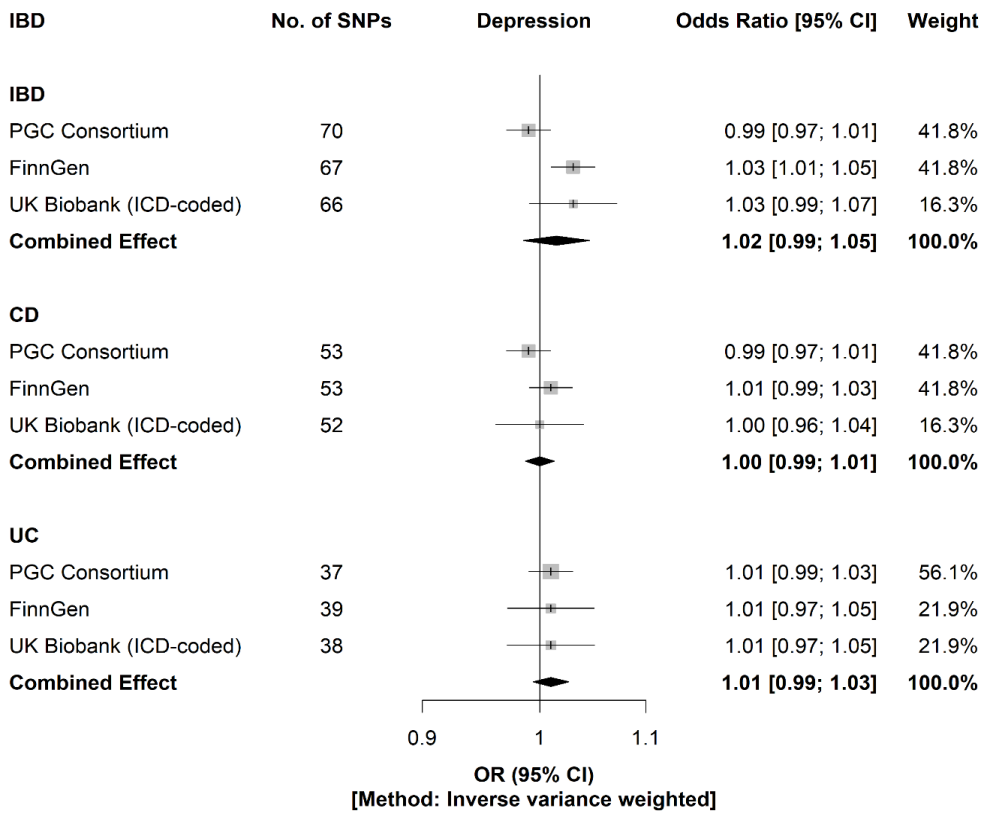


Figure 4 Association of IBD and depression in MR analyses

Estimated ORs represents the effect of per log-OR increase IBD on depression, obtained from an Inverse-variance weighted analysis, per outcome database separately and combined over the three databases using fixed-effect meta-analyses.

IBD: Inflammatory Bowel Disease; CD: Crohn's Disease; UC: Ulcerative Colitis; PGC: Psychiatric Genomics Consortium. OR: Odds Ratio; CI: Confidence Interval.

by depression may disrupt the homeostasis of the gastrointestinal tract and consequently lead to IBD.

Strength and limitations

There are two main strengths worth noting in the current study. First, we extracted gene-outcome associations from three independent sources of GWAS summary statistics, which included large sample sizes, particularly cases, for each exposure, with 534,635 participants (71,466 cases) for depression, 693,183 (36,507 cases) for IBD, 212,172 (13,714 cases) for CD, and 219,686 (15,691 cases) for UC. The effect estimations from various data sources all pointed to the same direction, and meta-analyses further strengthened the estimation. Second, we used clinical diagnosis-based depression as the exposure. A previous study indicated that effect size estimations between depression and IBD varied according to the definition used for depression, and self-reported depression tended to attenuate

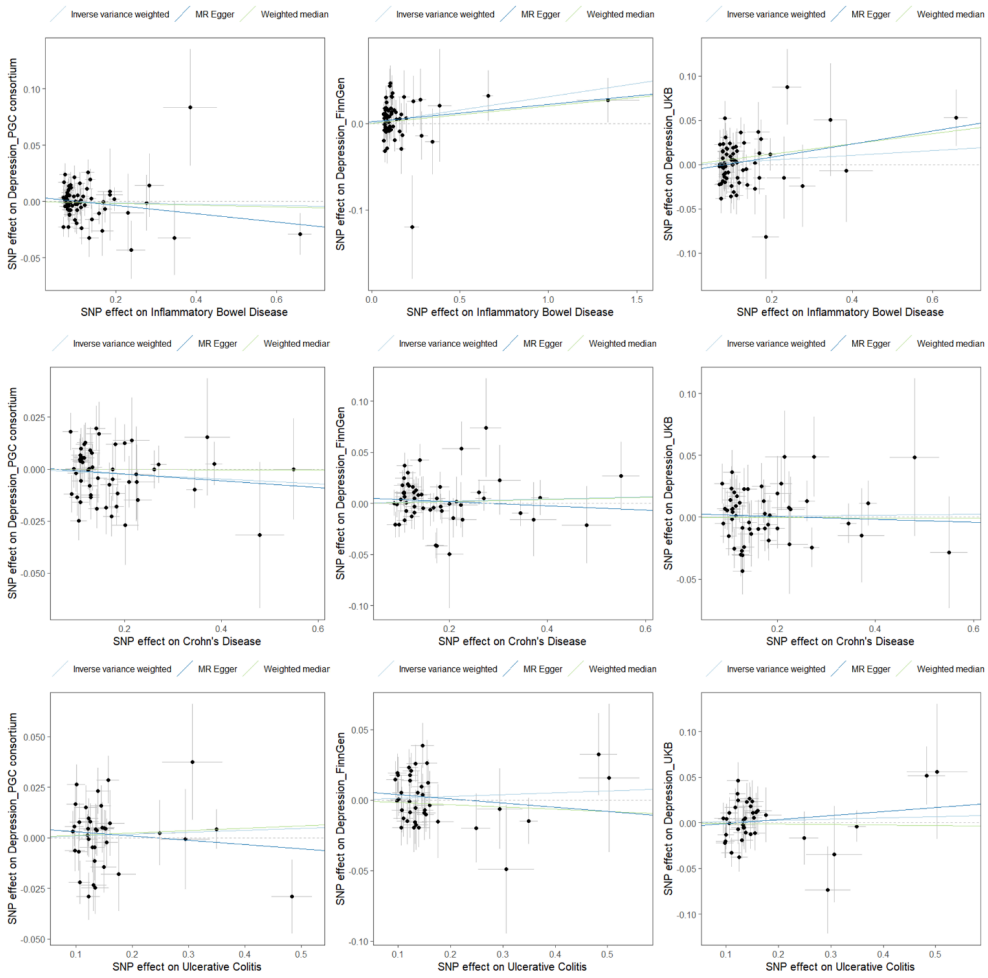


Figure 5 Scatter plot of MR analyses from Inflammatory Bowel Disease to Depression in each database

The figure caption is the same as in Figure 3.

association estimates [17]. Interestingly, 19 SNPs used in the analyses explained 0.41% of the total variation in clinically diagnosed depression, which was comparable to the total variations explained by 102 genetic variants (approximately 0.5%) identified by the GWAS using broad depression [19,38].

Some limitations should be acknowledged. Firstly, all data involved in the analyses were derived primarily from individuals of European ancestry, limiting the generalizability of our findings to other ethnic groups. Secondly, the statistical power may still be insufficient to detect an effect of depression on IBD in a single dataset, as indicated in the FinnGen study (**Supplementary Figure 5**). Thirdly, although we found no causal effect of IBD on depression, we only considered the dichotomous IBD diagnosis, i.e. the incidence, rather than the IBD disease

course. Previous studies showed that both the prevalence and incidence of depression are significantly higher among patients with active IBD than patients with inactive diseases [39,40]. While we acknowledge IBD has a dynamic natural course characterized by alternating periods of remission and relapse, and its episodic flares occur randomly and are largely unpredictable, dissecting the genetic makeup that is associated with IBD activity is still challenging. Due to the lack of GWAS on IBD disease activity to date, we were unable to investigate the effect of IBD activity on depression using the MR approach. However, the genetic determinants of IBD susceptibility may be distinct from mechanisms underlying disease activity. It might not be the disease per se, but rather the disease activity that is pivotal to depression. Therefore, we could not rule out the possibility of disease activity, but not incident IBD, as a causal risk factor for depression. In addition, depression is very heterogeneous and exhibits diverse symptoms. Patients with the same major depressive disorder diagnosis according to DSM-V may experience very different symptom profiles. However, depression is used as a binary variable in the GWAS, without taking symptomology into account. Therefore, it remains possible that IBD would have an effect on specific depression dimensions. We must also keep in mind that, IBD being rare, it can only be causal for a small fraction of patients with depression. Fourthly, when MR analyses were conducted from the direction of IBD to depression, there may be some overlap when depression data from UKB were used as outcome. The IBD GWAS included individuals from UK10K (4,686 IBD cases and 3,781 controls), which may have contributed to the UKB depression data. This accounted for 14% of the total population in the IBD GWAS data and represented the maximum overlap between the two datasets. Nevertheless, with such a low proportion of sample overlap, the bias is expected to be neglectable [41]. Additionally, the results remained consistent if using only PGC and FinnGen data, with an OR (95% CI) of 1.24. (1.06, 1.45). Finally, subgroup analyses are not feasible due to the use of summary statistics, particularly by sex, where previous studies pinpointed a higher prevalence of depressive symptoms in women than men [8], and specifically an increased risk for CD, but not UC, among women [14].

Clinical implications

To date, no consensus has been reached about the role of depression in the counseling and management of IBD. In the American College of Gastroenterology (ACG) clinical guideline for preventive health care, recommendations of screening for depression in patients with IBD is conditional due to a low level of evidence [42]. Meanwhile, in the latest update on the Selecting Therapeutic Targets in Inflammatory Bowel Disease (STRIDE), the absence of health-related anxiety and depression is removed as a therapeutic goal for treat-to-target strategies in IBD given the low (37%) endorsement [43]. Similarly, psychological therapies in IBD patients are rarely recommended to patients with IBD as adjunctive therapy to alleviate symptoms and improve quality of life given the very low quality of evidence by the British Society of Gastroenterology consensus guidelines [44].

The observed causal effect of depression on IBD identified in our study should raise the awareness of the involvement of tackling depression to achieve a better clinical outcome in IBD. On one hand, clinicians should raise the awareness of an index of suspicion about IBD in patients with depression; particularly, gastro-

intestinal symptoms in patients with depression should be spotted, as they are often considered as complaints of depression, such as chronic diarrhea, etc., but not as an early clinical manifestation of IBD and lead to "diagnostic overshadowing" [45]. On the other hand, the implementation of depression screening and the involvement of appropriate treatment for depression into the routine practice of IBD patients might help optimize the management of IBD and confer possibilities leading to better clinical outcomes for IBD patients.

Conclusion

Our findings corroborated a causal association between depression and IBD, which may have an influence on the clinical decision in the management of depression in patients with IBD. Though our results did not support a causal effect of IBD on depression, further investigations are needed to clarify the effect of IBD activity on depression (with different symptomology).

References

1. S. C. Ng, H. Y. Shi, N. Hamidi, et al. Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies. Lancet 2018; 390(10114): 2769-78.
2. H. Huang, M. Fang, L. Jostins, et al. Fine-mapping inflammatory bowel disease loci to single-variant resolution. Nature 2017; 547(7662): 173-8.
3. A. N. Ananthakrishnan. Epidemiology and risk factors for IBD. Nat Rev Gastroenterol Hepatol 2015; 12(4): 205-17.
4. G. P. Ramos, K. A. Papadakis. Mechanisms of Disease: Inflammatory Bowel Diseases. Mayo Clin Proc 2019; 94(1): 155-65.
5. D. J. Gracie, E. A. Guthrie, P. J. Hamlin, A. C. Ford. Bi-directionality of Brain-Gut Interactions in Patients With Inflammatory Bowel Disease. Gastroenterology 2018; 154(6): 1635-46.e3.
6. L. Keefer, S. V. Kane. Considering the Bidirectional Pathways Between Depression and IBD: Recommendations for Comprehensive IBD Care. Gastroenterol Hepatol (N Y) 2017; 13(3): 164-9.
7. D. J. Gracie, P. J. Hamlin, A. C. Ford. The influence of the brain-gut axis in inflammatory bowel disease and possible implications for treatment. Lancet Gastroenterol Hepatol 2019; 4(8): 632-42.
8. B. Barberio, M. Zamani, C. J. Black, E. V. Savarino, A. C. Ford. Prevalence of symptoms of anxiety and depression in patients with inflammatory bowel disease: a systematic review and meta-analysis. Lancet Gastroenterol Hepatol 2021; 6(5): 359-70.
9. H. Ritchie, Roser M. Mental Health. https://ourworldindata.org/mental-health (accessed January 22nd 2021).
10. R. A. Marrie, R. Walld, J. M. Bolton, et al. Rising incidence of psychiatric disorders before diagnosis of immune-mediated inflammatory disease. Epidemiol Psychiatr Sci 2019; 28(3): 333-42.
11. N. C. Momen, O. Plana-Ripoll, E. Agerbo, et al. Association between Mental Disorders and Subsequent Medical Conditions. N Engl J Med 2020; 382(18): 1721-31.
12. A. Mikocka-Walus, V. Pittet, J. B. Rossel, R. von Känel. Symptoms of Depression and Anxiety Are Independently Associated With Clinical Recurrence of Inflammatory Bowel Disease. Clin Gastroenterol Hepatol 2016; 14(6): 829-35.e1.
13. A. D. Frolkis, I. A. Vallerand, A. A. Shaheen, et al. Depression increases the risk of inflammatory bowel disease, which may be mitigated by the use of antidepressants in the treatment of depression. Gut 2019; 68(9): 1606-12.
14. A. N. Ananthakrishnan, H. Khalili, A. Pan, et al. Association between depressive symptoms and incidence of Crohn's disease and ulcerative colitis: results from the Nurses' Health Study. Clin Gastroenterol Hepatol 2013; 11(1): 57-62.
15. J. Blackwell, S. Saxena, I. Petersen, et al. Depression in individuals who subsequently develop inflammatory bowel disease: a population-based nested case-control study. Gut 2021; 70(9): 1642-8.
16. G. Davey Smith, G. Hemani. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. Hum Mol Genet 2014; 23(R1): R89-98.

17. Y. Wu, G. K. Murray, E. M. Byrne, J. Sidorenko, P. M. Visscher, N. R. Wray. GWAS of peptic ulcer disease implicates Helicobacter pylori infection, other gastrointestinal disorders and depression. Nat Commun 2021; 12(1): 1146.
18. N. Cai, J. A. Revez, M. J. Adams, et al. Minimal phenotyping yields genome-wide association signals of low specificity for major depression. Nat Genet 2020; 52(4): 437-47.
19. D. M. Howard, M. J. Adams, T. K. Clarke, et al. Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. Nat Neurosci 2019; 22(3): 343-52.
20. N. R. Wray, S. Ripke, M. Mattheisen, et al. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. Nat Genet 2018; 50(5): 668-81.
21. K. W. Choi, C. Y. Chen, M. B. Stein, et al. Assessment of Bidirectional Relationships Between Physical Activity and Depression Among Adults: A 2-Sample Mendelian Randomization Study. JAMA Psychiatry 2019; 76(4): 399-408.
22. K. M. de Lange, L. Moutsianas, J. C. Lee, et al. Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. Nat Genet 2017; 49(2): 256-61.
23. D. M. Howard, M. J. Adams, M. Shirali, et al. Genome-wide association study of depression phenotypes in UK Biobank identifies variants in excitatory synaptic pathways. Nat Commun 2018; 9(1): 1470.
24. L. Jostins. Mangrove: Risk Prediction on Trees. https://cranr-projectorg/web/packages/Mangrove/indexhtml 2017.
25. M. J. Brion, K. Shakhbazov, P. M. Visscher. Calculating statistical power in Mendelian randomization studies. Int J Epidemiol 2013; 42(5): 1497-501.
26. F. P. Hartwig, N. M. Davies, G. Hemani, G. Davey Smith. Two-sample Mendelian randomization: avoiding the downsides of a powerful, widely applicable but potentially fallible technique. Int J Epidemiol 2016; 45(6): 1717-26.
27. S. Burgess, A. Butterworth, S. G. Thompson. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol 2013; 37(7): 658-65.
28. J. Bowden, G. Davey Smith, P. C. Haycock, S. Burgess. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. Genet Epidemiol 2016; 40(4): 304-14.
29. J. Bowden, G. Davey Smith, S. Burgess. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol 2015; 44(2): 512-25.
30. M. Verbanck, C. Y. Chen, B. Neale, R. Do. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. Nat Genet 2018; 50(5): 693-8.
31. G. Hemani, K. Tilling, G. Davey Smith. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. PLoS Genet 2017; 13(11): e1007081.
32. J. Halfvarson, C. J. Brislawn, R. Lamendella, et al. Dynamics of the human gut microbiome in inflammatory bowel disease. Nat Microbiol 2017; 2: 17004.
33. E. A. Franzosa, A. Sirota-Madi, J. Avila-Pacheco, et al. Gut microbiome structure and metabolic activity in inflammatory bowel disease. Nat Microbiol 2019; 4(2): 293-305.

34. V. Osadchiy, C. R. Martin, E. A. Mayer. The Gut-Brain Axis and the Microbiome: Mechanisms and Clinical Implications. Clin Gastroenterol Hepatol 2019; 17(2): 322-32.
35. B. L. Bonaz, C. N. Bernstein. Brain-gut interactions in inflammatory bowel disease. Gastroenterology 2013; 144(1): 36-49.
36. A. D. Kostic, R. J. Xavier, D. Gevers. The microbiome in inflammatory bowel disease: current status and the future ahead. Gastroenterology 2014; 146(6): 1489-99.
37. E. A. Mayer, T. Savidge, R. J. Shulman. Brain-gut microbiome interactions and functional bowel disorders. Gastroenterology 2014; 146(6): 1500-12.
38. Tanya M. Teslovich, Kiran Musunuru, Albert V. Smith, et al. Biological, clinical and population relevance of 95 loci for blood lipids. Nature 2010; 466(7307): 707-13.
39. A. Mikocka-Walus, S. R. Knowles, L. Keefer, L. Graff. Controversies Revisited: A Systematic Review of the Comorbidity of Depression and Anxiety with Inflammatory Bowel Diseases. Inflamm Bowel Dis 2016; 22(3): 752-62.
40. B. Kochar, E. L. Barnes, M. D. Long, et al. Depression Is Associated With More Aggressive Inflammatory Bowel Disease. Am J Gastroenterol 2018; 113(1): 80-5.
41. S. Burgess, N. M. Davies, S. G. Thompson. Bias due to participant overlap in two-sample Mendelian randomization. Genet Epidemiol 2016; 40(7): 597-608.
42. F. A. Farraye, G. Y. Melmed, G. R. Lichtenstein, S. V. Kane. Corrigendum: ACG Clinical Guideline: Preventive Care in Inflammatory Bowel Disease. Am J Gastroenterol 2017; 112(7): 1208.
43. D. Turner, A. Ricciuto, A. Lewis, et al. STRIDE-II: An Update on the Selecting Therapeutic Targets in Inflammatory Bowel Disease (STRIDE) Initiative of the International Organization for the Study of IBD (IOIBD): Determining Therapeutic Goals for Treat-to-Target strategies in IBD. Gastroenterology 2021; 160(5): 1570-83.
44. C. A. Lamb, N. A. Kennedy, T. Raine, et al. British Society of Gastroenterology consensus guidelines on the management of inflammatory bowel disease in adults. Gut 2019; 68(Suppl 3): s1-s106.
45. S. Jones, L. Howard, G. Thornicroft. 'Diagnostic overshadowing': worse physical health care for people with mental illness. Acta Psychiatr Scand 2008; 118(3): 169-71.

Supplementary materials

Supplementary Table 1 Summary of Genetic Instruments identified for MR Analyses

Supplementary Table 2 Instrumental variables for depression (significant level of p < 1e-06)

Supplementary Table 3 Effect estimates of the associations of genetic instrumental variables for depression and risk of inflammatory bowel disease

Supplementary Table 4 Steiger filtering results from depression to IBD

Supplementary Table 5 Mendelian randomization results following Steiger filtering of genetic instrumental variables for depression and risk of inflammatory bowel disease

Supplementary Table 6 Instrumental variables for inflammatory bowel disease (significant level of p < 5e-08)

Supplementary Table 7 Effect estimates of the associations of genetic instrumental variables for inflammatory bowel disease and risk of depression

Supplementary Figure 1 Estimated ORs of the effect of per log-OR of depression on IBD, obtained from Weighted-median Estimator per outcome database separately and combined over the three databases for IBD and two databases (data are not available in the UK Biobank) for CD and UC using fixed-effect meta-analyses

Supplementary Figure 2 Leave-one-out plot of MR analyses from Depression to Inflammatory Bowel Disease in each database

Supplementary Figure 3 Estimated ORs for the effect of the effect of log-OR change of IBD on depression, obtained from Weighted-median estimator per outcome database separately and combined over the three databases using fixed-effect meta-analyses.

Supplementary Figure 4 Leave-one-out plot of MR analyses from Inflammatory Bowel Disease to Depression in each database

Supplementary Figure 5 Statistical power of MR analysis; (A) Depression to IBD. Outcome data were from de Lange et al. (black), FinnGen (red), and UK Biobank (blue) for IBD only; (B) IBD to depression. Outcome data were from PGC (black), FinnGen (red), and UK Biobank (blue). Dashed grey line indicates statistical power of 0.8. Depression in the UK Biobank are from GWAS conducted with ICD-coded cases only.

The Supplementary materials for this article can be found online at:

https://academic.oup.com/ecco-jcc/article-lookup/doi/10.1093/ecco-jcc/jjab191

CHAPTER 9

Systemic inflammatory markers in relation to cognitive function and measures of brain atrophy: a Mendelian randomization study

Jiao Luo, Saskia le Cessie, Gerard Jan Blauw, Claudio Franceschi, Raymond Noordam, Diana van Heemst

Geroscience. 2022 Jun 11. doi: 10.1007/s11357-022-00602-7. Epub ahead of print

Abstract

Background: Observational studies have implied associations between multiple cytokines and cognitive decline, anti-inflammatory drugs however did not yield any protective effects on cognitive decline. We aimed to assess the associations of systemic inflammation, as measured with multiple cytokine and growth factors concentrations with cognitive performance and brain atrophy using two-sample Mendelian randomization (MR).

Methods: Independent genetic instruments (p<5e-8 and p<5e-6) for 41 systemic inflammatory markers were retrieved from a genome-wide association study conducted in 8293 Finnish participants. Summary statistics for gene-outcome associations were obtained for cognitive performance (N = 257,841) and for brain atrophy measures of cerebral cortical surface area and thickness (N = 51,665) and hippocampal volume (N = 33,536). To rule out the heterogeneity in the cognitive performance, we additionally included three domains: the fluid intelligence score (N = 108,818), prospective memory result (N = 111,099), and reaction time (N = 330,069). Main results were computed by inverse-variance weighting, and sensitivity analyses taking pleiotropy and invalid instruments into account were performed by using Weighted-median estimator, MR-Egger and MR PRESSO.

Results: After correcting for multiple testing using false discovery rate, only genetically predicted (with p<5e-6 threshold) per-SD (standard deviation) higher IL-8 was associated with -0.103 (-0.155, -0.051, $p_{adjusted}$ = 0.004) mm^3 smaller hippocampal volume and higher intelligence fluid score [β: 0.103 SD (95% CI: 0.042, 0.165), $p_{adjusted}$ = 0.041]. Sensitivity analyses generally showed similar results, and no pleiotropic effect, heterogeneity, nor possible reverse causation was detected.

Conclusion: Our results suggested a possible causal association of high IL-8 levels with better cognitive performance but smaller hippocampal volume among the general healthy population, highlighting the complex role of inflammation in dementia-related phenotypes. Further research is needed to elucidate mechanisms underlying these associations.

Introduction

Dementia has become a major global health concern given the increased longevity and the increased number of people aged 60 years and older. The estimated population living with dementia was 47 million in 2015, which is estimated to triple by 2050 [1,2]. Changes in the brain start to occur several years before the first manifestation of clinical symptoms and diagnosis [3]. This indicates a large time window to delay, or even prevent the onset of clinically significant cognitive deficits. In the absence of any effective pharmacologic preventive and/or curative treatment to date [4], early markers would therefore provide insight into the pathogenesis and targets for potential preventive strategies.

Systemic inflammation has been hypothesized as a risk factor in cognitive decline and dementia [5-8]. Epidemiological studies found associations of elevated systemic inflammation and worse cognition in cross-sectional studies, and a steeper cognitive decline in prospective studies [9,10]. Moreover, anti-inflammatory therapy, particularly nonsteroidal anti-inflammatory drugs (NSAIDs), have been associated with a lower risk of cognitive decline in a meta-analysis of observational cohort studies [11]. However, in a recent systematic review of randomized clinical trials (RCTs), aspirin or other NSAIDs did not lower the risk of dementia [12]. Several potential explanations could have contributed to these inconsistencies. Most importantly, associations observed in epidemiological studies are prone to reverse causality and unmeasured and/or residual confounding. In addition, NSAIDs might be beneficial only when used in the very early stages of cognitive deficits [13]. Intervention in RCTs that have been carried out might be too late to modify the progression of cognitive decline. Therefore, it remains to be elucidated whether systemic inflammation is causally related to cognitive performance.

As such, Mendelian Randomization (MR) is an alternative approach using genetic instruments that are randomly allocated at conception as a proxy of exposure to infer the causality of life-long exposure on disease [14]. Three recent MR studies on inflammatory markers and Alzheimer's Disease (AD) yielded inconsistent results [15-17]. Despite being done in large populations, these studies could still suffer from limited power caused by a limited number of cases. Continuous traits are generally acknowledged to have more statistical power, and therefore cognitive function and measures of brain atrophy, which are hallmark characteristics of AD and other forms of dementia [18,19], might be better suitable phenotypes to dissect the potential causation of inflammation and dementia.

Given that inflammation is a complex process regulated through an integrated network of pro- and anti-inflammatory immune cells and cytokines, investigating multiple inflammatory markers simultaneously in the same study population could provide more insights into the role of inflammation in dementia. Therefore, in the present study, we leveraged a two-sample MR to assess causality in relation to a comprehensive amount of 41 genetically predicted circulating levels of systemic inflammatory markers with general cognitive performance (with three additional domains of the fluid intelligence score, prospective memory result, and reaction time) and measures of brain atrophy (cerebral cortical surface area and thickness and hippocampal volume).

Method

Study design

We conducted a two-sample MR study, and the overview of the study design is presented in **Figure 1**. MR builds on three principal assumptions: the instrumental variables should firstly be associated with the exposure; secondly not be associated with confounding factors in the relation between exposure and outcome; thirdly affect the outcome exclusively via the exposure, but not via other pathways. Data involved in the present study are publicly available summary statistics from genome-wide association study (GWAS). Specific ethical approval and informed consent were obtained in the original studies.

Selection of instrumental variables

Genetic instrumental variables associated with 41 circulating cytokines and growth factors were obtained from a recent GWAS which was conducted in 8293 randomly chosen participants from five geographical areas of Finland aged between 25-74 years, including The Cardiovascular Risk in Young Finns Study (YFS), FINRISK 1997 and FINRISK 2002 [20] (data were downloaded from *http://computationalmedicine.fi/data#Cytokine_GWAS*). All gene-exposure associations were reported as regression coefficients (β) in SD-scaled units, adjusted for age, sex, body mass index, and the first 10 genetic principal components.

To avoid pleiotropic bias, we first excluded SNPs that were associated with more than one cytokine and/or growth factor. Linkage disequilibrium (LD) between all SNPs for the same exposure was assessed in the European 1000 Genome Project reference panel. When LD presented (LD > 0.001), the variant with the smallest P-value was retained. To minimize weak instrumental bias, we considered an F-statistics of above 10 as indicative of a sufficiently strong instrument. Since many markers had no significant single nucleotide polymorphisms (SNPs) at the genome-wide significant level ($p < 5e\text{-}8$), we also adopted a more relaxed significance threshold ($p < 5e\text{-}6$) for instrumental variables selection. F statistics was calculated to assess the strength of each instrument, and a value of above 10 is considered sufficient. The proportion of total variation (R^2) explained by the individual genetic instrument was calculated using the formula , where β is the effect of the genetic variant on the respective cytokine levels and MAF is the minor allele frequency [21]. When MAF is not presented in the original dataset, we extracted the effect allele frequency from PhenoScanner GWAS database Version 2. R^2 for each exposure was calculated in an additive model of all included SNPs assuming no interaction between the individual SNPs. In addition, we calculated statistical power based on the online tools for continuous outcomes (*https://github.com/kn3in/mRnd*) (9), where the alpha level was set to 0.05. When the variance explained by the genetic variants was at the minimum of 2%, we had adequate power of 0.8 to detect about 0.04, 0.088, 0.11, 0.06, 0.06, 0.035 units difference in cognitive performance (in SD), cortical measures (surface area in mm^2 and thickness in mm), hippocampal volume (in mm^3), fluid intelligence score (SD), prospective memory results (SD), and reaction time (SD) respectively (**Figure 2**).

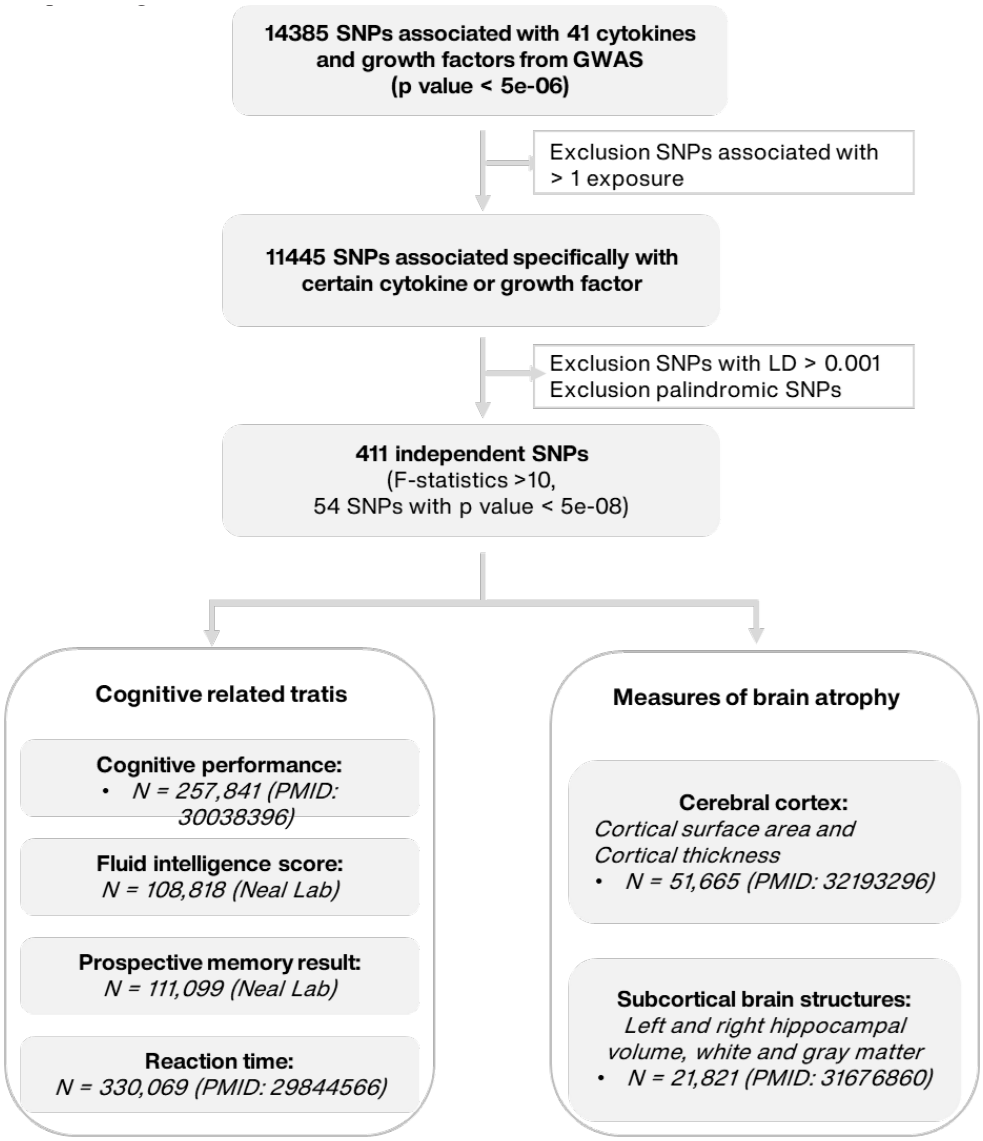


Figure 1 Schematic overview of the study design

Outcome data sources

Cognitive Performance

Summary statistics for the genetic associations with general cognitive performance were extracted from a recent GWAS (N = 257,841) with European-descent individuals which used a sample-size-weighted meta-analysis to combine data from Cognitive Genomics Consortium (COGENT, n = 35,295) and UK Biobank (n = 222,543) [22]. All individuals were aged between 16 and 102 years without

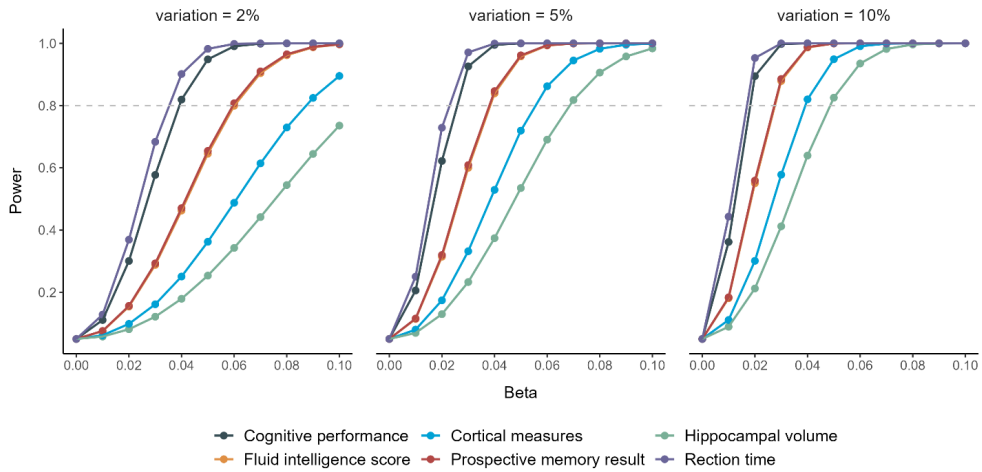


Figure 2 Statistical power for each outcome in MR analyses with different variations explained by the genetic instrumental variables

The calculation statistical power based on the online tools for continuous outcomes (*https://github.com/kn3in/mRnd*), where the alpha level was set to 0.05.

stroke or prevalent dementia. In COGENT, cognitive performance was defined as the score on the first unrotated component of the performance of at least three different neuropsychological tests within each included study. In the UK Biobank, verbal numerical reasoning (VNR) was assessed by 13 multiple-choice questions, and the VNR score was determined as the number of questions answered correctly with a time limit of two minutes, designed as a measure of a fluid intelligence test. This test has been demonstrated to have adequate reliability and validity [22,23]. Since the majority of the associations, as reflected in the number of participants, are derived from the UKB and thus the final effects are more representative for the executive function measured by the fluid intelligence test. Genetic association estimates are presented in SD units.

Due to the heterogenous definition of cognitive performance in the above GWAS, we additionally included three phenotypes to represent different domains, namely the fluid intelligence score (N = 108,818) and prospective memory result (N = 111,099) from the UK Biobank conducted by the Neal Lab (https://www.neallab.org/), and reaction time (N = 330,069) from the COGENT. Fluid intelligence score is a simple unweighted sum of the number of correct answers given to the 13 fluid intelligence questions. Therefore, we included this phenotype as a holistic measurement of multiple domains of 'fluid intelligence'. For prospective memory, participants were allowed up to 2 attempts to correctly recall the color/shape that was shown to them earlier in the touchscreen section, and it condenses the results into 3 group: instruction not recalled, either skipped or incorrect, correct recall on first attempt, and correct recall on the second attempt. Reaction time is based on 12 rounds of the card-game 'Snap'. The participant is shown two cards at a time; if both cards are the same, they press a button-box that is on the table in front of them as quickly as possible. For each of the 12 rounds, the following data were collected: the pictures shown on the cards, the number of

times the participant clicked the 'snap' button, and the time it took to first click the 'snap' button.

Cerebral cortical surface area and thickness
Genetic associations with cerebral cortical surface area (mm^2) and thickness (mm) were obtained from a genome-wide association meta-analysis of brain T1-weighted magnetic resonance imaging (MRI) data from 51,665 predominantly healthy individuals of European ancestry aged between 3.3 and 91.4 years across 60 cohorts by Enhancing Neuroimaging Genetics through Meta-analysis (ENIGMA) consortium [24]. Almost all participants are healthy, except for less than 1% with psychiatric disorders from included case-control studies out of the 60 cohorts. The cortical surface area was measured at the grey-white matter boundary and thickness was measured as the average distance between the white matter and pial surfaces. The total surface area and average thickness were computed for each participant separately. The genetic associations were calculated using an additive model within each cohort, adjusted for age, age squared, sex, sex-by-age interactions and age squared, the first four multidimensional scaling components, and diagnostic status (when the cohort followed a case-control design) and dummy variables for scanner when applicable.

Hippocampal volume
For gene-hippocampal volume associations, a GWAS meta-analysis from high-resolution brain MRI scans in 33,536 healthy individuals aged between 11 and 98 years at 65 sites between the ENIGMA and the CHARGE consortia was used; mean bilateral hippocampal volume (mm^3) was defined as the average of left and right [25]. Genetic associations in the study were assessed within each site, adjusted for age, age squared, sex, intracranial volume, four multidimensional scaling components, and diagnostic status when applicable; site effects were also adjusted for studies with data collected from multiple centers or scanners; mixed-effects models were additionally used to account for familial relationship with family data. Since only the z-statistic and p-value for each SNP were provided in the dataset, we calculated the corresponding estimate of the standardized regression coefficient for an outcome on an genetic variant () based on the equation = z statistic × se in the previous study; and se is computed by , where MAF is the minor allele frequency and N is the sample size [26].

Statistical analysis
We harmonized exposure and outcome GWAS summary statistics by making alignment of the summary statistics to the forward strand if the forward strand was known or could be inferred. Palindromic SNPs, that could not to be inferred to the forward strand and can introduce ambiguity into the identification of the effect allele in the exposure and outcome GWAS, were removed [27].

For the main analysis, Inverse-Variance weighted (IVW) regression analysis was used, which assumes no directional pleiotropic effects of individual instrumental variable [28]. This estimate combines the SNP-specific Wald ratios (gene-outcome divided by gene-exposure) by using a meta-analysis weighted by the inverse of the variance of the Wald estimates. Results are expressed as per unit change in regression coefficient (95% confidence interval) on the outcome of per standard deviation (SD) change in inflammatory markers. We also performed additional

sensitivity analyses that take pleiotropy into account, including MR-Egger, weighted-median estimator, and MR PRESSO (Pleiotropy Residual Sum and Outlier) [29-31]. In particular, the MR-Egger regression intercept estimates the average pleiotropic effect across the genetic variants if the MR assumption and the InSIDE (INstrument Strength Independent of Direct Effect) assumption hold [29]. An intercept that differs from zero indicates the presence of directional pleiotropy. A Weighted-Median estimator analysis can provide a valid estimate if at least half of the instrumental variables are valid [30]. MR-PRESSO was applied when there were sufficient number of genetic variants to detect and correct for horizontal pleiotropy through removing outliers with the assumption of more than 50% valid instruments and balanced pleiotropy and InSIDE [31]. We used Cochran's Q test statistic to examine the between-SNP heterogeneity; singe-SNP analysis was used to perform MR on each SNP individually and leave-one-out analysis was performed to assess if one particular variant could potentially have driven the association.

Reverse analysis

In order to examine possible presence of reverse causation, we additionally tested the associations between genetically influenced cognitive performance and measures of brain atrophy with any of the 41 inflammatory markers. To this end, we extracted independent genetic variants at genome-wide significant level for cognitive performance, cerebral cortical surface area and thickness, and hippocampal volume, from the same GWAS when these are used as outcomes in the main analyses. We excluded SNPs that were both associated with cerebral cortical surface area and thickness to maximally eliminate pleiotropic effect.

All the analyses were undertaken using R (v3.6.3) statistical software (The R Foundation for Statistical Computing, Vienna, Austria). MR analyses were performed using the R package "TwoSample MR" and "MR PRESSO". We used the false-discovery rate (FDR)-based multiple comparison, using Benjamini-Hochberg method, to correct for multiple testing. For cognitive performance and its domains, the adjustments were performed within each trait to avoid excessive stringency, whereas for cortical measure, the adjustment were performed for surface area and thickness simultaneously due to the same MRI measurement for these two traits. An adjusted p-value of 0.05 was considered as statistically significant.

Results

MR results

At the genome wide significance threshold ($p < 5e\text{-}08$) for instruments selection, 9 out of 41 inflammatory markers with no less than 3 SNPs. F-statistics were between 29.9 and 789.1, and the variation explained by the genetic variants for each marker ranged from 1.7% for VEGF to 19.8% for MIP1b, as shown in **Table 1**. No significant associations between inflammatory markers and any of the outcome measures were observed after stringent correction for multiple testing.

At a relaxed significance threshold ($p < 5e\text{-}06$) for instruments selection for the other 32 markers, F-statistics of individual variants ranged from 11.2 to 789.

Table 1 Summary information of genetic instrumental variables for each systemic inflammatory marker

Exposure	Full name	p < 5e-08			p < 5e-06		
		No. of SNPs	Variation (%)	F-statistics (range)	No. of SNPs	Variation (%)	F-statistics (range)
bNGF	Nerve Growth Factor Beta	1	1.1	36.5	8	5.3	20.8-36.5
CTACK	Cutaneous T-cell-attracting chemokine	3	6.3	29.9-142.7	11	12.7	21.5-142.7
Eotaxin	-	3	2.3	32.5-95.2	16	6.8	20.8-95.2
FGFBasic	Basic Fibroblast Growth Factor	0	-	-	6	2	20.8-24
GCSF	Granulocyte Colony-Stimulating Factor	0	-	-	9	6.3	20.4-25.2
GROa	Growth Related Oncogene Alpha	2	6.3	41.8-184.4	10	14.4	20.7-184.4
HGF	Interleukin 10	2	1.5	40.8-57.3	9	4.4	20.7-57.3
IFNg	Interleukin 12p70	0	-	-	9	2.5	21.9-23.9
IL10	Hepatocyte Growth Factor	0	-	-	6	2	21.3-25.3
IL12p70	Interferon Gamma	0	-	-	9	3.4	20.8-26.5
IL13	Interleukin 13	0	-	-	10	8.4	21.1-25.4
IL16	Interleukin 16	2	8.2	31.1-132	10	16.7	20.9-132
IL17	Interleukin 17	1	0.6	39	9	3.3	20.4-39
IL18	Interleukin 18	4	6.8	31.8-96.2	19	21.5	20.6-96.2
IL1b	Interleukin 1 Beta	0	-	-	6	3.9	13.6-31.6
IL1ra	Interleukin 1 Receptor Antagonist	0	-	-	7	4.9	21.2-23.6
IL2	Interleukin 2	0	-	-	10	7	20.9-23.5
IL2ra	Interleukin 2 Receptor Alpha Subunit	1	9.7	167.6-167.6	9	16.9	21.3-167.6
IL4	Interleukin 4	0	-	-	11	5.2	21.1-26.6
IL5	Interleukin5	0	-	-	5	3.8	22.1-24.9
IL6	Interleukin 6	0	-	-	7	2.7	21.6-23.3
IL7	Interleukin 7	0	-	-	12	12.3	20.6-26.4
IL8	Interleukin 8	0	-	-	3	4.8	22.1-24.6
IL9	Interleukin9	0	-	-	7	5.3	21.2-26.5
IP10	Interferon gamma-induced protein 10	2	1.9	31.1-32	12	10	21-32
MCP1	Monocyte Chemoattractant Protein-1	3	1.9	30.3-86.4	13	6	20.9-86.4
MCP3	Macrophage Colony-Stimulating Factor	0	-	-	4	8.3	22-25.7
MCSF	Macrophage migration Inhibitory Factor	1	1.6	31.6	8	13.3	21.3-31.6
MIF	Monokine Induced by interferon-Gamma	1	1.1	39-39	8	8.1	21.2-39
MIG	Interleukin 10	1	1	42.4	17	17.8	19-42.4
MIP1a	Macrophage Inflammatory Protein 1 Alpha	0	-	-	9	6.4	21-22.8
MIP1b	Macrophage Inflammatory Protein-1 Beta	7	19.8	58.2-789.1	22	26.2	20.6-789.1
PDGFbb	Platelet Derived Growth Factor BB	4	3.3	30.8-103.3	14	6.4	21-103.3
RANTES	Regulated on Activation, Normal T Cell Expressed and Secreted	1	0.4	30	11	7.6	20.9-30
SCF	Stem Cell Factor	1	0.4	31.8	9	3.9	20.8-31.8
SCGFb	Stem Cell Growth Factor Beta	3	6.5	43.3-99.4	14	13.7	20.7-99.4
SDF-1α	Stromal cell-Derived Factor 1 Alpha	0	-	-	9	3.6	11.2-29.2
TNFa	Tumor Necrosis Factor Alpha	0	-	-	5	6.8	22.1-24.9
TNFb	Tumor Necrosis Factor Beta	1	2	50.9	5	8.5	21.9-50.9
TRAIL	TNF-Related Apoptosis Inducing Ligand	7	18.5	99.4-370	19	24.4	19.9-370
VEGF	Vascular Endothelial Growth Factor	3	1.7	33.4-42.9	14	6.5	20.8-42.9

The proportion of total variance (R^2) explained by the individual genetic instrument was calculated using the formula $R^2=(\beta \times \sqrt{2 \times MAF(1-MAF)})$, where β is the effect of the genetic.

Instruments for each marker explained the proportional variance from 2.0% for FGFBasic to 26.2% for MIP1b. The results are generally similar comparing

Table 2 Associations of systemic inflammation markers with measures of cognitive performance and measures of brain atrophy

Exposure	Outcome	Method*	No. of SNPs	Beta (95%CI)	padjusted	Heterogeneity Q value (p value)	MR-Egger intercept (p value)
IL-8	Hippocampal volume (mm^3)	IVW	3	-0.103 (-0.155, -0.051)	0.004	1.1 (0.6)	
		Weighted-median	3	-0.099 (-0.169, -0.282)	0.22		
		MR-Egger	3	-0.073 (-0.154, 0.008)	0.92	0.2 (0.7)	-0.008 (0.5)
IL-8	Fluid intelligence score (SD)	IVW	3	0.103 (0.042, 0.165)	0.041	1.6 (0.6)	
		Weighted-median	3	0.102 (0.015, 0.186)	0.45		
		MR-Egger	3	0.103 (-0.001, 0.208)	0.86	1.6 (0.3)	-0.000 (1.0)

IVW: Inverse variance weighted.
Genetic variants selection was based on p < 5e-06. MR PRESSO (MR Pleiotropy RESidual Sum and Outlier) was not available due to limited number of SNPs.
$p_{adjusted}$ represents the FDR-based adjusted p-value.

to those obtained at p < 5e-08 for instruments selection of the inflammatory markers when available, with however narrower confidence intervals (**Supplementary Table 1 and 2**). In the primary analysis using IVW, genetically predicted one-SD higher IL-8 was associated with -0.103 (-0.155, -0.051, $p_{adjusted}$ = 0.004) mm^3 smaller hippocampal volume. For different cognitive performance domains, higher IL-8 was however associated with higher intelligence fluid score [β: 0.103 SD (95% CI: 0.042, 0.165), $p_{adjusted}$ = 0.041], as shown in **Table 2**. Estimations from the weighted-median method were generally consistent with those from IVW, and betas (95%CIs) were -0.099 (-0.169, -0.282, $p_{adjusted}$ = 0.23) and 0.10 (0.018, 0.182, $p_{adjusted}$ = 0.4), respectively. No pleiotropic effect was detected via the MR-Egger intercept (both p value > 0.5). No between-SNP heterogeneity observed with the Cochran's Q test statistic. Single-SNP analysis and leave-one-out analysis indicated that these associations were unlikely driven by a certain extreme variant (data not shown).

Reverse analyses

In total, respectively 182, 121, 6, and 4 SNPs for cognitive performance, cerebral cortical surface area and thickness, and hippocampal volume that also presented in the inflammatory marker GWAS were identified. Upon correcting for multiple testing using FDR, no significant association was detected using IVW method (**Supplementary Table 3**). Results from sensitivity MR methods did not differ substantially compared to the IVW analyses.

Discussion

In this two-sample MR analysis, we used genetic instruments for 41 systemic inflammatory markers to assess their potential causal associations with cognitive performance including its three domains (fluid intelligence score, prospective memory result, and reaction time), and measures of brain atrophy. Genetically predicted higher levels of IL-8 were associated with smaller hippocampal volume, but with better fluid intelligence score.

Systemic inflammation may result in neuronal consequences through several different pathways: 1) by interacting with blood-brain barrier function through receptor binding of inflammatory markers or through secretion of immune-active substances; 2) by neural afferent pathways bypassing the blood-brain barrier such as via the cranial nerve; and 3) through diffusion from blood through the perivascular spaces via the circumventricular organs which lack an endothelial blood-brain barrier [32-34]. IL-8 is a chemokine produced by several cell types and functions as a chemoattractant that recruits different types of immune cells to sites of inflammation by activating predominantly neutrophils and it can act as a potent angiogenic factor. Higher IL-8 levels, either in circulation or cerebrospinal fluid, have been associated with poor cognitive performance [35,36]. However, IL-8 measured in AD patients was found to be either elevated [5,37], decreased [38] or unchanged [39]. Similarly, we also found conflicting results of IL-8, as it associated with both smaller hippocampal volume and better fluid intelligence score upon correction for multiple testing. However, interestingly, in consistent with these finding, genetically determined higher levels of IL-8 levels were also associated with better general cognitive performance [β: 0.026 SD (0.007, 0.045), $p_{original}$ = 0.008] although before multiple testing correction only, and the same direction of effect on prospective memory score [β: -0.016 SD (95% CI: -0.045, 0.012)], $p_{adjusted}$ = 0.8], although insignificant.

This inconsistency may be explained by several hypotheses, firstly the different domains of cognitive performance. Specifically, a previous study specifically found that increased IL-8 was associated with worse cognitive performance in the memory and speed domains and in motor function [36]. While hippocampal volume is associated with amnestic, but not non-amnestic domain cognitive performance [40,41], fluid intelligence in the present study was more of a reflection of executive function. Alternatively, the higher pro-inflammatory status induced by increased IL-8, might be counter-balanced by the production of higher levels of anti-inflammatory components, which may differentially affect cognition and hippocampal volume. As an example, while cortisol, which has potent anti-inflammatory effect and is often elevated in response to inflammation, may enhance alertness and memory, long-term exposure to cortisol may severely damage the hippocampus [42]. In addition, inflammation may cast a dual role in the pathogenesis of dementia-related conditions particularly in AD [43,44]. Briefly, in the healthy state or preclinical stage of neurological diseases, a modest inflammatory response would be beneficial for the clearance of waste products, whereas in advanced stages, overreacted excessive inflammatory response that exceeds the capacity of self-repair would exacerbate the (neuro)inflammation. A fine example is that lower levels of both C-reactive protein and complement C3, representing a low inflammation profile, are causally associated with higher risk of AD in the general Danish population [45,46]. However, the dynamics of IL-8 in the pathogenesis of neurological diseases needs further investigation.

Strength and limitations

To the best of our knowledge, we firstly comprehensively tested the causal associations between multiple inflammatory markers with cognitive performance and measures of atrophy. By using two sample MR design, including the reverse MR study, we maximally avoided the drawbacks of reverse causation

and residual confounding in observational studies. However, several limitations need to be taken into account when interpreting the results. First, the activities of the inflammatory markers are complex, particularly cytokines that are highly pleiotropic and can act on many different cell types [47,48] and that depending on the context play roles in repair of tissue damage as well as in (chronic) inflammation. Moreover, individual cytokines can have many functions and different cytokines can share similar functions, which will induce a series of combined effects synergistically or antagonistically that functionally alter target cells [47,48]. Despite a series attempts have been made by excluding variants that were identified for multiple cytokines and by performing sensitivity analyses to ensure the elimination of potential pleotropic effect, no established methodology deals with such a complex orchestrated traits' network, and therefore we could not completely rule out the possibility of bias by directional pleiotropy using current methods. In addition, given the fact that cytokines rarely manifest their effects alone but rather work in regulatory networks, gene-gene interaction is important to disentangle the role of inflammatory markers in health-related conditions [49], such as risk of AD [50] and other forms of dementia. Second, we used a significance threshold of $p < 5e\text{-}6$ for the selection of instrumental variables, which might have included false-positive variants and consequently bias findings. Since most of the inflammatory markers are very expensive to measure, the included GWAS for these markers, although being the largest to date, comprises only 8293 European-ancestry individuals. Compared to the sample sizes of tens of thousands for other traits, for example, the outcome phenotypes, this might be too small to detect as many genome-wide significant genetic variants as possible. Therefore, the selection of a relaxed threshold this is a trade-off with statistical power. A more stringent threshold results in less available instrumental variables with subsequently decreased statistical power. Consequently, a null association identified might not be indicative of absence of evidence, but rather of insufficient power. In our analyses, for some markers with available instruments at both thresholds, estimates for the same outcome are comparable (**Supplementary Table 1 and 3**) but with smaller confidence interval given increased power. Furthermore, in previous MR studies, when it comes to complex traits, such as depression, with limited genetic variants at $p < 5e\text{-}08$ level, a more relaxed p-value ($p < 1e\text{-}6$) has been adopted and successfully disentangled the bidirectional causal relationships between physical activity and depression in MR analyses [51]. Taken together, we also used a relaxed threshold to identify any possible link between systemic inflammatory markers with outcome. However, more studies are needed to confirm these possible associations, particularly using genetic variants from GWAS with larger sample size. Lastly, this study is performed based on populations with European ancestry thus the results could be not representative of other groups with different ethnic backgrounds.

Conclusion

In conclusion, our MR study found some evidence to support a causal association of higher genetically determined IL-8 level and better cognitive performance and smaller hippocampal volume. Further research is needed to elucidate mechanisms underlying these associations, and to assess the suitability of these markers as potential preventive or therapeutic targets.

References

1. Martin Prince., Anders Wimo., Maëlenn Guerchet., Gemma-Claire Ali., Yu-Tzu Wu., Matthew Prina. World Alzheimer Report 2015: the global impact of dementia: an analysis of prevalence, incidence, cost and trends. 2015.
2. G. Livingston, A. Sommerlad, V. Orgeta, et al. Dementia prevention, intervention, and care. Lancet 2017; 390(10113): 2673-734.
3. L. L. Beason-Held, J. O. Goh, Y. An, et al. Changes in brain function occur years before the onset of cognitive impairment. J Neurosci 2013; 33(46): 18008-14.
4. H. A. Fink, E. Jutkowitz, J. R. McCarten, et al. Pharmacologic Interventions to Prevent Cognitive Decline, Mild Cognitive Impairment, and Clinical Alzheimer-Type Dementia: A Systematic Review. Ann Intern Med 2018; 168(1): 39-51.
5. X. N. Shen, L. D. Niu, Y. J. Wang, et al. Inflammatory markers in Alzheimer's disease and mild cognitive impairment: a meta-analysis and systematic review of 170 studies. J Neurol Neurosurg Psychiatry 2019; 90(5): 590-8.
6. S. K. L. Darweesh, F. J. Wolters, M. A. Ikram, F. de Wolf, D. Bos, A. Hofman. Inflammatory markers and the risk of dementia and Alzheimer's disease: A meta-analysis. Alzheimers Dement 2018; 14(11): 1450-9.
7. A. Koyama, J. O'Brien, J. Weuve, D. Blacker, A. L. Metti, K. Yaffe. The role of peripheral inflammatory markers in dementia and Alzheimer's disease: a meta-analysis. J Gerontol A Biol Sci Med Sci 2013; 68(4): 433-40.
8. K. S. P. Lai, C. S. Liu, A. Rau, et al. Peripheral inflammatory markers in Alzheimer's disease: a systematic review and meta-analysis of 175 studies. J Neurol Neurosurg Psychiatry 2017; 88(10): 876-82.
9. K. A. Walker, R. F. Gottesman, A. Wu, et al. Systemic inflammation during midlife and cognitive change over 20 years: The ARIC Study. Neurology 2019; 92(11): e1256-e67.
10. C. Serre-Miranda, S. Roque, N. C. Santos, et al. Cognition Is Associated With Peripheral Immune Molecules in Healthy Older Adults: A Cross-Sectional Study. Front Immunol 2020; 11: 2045.
11. W. Wang, Y. Sun, D. Zhang. Association Between Non-Steroidal Anti-Inflammatory Drug Use and Cognitive Decline: A Systematic Review and Meta-Analysis of Prospective Cohort Studies. Drugs Aging 2016; 33(7): 501-9.
12. F. Jordan, T. J. Quinn, B. McGuinness, et al. Aspirin and other non-steroidal anti-inflammatory drugs for the prevention of dementia. Cochrane Database of Systematic Reviews 2020; (4).
13. Bruno P. Imbimbo, Vincenzo Solfrizzi, Francesco Panza. Are NSAIDs useful to treat Alzheimer's disease or mild cognitive impairment? Frontiers in aging neuroscience 2010; 2: 19.
14. D. A. Lawlor, R. M. Harbord, J. A. Sterne, N. Timpson, G. Davey Smith. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. Stat Med 2008; 27(8): 1133-63.
15. C. H. C. Yeung, C. M. Schooling. Systemic inflammatory regulators and risk of Alzheimer's disease: a bidirectional Mendelian-randomization study. Int J Epidemiol 2020.

16. Panagiota Pagoni, Laura D. Howe, George Davey Smith, Yoav Ben-Shlomo, Evie Stergiakouli, Emma L. Anderson. Causal effects of circulating cytokine concentrations on risk of Alzheimer's disease: A bidirectional two-sample Mendelian randomization study. medRxiv 2020: 2020.11.18.20232629.
17. L. Fani, M. K. Georgakis, M. A. Ikram, M. K. Ikram, R. Malik, M. Dichgans. Circulating biomarkers of immunity and inflammation, risk of Alzheimer's disease, and hippocampal volume: a Mendelian randomization study. Transl Psychiatry 2021; 11(1): 291.
18. F. Jessen, R. E. Amariglio, M. van Boxtel, et al. A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease. Alzheimers Dement 2014; 10(6): 844-52.
19. L. Pini, M. Pievani, M. Bocchetta, et al. Brain atrophy in Alzheimer's Disease and aging. Ageing Res Rev 2016; 30: 25-48.
20. A. V. Ahola-Olli, P. Würtz, A. S. Havulinna, et al. Genome-wide Association Study Identifies 27 Loci Influencing Concentrations of Circulating Cytokines and Growth Factors. Am J Hum Genet 2017; 100(1): 40-50.
21. T. M. Teslovich, K. Musunuru, A. V. Smith, et al. Biological, clinical and population relevance of 95 loci for blood lipids. Nature 2010; 466(7307): 707-13.
22. J. J. Lee, R. Wedow, A. Okbay, et al. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. Nat Genet 2018; 50(8): 1112-21.
23. C. Fawns-Ritchie, I. J. Deary. Reliability and validity of the UK Biobank cognitive tests. PLoS One 2020; 15(4): e0231627.
24. K. L. Grasby, N. Jahanshad, J. N. Painter, et al. The genetic architecture of the human cerebral cortex. Science 2020; 367(6484).
25. D. P. Hibar, H. H. H. Adams, N. Jahanshad, et al. Novel genetic loci associated with hippocampal volume. Nat Commun 2017; 8: 13624.
26. Z. Zhu, F. Zhang, H. Hu, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet 2016; 48(5): 481-7.
27. Fernando Pires Hartwig, Neil Martin Davies, Gibran Hemani, George Davey Smith. Two-sample Mendelian randomization: avoiding the downsides of a powerful, widely applicable but potentially fallible technique. International journal of epidemiology 2016; 45(6): 1717-26.
28. S. Burgess, R. A. Scott, N. J. Timpson, G. Davey Smith, S. G. Thompson. Using published data in Mendelian randomization: a blueprint for efficient identification of causal risk factors. Eur J Epidemiol 2015; 30(7): 543-52.
29. J. Bowden, G. Davey Smith, S. Burgess. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol 2015; 44(2): 512-25.
30. J. Bowden, G. Davey Smith, P. C. Haycock, S. Burgess. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. Genet Epidemiol 2016; 40(4): 304-14.
31. M. Verbanck, C. Y. Chen, B. Neale, R. Do. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. Nat Genet 2018; 50(5): 693-8.

32. A. Fung, M. Vizcaychipi, D. Lloyd, Y. Wan, D. Ma. Central nervous system inflammation in disease related conditions: mechanistic prospects. Brain Res 2012; 1446: 144-55.
33. A. Varatharaj, I. Galea. The blood-brain barrier in systemic inflammation. Brain Behav Immun 2017; 60: 1-12.
34. T. Lin, G. A. Liu, E. Perez, et al. Systemic Inflammation Mediates Age-Related Cognitive Deficits. Front Aging Neurosci 2018; 10: 236.
35. D. Galimberti, N. Schoonenboom, P. Scheltens, et al. Intrathecal chemokine synthesis in mild cognitive impairment and Alzheimer disease. Arch Neurol 2006; 63(4): 538-43.
36. B. T. Baune, G. Ponath, J. Golledge, et al. Association between IL-8 cytokine and cognitive performance in an elderly general population--the MEMO-Study. Neurobiol Aging 2008; 29(6): 937-44.
37. R. Taipa, S. P. das Neves, A. L. Sousa, et al. Proinflammatory and anti-inflammatory cytokines in the CSF of patients with Alzheimer's disease and their correlation with cognitive decline. Neurobiol Aging 2019; 76: 125-32.
38. R. Hesse, A. Wahler, P. Gummert, et al. Decreased IL-8 levels in CSF and serum of AD patients and negative correlation of MMSE and IL-1β. BMC Neurol 2016; 16(1): 185.
39. W. Swardfager, K. Lanctôt, L. Rothenburg, A. Wong, J. Cappell, N. Herrmann. A meta-analysis of cytokines in Alzheimer's disease. Biol Psychiatry 2010; 68(10): 930-41.
40. I. A. Clark, A. M. Monk, V. Hotchin, et al. Does hippocampal volume explain performance differences on hippocampal-dependant tasks? Neuroimage 2020; 221: 117211.
41. M. Grundman, C. R. Jack, Jr., R. C. Petersen, et al. Hippocampal volume is associated with memory but not monmemory cognitive performance in patients with mild cognitive impairment. J Mol Neurosci 2003; 20(3): 241-8.
42. R. M. Sapolsky, L. M. Romero, A. U. Munck. How do glucocorticoids influence stress responses? Integrating permissive, suppressive, stimulatory, and preparative actions. Endocr Rev 2000; 21(1): 55-89.
43. R. Hohlfeld, M. Kerschensteiner, E. Meinl. Dual role of inflammation in CNS disease. Neurology 2007; 68(22 Suppl 3): S58-63; discussion S91-6.
44. C. Zheng, X. W. Zhou, J. Z. Wang. The dual roles of cytokines in Alzheimer's disease: update on interleukins, TNF-α, TGF-β and IFN-γ. Transl Neurodegener 2016; 5: 7.
45. S. H. Hegazy, J. Q. Thomassen, I. J. Rasmussen, B. G. Nordestgaard, A. Tybjaerg-Hansen, R. Frikke-Schmidt. C-reactive protein levels and risk of dementia-Observational and genetic studies of 111,242 individuals from the general population. Alzheimers Dement 2022.

Supplementary Materials

Supplementary Table 1 Information of the systemic inflammatory markers

Supplementary Table 2 Validation of significant association of systemic inflammation markers with measures of brain atrophy with the p value threshold of < 5e-08 for instrumental variables selection

Supplementary Table 3 Association of systemic inflammation markers with cognitive performance and brain atrophy in Inverse-variance Weighted analysis

Supplementary Figure 1 Power calculation for different outcomes

Supplementary Figure 2 Single SNP analyses

Supplementary Figure 3 Leave-one-out analysis

The Supplementary materials for this article can be found online at: *https://link.springer.com/article/10.1007/s11357-022-00602-7*

CHAPTER 10

Main findings and future perspectives

SUMMARY OF MAIN FINDINGS

The overall aim of this thesis was to disentangle the role of oxidative stress in chronic disease, with a focus on cardiovascular disease (CVD) and related cardiometabolic risk factors. The results derived from the studies described in this thesis addressed two main questions with the application of state-of-the-art epidemiological research methods: 1) at a population level, whether mitochondrial dysfunction is a causal risk factor in the development of atherosclerotic CVD and related intermediate risk factors (**Part II**), and 2) whether dietary antioxidants impose any clinical-relevant benefits in the prevention of atherosclerotic CVD (**Part III**). In addition, this thesis also sheds light on the role of inflammation in neurological diseases (**Part IV**). Furthermore, the implications of the work comprised in this thesis for future research are discussed in this part.

In **Chapter 2**, we reviewed the current biological knowledge of oxidative stress and its relation to ageing and age-related diseases in experimental and epidemiological studies. Despite reactive oxygen species being closely involved in the maintenance of cell function via a diverse array of signaling pathways, their overproduction leads to oxidative-related macromolecule damage and mitochondrial dysfunction which further result in altered lifespan and the manifestation of multiple age-related diseases. However, the associations between oxidative stress with diseases, such as CVD, are yet inconclusive, especially based on the findings from different antioxidants in observational studies and randomized clinical trials (RCTs). Nevertheless, several weaknesses in the study design and the concept of antioxidative capacity should be considered when interpreting those conflicting results. Moreover, novel and specific biomarkers of oxidative damage are warranted to monitor the effect of antioxidants supplementation.

In the general population, changes in leukocyte mitochondrial DNA copy number (mtDNA-CN) have been proposed to be a proxy for mitochondrial function and mtDNA-CN is becoming an attractive biomarker due to its relative ease of measurement. In **Chapter 3 and 4**, in line with the principles of triangulation in etiological studies, we examined the role of mtDNA-CN in CVD and metabolic traits using a combination of a prospective cohort design in the UK biobank followed by a Mendelian randomization (MR) approach. In **Chapter 3**, observational analyses among participants free of CVD at study inclusion showed that low mtDNA-CN was a risk factor of both incident coronary artery disease and heart failure. Using MR with summary-level data from the currently largest genetic consortia and biobanks, these associations were validated for coronary artery disease, but not for heart failure likely, which might be due to the phenotypic heterogeneity of heart failure. In **Chapter 4**, we further explored the association between mtDNA-CN with 168 blood-derived metabolomic measures of predominantly lipids and lipoproteins (sub)particles measured by using the Nightingale NMR-based platform. We observed associations between low mtDNA-CN and an atherogenic metabolomic profile, characterized by high levels of most lipids. These findings suggest that mitochondrial dysfunction, as proxied by mtDNA-CN, may have an influence on lipid dysregulation, a well-documented risk factor for atherosclerosis and a causal risk factor in the pathogenesis of atherosclerotic

CVD. Further studies are needed to test its validity in patients' risk classification and disease prevention.

In **Chapter 5**, we investigated the potential causal association between multiple circulating antioxidants and coronary heart disease using an MR design, including up to 768,121 participants with 93,230 cases from three large genetic consortia. We provided evidence that genetically predicted and thus lifelong higher circulating antioxidants levels, either as authentic circulating levels that are similar in magnitude to those achieved by dietary supplements or concentrations of corresponding metabolites, are unlikely to reduce the risk of coronary heart disease. However, the circulating antioxidant levels may not be equivalent to the authentic functional levels, particularly in the case of vitamin E with distinct catabolism upon oxidative modification. This highlights the need to link the markers that are specific for antioxidant capacity, i.e. antioxidants' functional levels, to CVD and related risk factors.

Vitamin E can be catabolized via either hepatic enzymatic pathways or oxidized in periphery[1]. In the hepatic pathway, vitamin E is enzymatically converted to a spectrum of enzymatic metabolites with successive shortening of the phytyl side chain. Alternatively, vitamin E acts as peroxyl lipid radicals scavenger and forms oxidized metabolites with the opening of the chromanol ring. These metabolites are predominantly excreted from the body via urine. In **Chapter 6 and 7**, we focused on the cross-sectional associations between circulating vitamin E, urinary enzymatic and oxidized metabolites, and cardiometabolic traits in approximately 500 middle-aged healthy individuals from the Netherlands Epidemiology of Obesity (NEO) Study. In **Chapter 6**, we specifically found that higher urinary levels of oxidized metabolites, but neither circulating vitamin E nor urinary enzymatic metabolites, were associated with lower insulin resistance. Similarly, in **Chapter 7**, we found that the associations of 147 NMR-based metabolomic measures, mostly consisting of lipids and lipoprotein (sub)particles, with enzymatic metabolites, have directions similar to those with circulating vitamin E. However, associations of metabolomic measures with oxidized metabolites were markedly different from those with both circulating vitamin E and enzymatic metabolites. These findings highlight that circulating vitamin E may be representative of the enzymatic catabolism but not the antioxidative function of vitamin E.

Inflammation is inextricably linked to oxidative stress. While the association between inflammation and atherosclerotic CVD is well-established, its role in neurological diseases is not clear. In **Chapter 8**, we investigated the bidirectional association between inflammatory bowel disease (IBD), as a disease model of sustained chronic inflammation, and depression using MR study in a combined sample size of 693 183 individuals (36 507 cases) for IBD and 534 635 individuals (71 466 cases) for depression, respectively. No association was observed between genetically influenced IBD and risk of depression, whereas genetically predicted depression was associated with a higher risk of IBD. In **Chapter 9**, we explored the causal effect of 41 systemic inflammatory markers on cognitive function and brain atrophy measures using an MR design. Similarly, no significant association after correcting for multiple testing was observed between 40 out of 41 inflammatory markers and any of the brain outcomes. These results however could not refute the role of inflammation in neurological diseases. In contrast, it

may indicate that rather than being a cause, an excessive inflammatory response is likely a consequence of neurological diseases.

FUTURE PERSPECTIVES

Antioxidants in chronic disease

The benefits of antioxidants have been widely hyped by the media and food industries. Their easy access largely facilitates the popularity of antioxidants supplementation in the general population. Based on the results derived from this thesis, some clinical experts now suggest stopping exploring the assumed protective effect of individual antioxidants in RCTs for atherosclerotic CVD[2]. However, from our perspective, it is still not the end of antioxidant supplementation as an easy and cheap method for chronic diseases prevention.

Antioxidant supplementation is often aimed at increasing circulating levels of antioxidants. However, the circulating levels are not necessarily corresponding to their functional levels in most of the cases due to several influential factors such as genetic background, individual health status, and the levels of other antioxidants in the circulation. For example, vitamin E shares common mechanisms for cellular uptake and efflux with cholesterol in many different cell types that are tightly regulated by genes[3]. Once taken up by cells, the intracellular distribution of antioxidants to organelles is regulated by different transport proteins binding to vitamin E, such as α-tocopherol transfer protein. In addition, vitamin C regenerates vitamin E by reducing vitamin E radicals formed during scavenging radicals. Hence, it could happen that an individual with higher supplementation with subsequently higher circulating levels has lower functional levels in certain target locations, e.g., mitochondria, to counteract overwhelming local production of oxidants. Due to the differences in antioxidant capacity among individuals, even identical supplementation strategies may possibly not lead to the same health outcomes. Therefore, monitoring the effect after supplementation with plausible biomarkers is of great interest. The ideal approach to monitor the supplementation effect is to measure the direct change of scavenging antioxidants. Although this is not possible for many antioxidants, it might be feasible in the case of vitamin E, where its urinary oxidized metabolites but not enzymatic ones are associated with CVD risk factors and thus represent the functional levels.

Furthermore, the selection of a proper population for supplementation is also critical in addition to rigorous monitoring strategies. Factors that could alter the total antioxidative capacity in the body should be considered. For example, metabolic syndrome patients had approximately 12% greater oxidation reduction potential and 59% lower total amount of readily oxidizable molecule reserves compared with healthy adults[4]. The bioavailability of vitamin E was shown to be reduced and the elimination delayed in these patients who have increased lipid peroxidation, independent of the co-ingested dairy fat amount[5]. Likewise, individuals with hyperlipidemia have been found to have reduced uptake of the newly absorbed vitamin E into blood[6]. Therefore, supplementation should be administrated to a population who are most likely to benefit. This group of population should be selected based on certain features (e.g., genotype, nutritional or health conditions, etc.) which could induce significant heterogeneous responses, for example, participants with different haptoglobin genotype. Haptoglobin binds hemoglobin with high affinity and stability, avoiding the release of heme iron from hemolysis into circulation, consequently preventing the production of

hydroxyl radical. Vitamin E supplementation has been shown to be associated with an approximately 35% reduction in CVD specifically in individuals with both diabetes and haptoglobin 2-2[7]. Furthermore, as mitochondria are among the most important vulnerable sites to oxidative stress, mitochondrial dysfunction may serve as the biomarker of oxidative damage for population selection for supplementation. Particularly, mtDNA-CN might be a proxy of mitochondrial dysfunction due to its possible causal role in atherosclerotic CVD. Indeed, in a previous study, adding mtDNA-CN to the 2013 American College of Cardiology/American Heart Association assessment tool further improves discrimination for CVD events that is mainly driven by the improvement for coronary heart disease risk, as well as improves sensitivity and specificity on initiating statin therapy for primary prevention of atherosclerotic CVD[8].

Notwithstanding, chronic diseases are multifactorial with heterogeneous pathophysiology and do not fit with the "one cause – one mechanism – one disease – one therapy" paradigm. A single strategy that only targets oxidative stress might not be sufficient to show significant pathophysiological effects on the complex pathways. Alternatively, "multifactor – multi-treatments" could be more effective since there are possible synergistic benefits from two or more agents with acceptable safety and efficacy. A combination of treatments, for example, traditional treatment plus antioxidants supplementation as adjuvant therapy, may achieve better outcomes than antioxidants supplementation only. For instance, middle-aged type 2 diabetic patients who received metformin treatment plus vitamin E and/or vitamin C had significant improvement of glucose measures as well as lipid profiles compared to patients with metformin treatment[9].

Collectively, further exploration of antioxidants in the prevention and treatment of CVD will remain important given their clear role in disease pathogenesis. Well-selected participants for supplementation and better biomarkers to monitor supplementation effects in trials can provide further insight into the role of antioxidants in CVD.

Inflammation in neurological diseases

Inflammation has emerged as an important mechanism in almost all neurological diseases, including depression and neurodegenerative diseases, such as Alzheimer's disease (AD)[10,11]. Inflammation contributes to disease pathogenesis via both the periphery and central nervous system, the latter of which is often referred to as neuroinflammation that is characterized by microglial activation and the presence of infiltrating leukocytes in the brain parenchyma. Genetic and epidemiological studies have shown robust associations between inflammation and neurological disease, however, therapeutic strategies against inflammation have been proven not very successful. This will be discussed taking AD as an example in this section.

In humans, several genetic variants associated with AD have been discovered in genome-wide association studies (GWAS) and are mapped to genes that are involved in regulating the innate immune response both within the central nervous system and in the blood. For example, genes that encode proteins involved in the regulation of complement activation and phagocytic function

of myeloid cells have been associated with AD, such as Triggering Receptor Expressed On Myeloid Cells (TREM2)[12,13], CD33[14], clusterin (CLU[15]), and complement component (3b/4b) receptor 1 (CR1)[15,16]. Similarly, in epidemiological case-control studies, peripheral inflammation, characterized by high levels of pro-inflammatory proteins such as interleukin-6 and C-reactive protein (CRP) in blood, is related to cognitive decline and neurodegeneration[17]. Moreover, long-term use of non-steroidal anti-inflammatory drugs also reduces the risk for developing AD in large prospective cohorts[18-20]. Nevertheless, in RCTs, both steroid and non-steroidal anti-inflammatory drugs generally failed to show any beneficial effect on cognition or overall AD severity[21,22].

Importantly, it is worth noting that there might be a dual role of inflammation in the progression of diseases, similar to the role of oxidative stress in chronic diseases. The inflammation dynamics frames the AD progress within a lifelong perspective of adaption to inflammatory insults induced by debris and misfolded proteins. Indeed, several studies have shown that the functions of microglia are stage-specific and change dynamically with AD progression. At the preclinical or early stages of AD, microglia can be neuroprotective assisting the clearance of accumulated amyloid beta protein, a major pathological hallmark of AD, whereas at the advanced stages of AD, microglia lose their protective function and shift to a more proinflammatory state via releasing numerous cytokines and chemoattractants, exacerbating neuroinflammation[23]. Similarly, blood monocytes that are capable to work in the blood and to infiltrate into the brain to clear amyloid-beta, share the same dynamics as microglia in the progression of AD[24,25]. This may partly explain the null effects in anti-inflammatory related trials in AD, where suppression of inflammation in patients at the early stages of AD would hypothetically unintendedly inhibit immune cell-mediated phagocytosis, in addition to other heterogeneities in RCTs such as drug dosage, administration time, and duration of the follow-up. Therefore, a too low systemic inflammatory response would probably lead to a less efficient clearance of waste products thus being associated with a high risk of AD.

Nevertheless, current evidence has been predominantly focused on a high inflammatory response in diseases, and few studies have explored the associations between a low inflammatory response and risk of neurological diseases including AD. Not surprisingly, in longitudinal studies in the general population, low levels of circulating CRP[26], complement component 3 (C3)[27], and apolipoprotein E[28,29] were observed to be associated with a higher risk of developing AD and all-cause dementia. This highlights the importance to consider the dynamics of inflammation in the development and progression of neurological diseases, particularly the shift from a beneficial to a harmful role. An in-depth understanding of the dual and complex role of inflammation may provide promising opportunities to identify novel biomarkers in the early detection of disease and facilitate trial designs for drug development.

Mendelian randomization in genetic epidemiology

The rapid technological development and decreased costs in genetic epidemiology have facilitated us to investigate the genetic makeup of multiple com-

plex traits. Notably, the sample size of the datasets for genetic associations has expanded largely in the last few years for the discovery and replication of genome-wide association studies (GWAS) findings, and more and more genetic variants are discovered. This covers a broad range of diseases, lifestyle-related phenotypes, and molecules measured using -omics platforms, including metabolomics[30,31] and proteomics[32,33]. On the one hand, tremendous summary-level data from GWAS are publicly available from most of the large genetic consortia. On the other hand, established mega-biobanks provide another fertile source to conduct research with well-genotyped and high-quality phenotypic data, such as UK Biobank[34], Estonian Biobank[35], China Kadoorie Biobank[36], Million Veteran Program[37], and FinnGen with longitudinal population-based cohort design. Importantly, individual-level data from biobanks with open access upon requirement offers opportunities for in-depth analyses, for example, generating summary statistics with desired covariates adjusted to avoid potential collider bias, conducting subgroup analyses, and investigating non-linear causal effects between exposure and outcome.

The unprecedented availability of these types of data leads to the extreme popularity of MR in research with relatively little effort and expense. Alongside, methodologies related to MR analyses have been considerably advanced over the past ten years. Various novel research methods and extensions of the basic MR design have been proposed in MR analyses attempting to account for possible violations of the instrumental variable assumptions and to overcome potential pitfalls[38]. These methodological approaches include sensitivity analyses of weighted-median estimator[39] and MR-Egger[40] to handle pleiotropy, bidirectional-MR to evaluate causal directions of effect between two traits[41], two-step MR[42] and network MR[43] to assess mediation, multivariable MR to deal with genetically correlated exposures[44], multifactorial MR to account for exposure interaction[45], and MR-clust to cluster genetic variants with similar causal effects[46], among others. These developments substantially raise the implementation and transparency with relative ease to perform MR analyses, especially using the two-sample approach[47]. In the last decade, there has been an exponential acceleration in the number of publications on MR, both in methodology and application, from 61 hits in 2010 to about 1300 in 2021 retrieved from PubMed.

Although available methods allow for rigorous analysis and robust causal inference with different situations, almost all methods assume a linear relationship between the exposure and outcome. It is worth emphasizing that the associations between oxidative stress with CVD are unlikely to be linear[48]. The contributions to a large proportion of CVD from oxidative stress are predominantly among individuals who have excessive amounts of oxidants or insufficient amounts of antioxidants. It is thus important to be aware that any of the relationships are more likely to be detected in specific groups that have a low antioxidants concentration or high oxidative damage. An example goes with the vitamin D debate: high circulating 25-hydroxyvitamin (25[OH]D) are associated with a decreased risk of several diseases and mortality in epidemiological observational studies, whereas RCTs of vitamin D supplementation or two-sample MR approach using summary statistics showed null findings. Interestingly, stratified MR analyses based on individual-level data according to baseline 25(OH)D concentrations suggest a causal relationship between 25(OH)D concentrations and mortality

for individuals with low vitamin D status[49]. Therefore, the inclusion of the overall population in MR studies might be inappropriate and will bias the estimate towards null unless a clear linear association has been shown. The use of proper data and methods[50,51] could identify causal non-linear relationships that might otherwise be undetected.

Furthermore, despite the identified risk factors that could be used for risk classification using prediction, the translation of the effects obtained from MR to clinical interventions should be done with caution. The MR estimates are the effects of genetically predicted risk factors on outcomes, thus representing the effects of lifelong differences in the level of the investigated risk factors. In the real-world studies, therapeutically induced changes in the risk factors by interventions in RCTs normally last only for the duration of the trial, thus representing a short- to medium- period effect of the differences in the risk factor. However, most causal risk factors seem to have cumulative effects on the outcome over time. Consequently, the anticipated effects of RCTs evaluating therapies targeted causal risk factors might be smaller than those from MR studies, especially in chronic diseases with a long development time. Notably, lifelong exposure to per unit lower plasma low-density lipoprotein showed a 3-fold greater reduction in the risk of coronary heart disease than that observed in RCTs with statin treatment started later in life[52].

Taken together, implementation of the best practical MR methodology is needed to infer causality between risk factors with chronic diseases. A cautious application of the results to intervention with optimal strategies will lead to potential targets for the prevention and treatment of chronic diseases.

References

1. R. Brigelius-Flohé, M. G. Traber. Vitamin E: function and metabolism. Faseb J 1999; 13(10): 1145-55.
2. G. S. Huggins, E. Simsolo. Application of Mendelian Randomization to the Oxidative Modification Hypothesis. J Am Coll Cardiol 2021; 77(1): 55-6.
3. M. Hacquebard, Y. A. Carpentier. Vitamin E: absorption, plasma transport and cell uptake. Curr Opin Clin Nutr Metab Care 2005; 8(2): 133-8.
4. G. Bobe, T. J. Cobb, S. W. Leonard, et al. Increased static and decreased capacity oxidation-reduction potentials in plasma are predictive of metabolic syndrome. Redox Biol 2017; 12: 121-8.
5. E. Mah, T. N. Sapper, C. Chitchumroonchokchai, et al. α-Tocopherol bioavailability is lower in adults with metabolic syndrome regardless of dairy fat co-ingestion: a randomized, double-blind, crossover trial. Am J Clin Nutr 2015; 102(5): 1070-80.
6. W. L. Hall, Y. M. Jeanes, J. K. Lodge. Hyperlipidemic subjects have reduced uptake of newly absorbed vitamin E into their plasma lipoproteins, erythrocytes, platelets, and lymphocytes, as studied by deuterium-labeled alpha-tocopherol biokinetics. J Nutr 2005; 135(1): 58-63.
7. I. Hochberg, E. M. Berinstein, U. Milman, C. Shapira, A. P. Levy. Interaction Between the Haptoglobin Genotype and Vitamin E on Cardiovascular Disease in Diabetes. Curr Diab Rep 2017; 17(6): 42.
8. F. N. Ashar, Y. Zhang, R. J. Longchamps, et al. Association of Mitochondrial DNA Copy Number With Cardiovascular Disease. JAMA Cardiol 2017; 2(11): 1247-55.
9. A. A. El-Aal, E. A. A. El-Ghffar, A. A. Ghali, M. R. Zughbur, M. M. Sirdah. The effect of vitamin C and/or E supplementations on type 2 diabetic adult males under metformin treatment: A single-blinded randomized controlled clinical trial. Diabetes Metab Syndr 2018; 12(4): 483-9.
10. A. H. Miller, C. L. Raison. The role of inflammation in depression: from evolutionary imperative to modern treatment target. Nat Rev Immunol 2016; 16(1): 22-34.
11. J. W. Kinney, S. M. Bemiller, A. S. Murtishaw, A. M. Leisgang, A. M. Salazar, B. T. Lamb. Inflammation as a central mechanism in Alzheimer's disease. Alzheimers Dement 2018; 4: 575-90.
12. C. Reitz, R. Mayeux. TREM2 and neurodegenerative disease. N Engl J Med 2013; 369(16): 1564-5.
13. T. Jonsson, H. Stefansson, S. Steinberg, et al. Variant of TREM2 associated with the risk of Alzheimer's disease. N Engl J Med 2013; 368(2): 107-16.
14. E. M. Bradshaw, L. B. Chibnik, B. T. Keenan, et al. CD33 Alzheimer's disease locus: altered monocyte function and amyloid biology. Nat Neurosci 2013; 16(7): 848-50.
15. J. C. Lambert, S. Heath, G. Even, et al. Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. Nat Genet 2009; 41(10): 1094-9.
16. J. C. Lambert, C. A. Ibrahim-Verbaas, D. Harold, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet 2013; 45(12): 1452-8.
17. K. A. Walker, B. N. Ficek, R. Westbrook. Understanding the Role of Systemic Inflammation in Alzheimer's Disease. ACS Chem Neurosci 2019; 10(8): 3340-2.

18. B. A. in t' Veld, A. Ruitenberg, A. Hofman, et al. Nonsteroidal antiinflammatory drugs and the risk of Alzheimer's disease. N Engl J Med 2001; 345(21): 1515-21.
19. P. P. Zandi, J. C. Anthony, K. M. Hayden, K. Mehta, L. Mayer, J. C. Breitner. Reduced incidence of AD with NSAID but not H2 receptor antagonists: the Cache County Study. Neurology 2002; 59(6): 880-6.
20. A. G. Yip, R. C. Green, M. Huyck, L. A. Cupples, L. A. Farrer. Nonsteroidal anti-inflammatory drug use and Alzheimer's disease risk: the MIRAGE Study. BMC Geriatr 2005; 5: 2.
21. D. Jaturapatporn, M. G. Isaac, J. McCleery, N. Tabet. Aspirin, steroidal and non-steroidal anti-inflammatory drugs for the treatment of Alzheimer's disease. Cochrane Database Syst Rev 2012; (2): Cd006378.
22. M. Miguel-Álvarez, A. Santos-Lozano, F. Sanchis-Gomar, et al. Non-steroidal anti-inflammatory drugs as a treatment for Alzheimer's disease: a systematic review and meta-analysis of treatment effect. Drugs Aging 2015; 32(2): 139-47.
23. K. M. Onuska. The Dual Role of Microglia in the Progression of Alzheimer's Disease. J Neurosci 2020; 40(8): 1608-10.
24. A. D. Thome, A. Faridar, D. R. Beers, et al. Functional alterations of myeloid cells during the course of Alzheimer's disease. Mol Neurodegener 2018; 13(1): 61.
25. B. J. Gu, X. Huang, A. Ou, et al. Innate phagocytosis by peripheral blood monocytes is altered in Alzheimer's disease. Acta Neuropathol 2016; 132(3): 377-89.
26. S. H. Hegazy, J. Q. Thomassen, I. J. Rasmussen, B. G. Nordestgaard, A. Tybjaerg-Hansen, R. Frikke-Schmidt. C-reactive protein levels and risk of dementia-Observational and genetic studies of 111,242 individuals from the general population. Alzheimers Dement 2022; Online ahead of print.(doi: 10.1002/alz.12568).
27. K. L. Rasmussen, B. G. Nordestgaard, R. Frikke-Schmidt, S. F. Nielsen. An updated Alzheimer hypothesis: Complement C3 and risk of Alzheimer's disease-A cohort study of 95,442 individuals. Alzheimers Dement 2018; 14(12): 1589-601.
28. K. L. Rasmussen, A. Tybjaerg-Hansen, B. G. Nordestgaard, R. Frikke-Schmidt. APOE and dementia - resequencing and genotyping in 105,597 individuals. Alzheimers Dement 2020; 16(12): 1624-37.
29. K. L. Rasmussen, A. Tybjærg-Hansen, B. G. Nordestgaard, R. Frikke-Schmidt. Plasma apolipoprotein E levels and risk of dementia: A Mendelian randomization study of 106,562 individuals. Alzheimers Dement 2018; 14(1): 71-80.
30. J. Adamski. Genome-wide association studies with metabolomics. Genome Med 2012; 4(4): 34.
31. G. Kastenmüller, J. Raffler, C. Gieger, K. Suhre. Genetics of human metabolism: an update. Hum Mol Genet 2015; 24(R1): R93-r101.
32. K. Suhre, M. I. McCarthy, J. M. Schwenk. Genetics meets proteomics: perspectives for large population-based studies. Nat Rev Genet 2021; 22(1): 19-37.
33. N. Brandes, N. Linial, M. Linial. PWAS: proteome-wide association study-linking genes and phenotypes by functional variation in proteins. Genome Biol 2020; 21(1): 173.
34. C. Sudlow, J. Gallacher, N. Allen, 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(3): e1001779.

35. L. Leitsalu, T. Haller, T. Esko, et al. Cohort Profile: Estonian Biobank of the Estonian Genome Center, University of Tartu. Int J Epidemiol 2015; 44(4): 1137-47.
36. Z. Chen, J. Chen, R. Collins, et al. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. Int J Epidemiol 2011; 40(6): 1652-66.
37. J. M. Gaziano, J. Concato, M. Brophy, et al. Million Veteran Program: A mega-biobank to study genetic influences on health and disease. J Clin Epidemiol 2016; 70: 214-23.
38. R. C. Richmond, G. Davey Smith. Mendelian Randomization: Concepts and Scope. Cold Spring Harb Perspect Med 2022; 12(1).
39. J. Bowden, G. Davey Smith, P. C. Haycock, S. Burgess. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. Genet Epidemiol 2016; 40(4): 304-14.
40. J. Bowden, G. Davey Smith, S. Burgess. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol 2015; 44(2): 512-25.
41. N. J. Timpson, B. G. Nordestgaard, R. M. Harbord, et al. C-reactive protein levels and body mass index: elucidating direction of causation through reciprocal Mendelian randomization. Int J Obes (Lond) 2011; 35(2): 300-8.
42. C. L. Relton, G. Davey Smith. Two-step epigenetic Mendelian randomization: a strategy for establishing the causal role of epigenetic processes in pathways to disease. Int J Epidemiol 2012; 41(1): 161-76.
43. S. Burgess, R. M. Daniel, A. S. Butterworth, S. G. Thompson. Network Mendelian randomization: using genetic variants as instrumental variables to investigate mediation in causal pathways. Int J Epidemiol 2015; 44(2): 484-95.
44. S. Burgess, S. G. Thompson. Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects. Am J Epidemiol 2015; 181(4): 251-60.
45. J. M. B. Rees, C. N. Foley, S. Burgess. Factorial Mendelian randomization: using genetic variants to assess interactions. Int J Epidemiol 2020; 49(4): 1147-58.
46. C. N. Foley, A. M. Mason, P. D. W. Kirk, S. Burgess. MR-Clust: clustering of genetic variants in Mendelian randomization with similar causal estimates. Bioinformatics 2021; 37(4): 531-41.
47. S. Burgess, A. Butterworth, S. G. Thompson. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol 2013; 37(7): 658-65.
48. F. Petermann-Rocha, F. K. Ho, H. Foster, et al. Nonlinear Associations Between Cumulative Dietary Risk Factors and Cardiovascular Diseases, Cancer, and All-Cause Mortality: A Prospective Cohort Study From UK Biobank. Mayo Clin Proc 2021; 96(9): 2418-31.
49. Emerging Risk Factors Collaboration/EPIC-CVD/Vitamin D Studies Collaboration. Estimating dose-response relationships for vitamin D with coronary heart disease, stroke, and all-cause mortality: observational and Mendelian randomisation analyses. Lancet Diabetes Endocrinol 2021; 9(12): 837-46.
50. S. Burgess, N. M. Davies, S. G. Thompson. Instrumental variable analysis with a nonlinear exposure-outcome relationship. Epidemiology 2014; 25(6): 877-85.
51. J. R. Staley, S. Burgess. Semiparametric methods for estimation of a nonlinear exposure-outcome relationship using instrumental variables with application to Mendelian randomization. Genet Epidemiol 2017; 41(4): 341-52.

52. B. A. Ference, W. Yoo, I. Alesh, et al. Effect of long-term exposure to lower low-density lipoprotein cholesterol beginning early in life on the risk of coronary heart disease: a Mendelian randomization analysis. J Am Coll Cardiol 2012; 60(25): 2631-9.

Nederlandse samenvatting

Het hoofddoel van dit proefschrift is het ontrafelen van de causale rol van oxidatieve stress, gemeten door antioxidanten spiegels in bloed en mitochondriaal DNA-kopieaantal (mtDNA-CN), bij het ontstaan van chronische ziekten. In dit proefschrift is de nadruk gelegd op het onderzoek naar cardiometabole ziekten en verwante risicofactoren door het toepassen van conventionele epidemiologische technieken in combinatie met Mendeliaanse Randomisatie. De studies beschreven in dit proefschrift hebben antwoord gegeven op twee hoofdvragen na toepassing van innovatieve epidemiologische onderzoeksmethoden op populatieniveau, namelijk: 1) of mitochondriale disfunctie een causale risicofactor is in de ontwikkeling van atherosclerotische cardiovasculaire ziekten en verwante risicofactoren (**Deel II**), en 2) of antioxidanten in de voeding klinisch relevante voordelen opleveren bij de preventie van atherosclerotische cardiovasculare ziekten (**Deel III**). Daarnaast werpt dit proefschrift licht op de rol van ontsteking, welke een bron kan zijn van verhoogde oxidatieve stress , bij neurologische aandoeningen (**Deel IV**).

De Mendeliaanse Randomisatie techniek gebruikt genetische varianten die sterk verband houden met de blootstelling om het causale effect van modificeerbare risicofactoren op uitkomsten te onderzoeken. Genetische varianten die reeds bij de conceptie worden geërfd worden willekeurig doorgegeven tijdens de meiose en kunnen worden gebruikt als proxies van de blootstellingsniveaus. Als gevolg van het willekeurig doorgeven van overerfbaar materiaal worden individuen verdeeld in twee vergelijkbare groepen. Degenen die het effect-allel dragen (bijvoorbeeld met een verhoogd blootstellingsniveau, zoals een verhoogd LDL cholesterol) worden toegewezen aan de groep met hogere blootstellingsniveaus, terwijl degenen die het alternatieve allel dragen, worden toegewezen aan de groep met lagere blootstellingsniveaus. Omdat al het genetisch materiaal willekeurig wordt doorgegeven, kunnen genetische varianten niet gerelateerd zijn aan mogelijk verstorende factoren die de relatie tussen de blootstelling en de uitkomst beïnvloeden en kan elk verschil in uitkomsten tussen genetisch gedefinieerde groepen direct worden toegeschreven aan de blootstelling. Om deze reden wordt Mendeliaanse Randomizatie beschouwd als een ”natuurlijk experiment”, dat een gerandomizeerde klinische studie kan nabootsen, waarbij individuen worden gerandomiseerd op basis van een genetische variant in plaats van een bepaalde interventie (bijvoorbeeld een farmeutisch middel).

In **Hoofdstuk 2** wordt de huidige biologische kennis van oxidatieve stress en de relatie met veroudering en ouderdomsgerelateerde ziekten in experimentele en epidemiologische studies besproken. Ondanks dat reactieve zuurstofradicalen nauw betrokken zijn bij het in stand houden van de normale celfunctie via een breed scala aan signaalroutes, leidt overproductie tot oxidatieve schade aan macromoleculen en verstoort het de mitochondriële functie, wat kan leiden tot een kortere levensduur en de manifestatie van meerdere ouderdomsgerelateerde ziekten. De rol van oxidatieve stress in het ontstaan van ziekten, zoals hart- en vaatziekten, is echter nog niet eenduidig; vooral op basis van de tegenstrijdige bevindingen van verschillende antioxidanten in observationele epidemiologische cohort studies en gerandomiseerde klinische onderzoeken. Desalniettemin

moeten verschillende zwakke punten in het onderzoeksontwerp en het concept van antioxidatieve capaciteit in overweging worden genomen bij het interpreteren van de tegenstrijdige resultaten. Bovendien zijn nieuwe en specifieke biomarkers van oxidatieve schade gerechtvaardigd om het effect van suppletie met antioxidanten beter te kunnen volgen in het lichaam.

In de algemene bevolking, worden veranderingen in de hoeveelheid mitochondriaal DNA (mtDNA-CN) in witte bloed cellen verondersteld een goede proxy te zijn voor de mitochondriale functie. Om deze reden wordt het mtDNA-CN gezien als een aantrekkelijke biomarker omdat dit relatief gemakkelijk in grote aantallen kan worden gemeten. In **Hoofdstuk 3 en 4** is, in lijn met de principes van triangulatie in etiologische (epidemiologische) studies, de rol van mtDNA-CN in cardiovasculaire ziekten en bloed metabolieten onderzocht met behulp van een combinatie van een prospectief cohortontwerp in de grote UK Biobank, entoepassing van de Mendeliaanse Randomisatie techniek. In **Hoofdstuk 3** lieten observationele prospectieve analyses onder deelnemers die vrij waren van hart- en vaatziekten bij inclusie van de studie zien dat een laag mtDNA-CN een risicofactor was voor zowel het ontwikkelen van coronair vaatlijden als hartfalen. Met behulp van Mendeliaanse Randomisatie, op basis van gegevens die zijn gegenereerd door de momenteel grootste genetische consortia en biobanken, werden deze associaties gevalideerd voor coronair vaatlijden, maar niet voor hartfalen. Het uitblijven van het vinden van een relatie tussen mtDNA-CN en hartfalen met Mendeliaanse Randomisatie is waarschijnlijk te wijten aan de fenotypische heterogeniteit van hartfalen. In **Hoofdstuk 4** hebben we de associatie tussen mtDNA-CN met 168 bloedmetabolieten van voornamelijk lipiden en lipoproteïnen (sub)deeltjes verder onderzocht met behulp van het Nightingale NMR-gebaseerde metabolomics platform. We hebben associaties waargenomen tussen laag mtDNA-CN en een meer atherogeen metabolieten profiel, dat gekenmerkt wordt door hogere niveaus van de meeste lipiden in bloed. Deze bevindingen suggereren dat mitochondriale dysfunctie, zoals benaderd door mtDNA-CN, een invloed kan hebben op lipidendysregulatie, welke een goed gedocumenteerde risicofactor is voor het ontstaan van atherosclerose en een causale risicofactor is in de pathogenese van atherosclerotische cardiovasculaire ziekten. Verdere studies zijn nodig om de validiteit hiervan in de risicoclassificatie van patiënten en ziektepreventie te onderzoeken.

In **Hoofdstuk 5** onderzochten we het mogelijke causale verband tussen meerdere circulerende antioxidanten en coronaire hartziekten met behulp van Mendeliaanse Randomisatie op basis van data gegenereerd door 3 grote studies en op basis van 768.121 deelnemers van wie 93.230 mensen coronaire hartziekten hadden ontwikkeld. In dit hoofdstuk is bewijs geleverd dat genetisch voorspelde en dus levenslang hogere circulerende antioxidantenniveaus, hetzij als authentieke circulerende niveaus die vergelijkbaar zijn met de spiegels die bereikt kunnen worden door voedingssupplementen of als concentraties van overeenkomstige metabolieten, het risico op coronaire hartziekte niet verlagen. Het is echter mogelijk dat de circulerende antioxidantniveaus niet gelijk zijn aan de authentieke functionele niveaus, vooral in het geval van vitamine E met zijn duidelijk katabolisme na oxidatieve modificatie. Dit benadrukt de noodzaak om markers die specifiek zijn voor antioxidantcapaciteit, dat wil zeggen de

functionele niveaus van antioxidanten, te koppelen aan hart- en vaatziekten en gerelateerde risicofactoren.

Vitamine E kan worden afgebroken via enzymatische routes in de lever of worden geoxideerd in de periferie. In de leverroute wordt vitamine E enzymatisch omgezet in een spectrum van enzymatische metabolieten. Als alternatief werkt vitamine E als peroxyllipidenradicalenvanger en vormt het geoxideerde metabolieten met de opening van de chromanolring. Deze metabolieten worden voornamelijk via de urine door het lichaam uitgescheiden. In **Hoofdstuk 6 en 7** hebben we ons gericht op de cross-sectionele associaties tussen circulerende vitamine E niveaus in bloed en enzymatische en geoxideerde metabolieten niveaus in de urine, en cardiometabolische eigenschappen bij ongeveer 500 gezonde deelnemers van middelbare leeftijd uit de Nederlandse Epidemiologie van Obesitas Studie. In **Hoofdstuk 6** wordt beschreven dat hogere niveaus van geoxideerde metabolieten in de urine, maar niet de circulerende vitamine E niveaus in bloed en de enzymatisch gemetaboliseerde metabolieten niveaus in urine, geassocieerd waren met een lagere insulineresistentie (als voorloper voor het onstaan van type 2 diabetes). Daarnaast beschrijft **Hoofdstuk 7** dat de associaties van 147 NMR-gebaseerde metabole biomarker metingen, meestal bestaande uit lipiden en lipoproteïne (sub)deeltjes, met enzymatische metabolieten niveaus, dezelfde richting hebben als die met circulerende vitamine E. Echter, associaties van de metabole biomarkers met geoxideerde metabolieten waren duidelijk verschillend van die met zowel circulerende vitamine E als enzymatische metabolieten. Deze bevindingen benadrukken dat circulerend vitamine E representatief kan zijn voor het enzymatische katabolisme, maar niet voor de antioxidatieve functie van vitamine E.

Ontsteking is onlosmakelijk verbonden met oxidatieve stress. Hoewel de associatie tussen ontsteking en atherosclerotische hart- en vaatziekten welbekend is, is de rol hiervan bij neurologische en neurodegeneratieve aandoeningen niet duidelijk. In **Hoofdstuk 8** wordt een bidirectionele associatie beschreven tussen inflammatoire darmziekte, als een ziektemodel van aanhoudende chronische ontsteking, en depressie met behulp van Mendeliaanse Randomisatie in een gecombineerde steekproefomvang van 693.183 individuen (36.507 ziektegevallen) voor inflammatoire darmziekte en 534.635 individuen (71.466 geziektegevallen) voor depressie. Er werd geen verband gevonden tussen genetisch beïnvloede inflammatoire darmziekte en het risico op depressie, terwijl genetisch voorspelde depressie geassocieerd was met een hoger risico op inflammatoire darmziekte. In **Hoofdstuk 9** hebben we het causale effect van 41 systemische inflammatoire markers op cognitieve functie en hersenatrofie onderzocht met behulp van Mendeliaanse Randomisatie. Na correctie voor het doen van meerdere statistische toetsen werd geen significante associatie waargenomen tussen 40 van de 41 ontstekingsmarkers en één van de onderzochte uitkomsten. Deze resultaten konden echter de rol van ontsteking bij neurologische aandoeningen niet weerleggen. Het is mogelijk dat in plaats van een oorzaak, een overmatige ontstekingsreactie ook een gevolg kan zijn van neurodegeneratieve aandoeningen.

In het laatste hoofdstuk (**Hoofdstuk 10, Deel V**) heb ik de belangrijkste bevindingen van dit proefschrift samengevat (zoals hierboven vermeld), en de toekomst-

perspectieven besproken op basis van de implicaties van het werk beschreven in dit proefschrift. Voor antioxidanten zal verdere verkenning van antioxidanten bij de preventie en behandeling van hart- en vaatziekten belangrijk blijven gezien hun duidelijke rol in de pathogenese van ziekten. Goed geselecteerde deelnemers voor suppletie en betere biomarkers om suppletie-effecten in onderzoeken te volgen, kunnen meer inzicht geven in de rol van antioxidanten bij hart- en vaatziekten. Bovendien kan ontsteking op een complexe manier geassocieerd worden met neurologische aandoeningen. Dit benadrukt het belang om rekening te houden met de dynamiek van ontstekingen bij de ontwikkeling en progressie van ziekten, met name de verschuiving van een gunstige naar een schadelijke rol. Een diepgaand begrip van de dubbele en complexe rol van inflammatie kan veelbelovende kansen bieden om nieuwe biomarkers te identificeren voor de vroege detectie van ziekten en om proefontwerpen voor de ontwikkeling van geneesmiddelen te vergemakkelijken. Belangrijk is dat implementatie van de beste praktische Mendeliaanse randomisatiemethoden nodig is om causaliteit tussen risicofactoren met chronische ziekten af te leiden. Een voorzichtige toepassing van de resultaten op interventies met optimale strategieën zal leiden tot potentiële doelen voor de preventie en behandeling van chronische ziekten.

List of publications

Published articles included in this thesis

1. **Luo J**, le Cessie S, Blauw GJ, et al. Systemic inflammatory markers in relation to cognitive function and measures of brain atrophy: A Mendelian Randomization study. Geroscience. 2022. Epub ahead of print.
2. **Luo J**, Hashimoto Y, Martens LG et al. Associations of metabolomic profiles with circulating vitamin E and urinary vitamin E metabolites in middle-aged individuals. Nutrition 2022; 93:111440.
3. **Luo J**, Xu Z, Noordam R, van Heemst D, Li-Gao R. Depression and inflammatory bowel disease: A bidirectional two-sample Mendelian randomization study. J Crohns Colitis. 2022. 16(4): p. 633-642.
4. **Luo J**, le Cessie S, van Heemst D, Noordam R. Diet-derived circulating antioxidants and risk of coronary heart disease: A Mendelian randomization study. J Am Coll Cardiol. 2021; 77:45-54.
5. **Luo J**, Meulmeester FL, Martens LG, 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; 40:4192-4200.
6. **Luo J**, Mills K, le Cessie S, Noordam R, van Heemst D. Ageing, age-related diseases and oxidative stress: What to do next? Ageing Res Rev. 2020; 57:100982.

Other publications

7. 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;10: e022567. (*co-first author)
8. Wang W, **Luo J**, Willems van Dijk K, et al. Assessment of the bi-directional relationship between blood mitochondrial DNA copy number and type 2 diabetes mellitus: a multivariable-adjusted Mendelian Randomization study. Diabetologia. 2022.
9. Martens LG, **Luo J**, Willems van Dijk K, et al. The association between lymphocyte mitochondrial DNA abundance and Stroke: a combination of multivariable-adjusted survival and mendelian randomization analyses. Atherosclerosis. 2022.
10. Meulmeester FL, **Luo J**, Martens LG, et al. Association of measures of body fat with serum alpha-tocopherol and its metabolites in middle-aged individuals. Nutr Metab Cardiovasc Dis. 2021;31:2407-2415.
11. Martens LG, **Luo J**, Meulmeester FL, et al. Associations between Lifestyle Factors and Vitamin E Metabolites in the General Population. Antioxidants. 2020 Dec 15;9(12):1280.
12. Faquih T, van Smeden M, **Luo J** et al. A Workflow for missing values imputation of untargeted metabolomics data. Metabolites 2020;10 (12).

Other manuscripts

13. **Luo J**, Noordam R, Jukema JW, et al. Low mitochondrial DNA abundance drives atherosclerotic cardiovascular disease: cohort and genetic studies. In revision. 2022.
14. **Luo J**, Noordam R, et al. Associations of mtDNA copy number with circulating lipoprotein, lipid, and metabolite levels. Under submission.

PhD Portfolio

Courses	Years	Hours
PhD Introductory Meeting	2018	5
Introduction to Clinical Epidemiology	2018	56
Basic Methods and Reasoning in Biostatistics	2018	42
Academic Writing for PhDs	2019	60
Regression Analysis	2019	42
Clinical Epidemiology	2019	56
Prediction modelling and intervention research	2019	84
Causal Inference	2020	84
Statistical Aspects of Clinical Trials	2021	42
Meta-analysis	2021	28
Survival Analysis	2021	42
Analysis of Repeated Measurements	2021	42
Congress attendance	**Years**	**Hours**
Oral presentation 45th Annual Dutch Diabetes Research Meeting (NVDO)	2019	7
Oral presentation 6th Junior Dutch Endocrinology Meeting (JNVE)	2019	7
Poster presentation 26th European and International Congress on Obesity (ECOICO)	2020	14
Poster presentation 5th International conference on Mendelian randomization	2021	14
Oral presentation The Future of Mendelian Randomization Studies 2020	2021	21
Student supervision	**Years**	**Hours**
Yasufumi Hashimoto (exchange project)	2021	28
Maël Thielman (master thesis)	2021	42
Romy Kwakernaak (master thesis)	2021	42
Personal grant	**Years**	**Hours**
Chinese Scholarship Council	2018	-
Leiden University Fonds	2021	-

Curriculum Vitae

Jiao Luo was born on the 23rd of December 1991 in Hongya, China. She attended medical school at Sichuan University (China) from 2009 to 2014, whereafter she obtained her Master of Public Health degree in 2017 at the same university. She worked as a junior researcher handling data from clinical trials in the National Engineering Research Center of Immunological Products in Chongqing, China, between 2017 and 2018.

In 2018, she received a scholarship from the Chinese Scholarship Council to start as a PhD student at the Departments of Clinical Epidemiology and Internal Medicine (Section of Gerontology and Geriatrics) at the Leiden University Medical Center (LUMC), under the supervision of Prof. dr Saskia le Cessie, Dr. Diana van Heemst, and. Dr. Raymond Noordam. She spent part of her PhD abroad in Denmark at the Department of Clinical Biochemistry, Rigshospitalet - Copenhagen University Hospital, under the supervision of Prof. dr. Ruth Frikke-Schmidt. This research period was supported by a grant from the Leiden University Fonds (LUF).

From April 2022 onwards, she started as a postdoctoral researcher in the group of Prof. dr. Ruth Frikke-Schmidt at the Department of Clinical Biochemistry, Rigshospitalet - Copenhagen University Hospital.

Acknowledgements

I would like to thank all participants and staff of the Netherlands Epidemiology of Obesity study and UK Biobank as well as the genetics consortia on which the studies in this thesis are based.

This thesis would not have been possible without the help in many ways from many people.

Foremost, I would like to express my greatest appreciation to my promotor, prof. dr. Saskia le Cessie for your guidance, advice, and support. Saskia, if there would ever be an award to the best promotor, you are the one, being exceptionally professional but extremely nice and approachable.

I would like to extend my gratitude to my co-promotor, dr. Diana van Heemst. Diana, I am always amazed by how biologically knowledgeable you are, and thanks for your tremendous help in all discussions and for always being supportive of any projects I wanted to explore.

To my co-promotor, dr. Raymond Noordam. Raymond, thank you for taking me on this PhD project, for providing your valuable daily supervision, hands-on mentoring, and for allowing me to develop as an independent researcher. Thanks for your support in exploring new topics and for answering all my questions.

To my lovely colleagues at the epi and ouderengeneeskunde departments. I am so thankful for the 'gezellige' time throughout this journey, especially roomies from C7-125 who helped me learn and practice the word 'gezellig'.

I am grateful to my friends both from Leiden and back home in China for your company and support. A special acknowledgement to 胡静, who takes care of my parents during my absence. You know that we are more than friends, but non-blood-related sisters.

I would like to thank my parents, who offered me the priceless treasure in this world - education. Thank you for your unwavering encouragement and for shaping me into an independent, resilient, and thoughtful person. 我要诚挚地感谢我的父母。感谢你们为我提供了良好的家庭和教育环境，并把我塑造成了一个独立、坚忍和思辨的人。

Finally, to my best friend and beloved, Zhongwei. You have lightened up my life with love, support, encouragement, and most importantly, twisted humor. Without you, I would not have become the person I am today. I am glad you are here by my side on this special day. You are the incomparably best.