Nutrition and cardiometabolic health:
The role of DNA methylation

Kim V.E. Braun

ACKNOWLEDGEMENTS

The studies described in this thesis were performed within the Generation R Study, The Rotterdam Study, the ARIC study, the B-PROOF Study, Nurses' Health Study, Nurses' Health Study II, and Health Professionals Follow-Up Study. We gratefully acknowledge the contributions of participants, research staff, data management, and health professionals of all studies.

The work presented in this thesis was conducted within ErasmusAGE at the Department of Epidemiology of Erasmus Medical Center, Rotterdam, The Netherlands. ErasmusAGE is a center for aging research across the life course and is funded by Nestlé Nutrition (Nestec Ltd.) and Metagenics Inc. The funders had no role in design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review or approval in the manuscripts included in this thesis.

Publication of this thesis was kindly supported by the Department of Epidemiology of Erasmus Medical Center and by Erasmus University Rotterdam. Financial support was also kindly provided by ChipSoft and Danone Nutricia Research. Additional financial support by the Dutch Heart Foundation for the publication of this thesis is gratefully acknowledged.

ISBN: 978-94-6361-053-7

Layout and print: Optima Grafische Communicatie, Rotterdam, The Netherlands
Cover design: Erwin Timmerman, Optima Grafische Communicatie

Nutrition and cardiometabolic health:
The role of DNA methylation

Voeding en cardiometabole gezondheid:
De rol van DNA methylatie

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus

Prof.dr. H.A.P. Pols

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
woensdag 7 maart 2018 om 13.30 uur

door

Kim Valeska Emilie Braun

geboren te Rotterdam

Erasmus University Rotterdam

Erasmus

PROMOTIECOMMISSIE

Promotor:	Prof.dr. O.H. Franco Duran
Overige leden:	Prof.dr. H. Boersma
	Prof.dr. E.F.C. van Rossum
	Prof.dr. J.M. Geleijnse
Copromotoren:	Dr.ir. R.G.Voortman
	Dr. A. Dehghan

Paranimfen:	V. Jen
	M.A. Berghout

To my father

TABLE OF CONTENTS

MANUSCRIPT THAT FORM THE BASIS OF THIS THESIS

Chapter 1: Introduction

Partly based on:

Braun KVE*, Portilla E*, Chowdhury R, Nano J, Troup J, Voortman T, Franco OH, Muka T. "*The role of epigenetic modifications in cardiometabolic diseases*", in: Moskalev A, Vaiserman AM. "*Epigenetics of Aging and Longevity: Translational Epigenetics*", Academic Press; 2017; p. 347-64.

Chapter 2: Nutrition & DNA methylation

Braun KVE*, Mandaviya P*, Franco OH, Nano J, Girschik C, Bramer WM, Muka T, Troup J, van Meurs JBJ, Heil SG, Voortman T. Nutrients and DNA methylation across the life course: a systematic review. *Submitted for publication.*

Chapter 3: DNA methylation & cardiometabolic health

Braun KVE*, Voortman T*, Dhana K, Troup J, Bramer WM, Troup J, Chowdhury R, Dehghan A, Muka T, Franco OH. The role of DNA methylation in dyslipidaemia: A systematic review. *Progress in Lipid Research. 2016;64:178-91.*

Braun KVE, Dhana K, de Vries PS, Voortman T, van Meurs JBJ, Uitterlinden AG, Hofman A, Hu FB, BIOS consortium, Franco OH, Dehghan A. Epigenome-wide association study (EWAS) on lipids: the Rotterdam Study. *Clinical Epigenetics. 2017;9:15.*

Dhana K*, **Braun KVE***, Nano J, Voortman T, Demerath EW, Guan W, Fornage M, van Meurs JBJ, Uitterlinden AG, Hofman A, Franco OH, Dehghan A. Epigenome-wide association study (EWAS) on obesity-related traits. *Accepted for publication in American Journal of Epidemiology.*

Chapter 4: Nutrition & cardiometabolic health in early life

Braun KVE*, Voortman T*, Kiefte-de Jong JC, Jaddoe VWV, Hofman A, Franco OH, van den Hooven EH. Dietary Intakes of Folic Acid and Methionine in Early Childhood Are Associated with Body Composition at School Age. *Journal of Nutrition. 2015;145(9):2123-9.*

Braun KVE, Erler NS, Kiefte-de Jong JC, Jaddoe VW, van den Hooven EH, Franco OH, Voortman T. Dietary Intake of Protein in Early Childhood Is Associated with Growth Trajectories between 1 and 9 Years of Age. *Journal of Nutrition. 2016;146(11):2361-7.*

Voortman T, **Braun KVE**, Kiefte-de Jong JC, Jaddoe VW, Franco OH, van den Hooven EH. Protein intake in early childhood and body composition at the age of 6 years: The Generation R Study. *International Journal of Obesity (Lond).* 2016;40(6):1018-25.

Stroobant W*, **Braun KVE***, Kiefte-de Jong JC, Moll HA, Jaddoe VWV, Brouwer IA, Franco OH, Voortman T. Intake of Different Types of Fatty Acids in Infancy Is Not Associated with Growth, Adiposity, or Cardiometabolic Health up to 6 Years of Age. *Journal of Nutrition. 2017;147(3):413-20.*

Chapter 5: Nutrition & cardiometabolic health in adults

Oliai Araghi S*, **Braun KVE***, van der Velde N, van Dijk SC, van Schoor NM, Zillikens C, de Groot L, Uitterlinden A, Stricker B, Voortman T*, Kiefte-de Jong JC*. Associations of serum folate and vitamin B12 with body composition in elderly: The B-PROOF study. *Submitted for publication.*

Braun KVE, Satija A, Voortman T, Franco OH, Sun Q, Bhupathiraju SN, Hu FB. Methyl donor nutrient intake and incidence of type 2 diabetes mellitus: results from the Nurses' Health Study and Health Professionals Follow-Up Study. *Manuscript in preparation*

Girschik C*, **Braun KVE***, Franco OH, Voortman T. Associations between macronutrient intake and incidence of coronary heart disease (CHD): The Rotterdam Study. *Manuscript in preparation*

**Denotes equal contribution*

Chapter 1

Introduction

Partly based on:

Braun KVE*, Portilla E*, Chowdhury R, Nano J, Troup J, Voortman T, Franco OH, Muka T. "*The role of epigenetic modifications in cardiometabolic diseases*", in: Moskalev A, Vaiserman AM. "*Epigenetics of Aging and Longevity: Translational Epigenetics*", Academic Press; 2017; p. 347-64.

**denotes equal contribution*

CARDIOMETABOLIC HEALTH

Cardiometabolic health plays an important role in healthy aging and longevity. Despite improvements in prevention, the prevalence of type 2 diabetes (T2D) continues to increase, and cardiovascular disease (CVD) remains the leading cause of death worldwide.[1, 2] Addressing risk factors of T2D and CVD would help to further improve prevention, but this requires a better understanding of the etiology of these cardiometabolic diseases. One of the major risk factors for CVD and T2D is obesity.[3] In addition, dyslipidemia, defined as decreased high-density lipoprotein-cholesterol (HDL-C), elevated low-density lipoprotein-cholesterol (LDL-C), and/or elevated triacylglycerol (TAG) concentrations,[4] is recognized as a prominent risk factor for CVD.[5, 6] Cardiometabolic risk factors such as obesity and dyslipidemia, are not only important for cardiometabolic disease risk in adulthood, but already during childhood. Several studies have shown that the development of cardiometabolic risk factors already begins in early life and that these risk factors track into adulthood.[7-9] As the prevalence of overweight and obesity is rising among children, these children are also at higher risk of obesity, T2D, and CVD in adulthood.[10] Therefore, gaining knowledge about factors that may influence cardiometabolic health across different stages of the life course is very relevant for early prevention of T2D and CVD.[11]

EPIGENETIC INFLUENCES ON CARDIOMETABOLIC HEALTH

Both T2D and CVD are influenced by environmental and genetic factors. Several genome-wide association studies (GWAS) have identified loci that explain a fraction of the variance in T2D and CVD or their related risk factors.[12, 13] Beyond this, the role of epigenetic determinants is increasingly recognized as a potential important link between environmental exposure and disease risk. Thus, epigenetic determinants may be a benchmark to capture the influences of environmental exposures and disease risk in cardiometabolic health.[14] Epigenetics refers to the mechanisms that affect gene expression, without changing the sequence of DNA.[15] The best understood and most studied epigenetic mechanism is DNA methylation, the attachment of a methyl group to a CpG site. Several prominent risk factors for T2D and CVD, including dyslipidemia and obesity may be regulated by DNA methylation. Expanding this knowledge can help to further unravel our understanding of underlying mechanisms that are leading to T2D and CVD. Recently, epigenome-wide association studies (EWAS) have become available, providing further insights into the DNA methylation alterations associated with complex traits and diseases, and providing an opportunity to identify epigenetic profiles that underlie these traits and disease. The study of epigenetic markers is emerging as a promising molecular strategy for risk stratification for complex diseases, and, when implemented, it could have a sizable public health and clinical impact.[16] Epidemiological studies have mainly investigated the relation between DNA methylation and cardiometabolic risk factors using a candidate-gene approach, reporting that lipid concentrations are associated with DNA methylation of several CpG sites, such as *APOE* and *ABCA1*.[17, 18] In addition to the candidate-gene approach, gene-specific DNA methylation was also studied for the whole genome. Results of these EWASs confirmed associations of known lipid-associated genes, such as *ABCG1*, but novel CpG sites were also identified. To date, only a few

studies have examined blood lipids in relation to differentially methylated sites on a genome-wide level.[19-21] However, these studies have mainly been performed in patient populations, while only one study has been performed within a population-based study. Although promising results have been reported in the field of epigenetics and dyslipidemia, our understanding of the role of epigenetics in regulating cardiometabolic risk in the general population is still limited.

THE ROLE OF NUTRITION IN DNA METHYLATION AND CARDIO-METABOLIC HEALTH

Considering that DNA methylation is reversible and can be influenced by environmental factors, future therapies targeting the epigenome may be a novel strategy to prevent and treat dyslipidemia, obesity, and subsequently decreasing the risk of T2D and CVD. Nutrition is one of the environmental factors which can affect DNA methylation, and consequently several health outcomes.[22-24] Some nutrients are directly involved in methylation of DNA, such as vitamin B2, vitamin B6, vitamin B12, folate and methionine, which are also known as methyl donor nutrients. These nutrients act as co-factors in the one-carbon metabolism, resulting in the forming of s-adenosylmethionine (SAM), which is the primary methyl donor.[25] Deficiency of these methyl donor nutrients could cause dysregulation of DNA methylation and may lead to disturbed energy and lipid metabolism, increasing the risk of cardiometabolic diseases.[26-28] Findings from animal studies show that DNA methylation is prone to modification by external factors in early life. The effect of methyl donor nutrients on DNA methylation was studied for example in agouti mice, which are genetically predisposed to diabetes and obesity. When these mice received methyl donor nutrient supplementation during pregnancy, their offspring were at lower risk of obesity and diabetes compared to the offspring of which their mothers did not receive methyl donor nutrients.[29] This difference in phenotype was caused by differential DNA methylation. Results from other animal studies showed that supplementation of methyl donor nutrients affected DNA methylation and subsequenlty reduced liver fat accumulation,[30] and have a protective effect on the development of obesity.[31, 32] In addition to methyl donor nutrients, also other nutrients, such as protein and fatty acids, may have an effect on DNA methylation. For example, offspring of pigs fed a low-protein diet during gestation had lower DNA methylation at cardiometabolic genes compared to those fed a high-protein diet.[33-35] Furthermore, also high-fat diets have been shown to alter DNA methylation at metabolic genes in mice.[36] The influence of nutrition on DNA methylation has also been studied in humans, both during pregnancy or in other lifecourse stages, such as adolescence and adulthood. These studies suggest that higher intake of methyl donor nutrients, such as folate or vitamin B12, is associated with differential DNA methylation.[25] However, there is still a lot of inconsistency in these results and further studies are needed to elucidate these associations. As epigenetic mechanisms may be affected by nutrition, more research should be performed on epigenetic therapy strategies to reduce the high burden of T2D and CVD, establishing potential novel therapeutic and preventive strategies in cardiometabolic risk.[37] Due to the modifiability of diet, optimizing nutrition is an eminently suitable strategy for prevention of cardiometabolic diseases.

OBJECTIVES

The overall aim of this thesis was to investigate associations between nutritional factors, DNA methylation, and cardiometabolic health in children as well as adults. (**Figure 1.1.**). Therefore, the following objectives were:

1. To investigate nutritional factors associated with DNA methylation.
2. To identify differentially methylated CpG sites in relation to cardiometabolic risk factors.
3. To examine associations between nutrition and cardiometabolic risk factors in childhood.
4. To examine associations between nutrition and cardiometabolic diseases in adulthood.

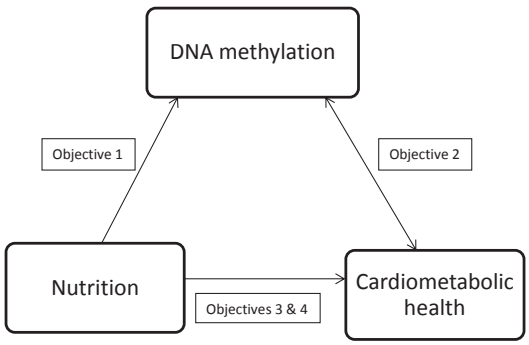


Figure 1.1. Overview of objectives included in this thesis

STUDY POPULATIONS

The studies presented in this thesis were embedded in the Generation R Study, The Rotterdam Study, the B-PROOF Study, Nurses' Health Study, Nurses' Health Study II, and Health Professionals Follow-Up Study.

The Generation R Study

The Generation R Study is a population-based prospective cohort from early fetal life onward in Rotterdam, the Netherlands.[38] Pregnant women with a delivery date between April 2002 and January 2006 were enrolled in the study, and data on follow-up in early childhood was available for 7,893 children. In this thesis, we included data on dietary intake in early childhood and growth, body composition and cardiometabolic health during follow-up. Food intake was assessed when the children had a median age of 1 year using a 211-item semi-quantitative food frequency questionnaire (FFQ), which was specifically designed for this age group.[39] Growth, including height, weight, and BMI were measured repeatedly during routine visits to Child Health Centers up to the age of 4 years. At the age of 6 years, children visited our research center in Erasmus Medical Center for a detailed physical examination.[40] During this visit we measured not only height, weight, and BMI, but also body fat mass and fat-free mass, and blood samples were obtained to determine concentrations of insulin, TAG, total cholesterol, HDL-cholesterol, LDL-cholesterol, and C-peptide.

The Rotterdam Study

The Rotterdam Study is a large prospective, population-based cohort aimed at assessing the occurrence of and risk factors for chronic diseases (cardiovascular, endocrine, hepatic, neurological, ophthalmic, psychiatric, dermatological, oncological, and respiratory) in the middle-aged and elderly.[41] A total of 14,926 subjects, living in the well-defined Ommoord district in the city of Rotterdam in the Netherlands are included in this study. The first sub-cohort, Rotterdam Study-I (RS-I), started in 1990 and comprised of 7,983 subjects with age 55 years or above. The second sub-cohort (RS-II), started in 2000 and included 3,011 subjects who had reached an age of 45 years since 1989. The third sub-cohort, Rotterdam Study-III (RS-III), started in 2006 and consisted of 3,932 subjects aged 45 years and above. Dietary intake was measured at baseline visits of all three cohorts using validated semi-quantitative FFQs. During the visits to the research center, BMI and waist circumference were measured and participants had blood samples taken to determine concentrations of triglycerides, HDL-C, and total cholesterol. DNA methylation was measured in a random sample of 1,454 participants from the third visit of the second cohort (RS-II-3) and first and second visit of the third cohort (RS-III-1, RS-III-2).

The B-PROOF study

In Chapter 5.1 baseline data of the B-PROOF (B-vitamins for the Prevention Of Osteoporotic Fractures) study were used. The B-PROOF study is a multi-center, randomized, placebo-controlled, double-blind intervention study, investigating the effect of a 2-year daily oral vitamin B12 (500 μg) and folic acid (400 μg) supplementation on fracture incidence. The study was conducted in three research centers in the Netherlands: VU University Medical Center (Amsterdam), Wageningen University (Wageningen), and Erasmus Medical Center (Rotterdam). This study included 2919 individuals, aged 65 years and older with elevated homocysteine levels (12 - 50 μmol/l).[42] At baseline, venous blood samples were obtained and serum folate, vitamin B12, methylmalonic acid (MMA), and holo-transcobalamin (HoloTC) were determined.[42, 43] Furthermore, height, weight, and BMI were measured, and in a subsample of participants from the Amsterdam and Rotterdam region (n=1227) body fat mass and fat-free mass were measured by Dual Energy X-ray assessment (DXA). Dietary data were collected in a subsample of the Wageningen region (n=603), using a Food Frequency Questionnaire (FFQ).

Nurses' Health Study, Nurses' Health Study II, and Health Professionals Follow-Up Study

Data from three prospective cohort studies in the USA were used in chapter 5.2 of this thesis: the Nurses' Health Study (NHS), Nurses' Health Study II (NHS2), and the Health Professionals' Follow-Up Study (HPFS). The NHS started in 1976 with 121,701 female nurses aged 30-55 years, the NHS2 started in 1989 with 116,430 female nurses aged 25-42 years, and the HPFS started in 1986 with 51,529 male health professionals aged 40-75 years. In all three cohorts, information on lifestyle and medical history was obtained by questionnaires at baseline and every 2 years during follow-up, with a response rate of ~90% per cycle. For the study included in this thesis, we used data on diet and diabetes. Dietary data were collected every 2-4 years using a validated semi-quantitative FFQ consisting of ~130 food items.[44, 45] Participants in all three cohorts were asked whether they were diagnosed with

diabetes by a physician every two years. Participants who self-reported physician diagnosed diabetes were sent a supplementary questionnaire to confirm diagnosis.[46, 47]

THESIS OUTLINE

Chapter 2 provides an overview of the current evidence on the associations of different nutrients with DNA methylation in humans across all stages of the life course (i.e. during pregnancy, infancy, childhood, adolescence, and adulthood).

Chapter 3 focuses on associations of DNA methylation and cardiometabolic risk factors. **Chapter 3.1** provides an overview of the current literature on the association between DNA methylation and lipid levels. In **Chapter 3.2** results of an EWAS on lipid levels in the Rotterdam Study are presented. **Chapter 3.3** presents the findings of an EWAS on BMI and WC in the Rotterdam Study and replication of these findings in a USA-based study.

Chapter 4 focuses on associations of nutrition in early childhood with cardiometabolic health in children at school age from The Generation R Study. In **Chapter 4.1** the associations of intake of methyl donor nutrients, including vitamin B6, vitamin B12, folate, and methionine with growth and body composition are described. **Chapter 4.2** describes the association between protein intake and repeatedly measured growth. **Chapter 4.3** presents associations between protein intake and detailed body composition. **Chapter 4.4** describes associations of different types of fatty acids with body composition and cardiometabolic health in childhood.

Chapter 5 focuses on associations between nutrition and cardiometabolic health in adults. **Chapter 5.1** presents the associations of folate and vitamin B12 with body composition in the B-PROOF study. **Chapter 5.2** studied the role of methyl donor nutrients and the risk of diabetes using data from the Nurses' Health Study, Nurses' Health Study II, and Health Professionals Follow-Up Study. **Chapter 5.3** presents the associations between macronutrient intake and the risk of CHD in The Rotterdam Study.

In **Chapter 6**, the main findings are discussed as well as the methodological considerations, implications, and recommendation for future studies.

REFERENCES

1. Yusuf S, Reddy S, Ounpuu S, Anand S. Global burden of cardiovascular diseases: Part II: variations in cardiovascular disease by specific ethnic groups and geographic regions and prevention strategies. Circulation. 2001;104(23):2855-64.
2. Hunter DJ, Reddy KS. Noncommunicable diseases. The New England journal of medicine. 2013; 369(14):1336-43.
3. Berrington de Gonzalez A, Hartge P, Cerhan JR, Flint AJ, Hannan L, MacInnis RJ, et al. Body-mass index and mortality among 1.46 million white adults. The New England journal of medicine. 2010; 363(23):2211-9.
4. Otocka-Kmiecik A, Mikhailidis DP, Nicholls SJ, Davidson M, Rysz J, Banach M. Dysfunctional HDL: a novel important diagnostic and therapeutic target in cardiovascular disease? Progress in lipid research. 2012;51(4):314-24.
5. Yusuf S, Hawken S, Ôunpuu S, Dans T, Avezum A, Lanas F, et al. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. The Lancet. 2004;364(9438):937-52.
6. Mooradian AD. Dyslipidemia in type 2 diabetes mellitus. Nature clinical practice Endocrinology & metabolism. 2009;5(3):150-9.
7. Morrison JA, Glueck CJ, Woo JG, Wang P. Risk factors for cardiovascular disease and type 2 diabetes retained from childhood to adulthood predict adult outcomes: the Princeton LRC Follow-up Study. International journal of pediatric endocrinology. 2012;2012(1):6.
8. Katzmarzyk PT, Perusse L, Malina RM, Bergeron J, Despres JP, Bouchard C. Stability of indicators of the metabolic syndrome from childhood and adolescence to young adulthood: the Quebec Family Study. J Clin Epidemiol. 2001;54(2):190-5.
9. Morrison JA, Friedman LA, Gray-McGuire C. Metabolic syndrome in childhood predicts adult cardiovascular disease 25 years later: the Princeton Lipid Research Clinics Follow-up Study. Pediatrics. 2007; 120(2):340-5.
10. Weiss R, Dziura J, Burgert TS, Tamborlane WV, Taksali SE, Yeckel CW, et al. Obesity and the metabolic syndrome in children and adolescents. New England Journal of Medicine. 2004;350(23):2362-74.
11. L. Guariguata DRW, I. Hambleton, J. Beagley, U. Linnenkamp, J.E. Shaw. Global estimates of diabetes prevalence for 2013 and projections for 2035. Diabetes research and clinical practice. 2014;103(2): 137-49.
12. Consortium CAD, Deloukas P, Kanoni S, Willenborg C, Farrall M, Assimes TL, et al. Large-scale association analysis identifies new risk loci for coronary artery disease. Nat Genet. 2013;45(1):25-33.
13. International Consortium for Blood Pressure Genome-Wide Association S, Ehret GB, Munroe PB, Rice KM, Bochud M, Johnson AD, et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. Nature. 2011;478(7367):103-9.
14. Muka T, Koromani F, Portilla E, O'Connor A, Bramer WM, Troup J, et al. The role of epigenetic modifications in cardiovascular disease: A systematic review. Int J Cardiol. 2016;212:174-83.
15. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nature Reviews Genetics. 2012;13(7):484-92.
16. Muka T, Koromani F, Portilla E, O'Connor A, Bramer WM, Troup J, et al. The role of epigenetic modifications in cardiovascular disease: A systematic review. International journal of cardiology. 2016;212: 174-83.

17. Guay S-P, Brisson D, Lamarche B, Gaudet D, Bouchard L. Epipolymorphisms within lipoprotein genes contribute independently to plasma lipid levels in familial hypercholesterolemia. Epigenetics. 2014; 9(5):718-29.
18. Ma Y, Smith CE, Lai CQ, Irvin MR, Parnell LD, Lee YC, et al. Genetic variants modify the effect of age on APOE methylation in the Genetics of Lipid Lowering Drugs and Diet Network study. Aging Cell. 2015;14(1):49-59.
19. Irvin MR, Zhi D, Joehanes R, Mendelson M, Aslibekyan S, Claas SA, et al. Epigenome-wide association study of fasting blood lipids in the genetics of lipid-lowering drugs and diet network study. Circulation. 2014;130(7):565-72.
20. Pfeifferm L, Wahl S, Pilling LC, Reischl E, Sandling JK, Kunze S, et al. DNA Methylation of Lipid-Related Genes Affects Blood Lipid Levels. Circ Cardiovasc Genet. 2015.
21. Guay SP, Voisin G, Brisson D, Munger J, Lamarche B, Gaudet D, et al. Epigenome-wide analysis in familial hypercholesterolemia identified new loci associated with high-density lipoprotein cholesterol concentration. Epigenomics. 2012;4(6):623-39.
22. Anderson OS, Sant KE, Dolinoy DC. Nutrition and epigenetics: an interplay of dietary methyl donors, one-carbon metabolism and DNA methylation. The Journal of nutritional biochemistry. 2012;23(8): 853-9.
23. Park LK, Friso S, Choi SW. Nutritional influences on epigenetics and age-related disease. Proc Nutr Soc. 2012;71(1):75-83.
24. Glier MB, Green TJ, Devlin AM. Methyl nutrients, DNA methylation, and cardiovascular disease. Mol Nutr Food Res. 2014;58(1):172-82.
25. Friso S, Udali S, De Santis D, Choi SW. One-carbon metabolism and epigenetics. Mol Aspects Med. 2017;54:28-36.
26. Obeid R. The Metabolic Burden of Methyl Donor Deficiency with Focus on the Betaine Homocysteine Methyltransferase Pathway. Nutrients. 2013;5(9):3481-95.
27. Glier MB, Green TJ, Devlin AM. Methyl nutrients, DNA methylation, and cardiovascular disease. Molecular Nutrition & Food Research. 2014;58(1):172-82.
28. Fardet A. New hypotheses for the health-protective mechanisms of whole-grain cereals: what is beyond fibre? Nutrition Research Reviews. 2010;23(01):65-134.
29. Dolinoy DC. The agouti mouse model: an epigenetic biosensor for nutritional and environmental alterations on the fetal epigenome. Nutr Rev. 2008;66 Suppl 1:S7-11.
30. Cordero P, Campion J, Milagro FI, Martinez JA. Transcriptomic and epigenetic changes in early liver steatosis associated to obesity: effect of dietary methyl donor supplementation. Molecular genetics and metabolism. 2013;110(3):388-95.
31. Waterland RA, Travisano M, Tahiliani KG, Rached MT, Mirza S. Methyl donor supplementation prevents transgenerational amplification of obesity. International Journal of Obesity. 2008;32(9):1373-9.
32. Carlin J, George R, Reyes TM. Methyl donor supplementation blocks the adverse effects of maternal high fat diet on offspring physiology. PLoS ONE. 2013;8(5):e63549.
33. Altmann S, Murani E, Schwerin M, Metges CC, Wimmers K, Ponsuksili S. Maternal dietary protein restriction and excess affects offspring gene expression and methylation of non-SMC subunits of condensin I in liver and skeletal muscle. Epigenetics. 2012;7(3):239-52.
34. Altmann S, Murani E, Schwerin M, Metges CC, Wimmers K, Ponsuksili S. Dietary protein restriction and excess of pregnant German Landrace sows induce changes in hepatic gene expression and promoter methylation of key metabolic genes in the offspring. J Nutr Biochem. 2013;24(2):484-95.

35. Lillycrop KA, Phillips ES, Jackson AA, Hanson MA, Burdge GC. Dietary protein restriction of pregnant rats induces and folic acid supplementation prevents epigenetic modification of hepatic gene expression in the offspring. J Nutr. 2005;135(6):1382-6.
36. Zwamborn RA, Slieker RC, Mulder PC, Zoetemelk I, Verschuren L, Suchiman HE, et al. Prolonged high-fat diet induces gradual and fat depot-specific DNA methylation changes in adult mice. Sci Rep. 2017;7:43261.
37. Milagro FI, Mansego ML, De Miguel C, Martinez JA. Dietary factors, epigenetic modifications and obesity outcomes: progresses and perspectives. Mol Aspects Med. 2013;34(4):782-812.
38. Jaddoe VWV, van Duijn CM, Franco OH, van der Heijden AJ, van Iizendoorn MH, de Jongste JC, et al. The Generation R Study: design and cohort update 2012. European journal of epidemiology. 2012; 27(9):739-56.
39. Kiefte-de Jong JC, de Vries JH, Bleeker SE, Jaddoe VW, Hofman A, Raat H, et al. Socio-demographic and lifestyle determinants of 'Western-like' and 'Health conscious' dietary patterns in toddlers. The British journal of nutrition. 2013;109(1):137-47.
40. Kruithof CJ, Kooijman MN, van Duijn CM, Franco OH, de Jongste JC, Klaver CC, et al. The Generation R Study: Biobank update 2015. Eur J Epidemiol. 2014;29(12):911-27.
41. Hofman A, Murad SD, van Duijn CM, Franco OH, Goedegebure A, Ikram MA, et al. The Rotterdam Study: 2014 objectives and design update. Eur J Epidemiol. 2013;28(11):889-926.
42. van Wijngaarden JP, Dhonukshe-Rutten RA, van Schoor NM, van der Velde N, Swart KM, Enneman AW, et al. Rationale and design of the B-PROOF study, a randomized controlled trial on the effect of supplemental intake of vitamin B12 and folic acid on fracture incidence. BMC Geriatr. 2011;11:80.
43. van Wijngaarden JP, Swart KM, Enneman AW, Dhonukshe-Rutten RA, van Dijk SC, Ham AC, et al. Effect of daily vitamin B-12 and folic acid supplementation on fracture incidence in elderly individuals with an elevated plasma homocysteine concentration: B-PROOF, a randomized controlled trial. Am J Clin Nutr. 2014;100(6):1578-86.
44. Willett WC, Sampson L, Browne ML, Stampfer MJ, Rosner B, Hennekens CH, et al. The use of a self-administered questionnaire to assess diet four years in the past. Am J Epidemiol. 1988;127(1):188-99.
45. Rimm EB, Giovannucci EL, Stampfer MJ, Colditz GA, Litin LB, Willett WC. Reproducibility and validity of an expanded self-administered semiquantitative food frequency questionnaire among male health professionals. Am J Epidemiol. 1992;135(10):1114-26; discussion 27-36.
46. Hu FB, Leitzmann MF, Stampfer MJ, Colditz GA, Willett WC, Rimm EB. Physical activity and television watching in relation to risk for type 2 diabetes mellitus in men. Arch Intern Med. 2001;161(12):1542-8.
47. Manson JE, Rimm EB, Stampfer MJ, Colditz GA, Willett WC, Krolewski AS, et al. Physical activity and incidence of non-insulin-dependent diabetes mellitus in women. Lancet. 1991;338(8770):774-8.

Chapter 2

Nutrition & DNA methylation

2.1

Nutrients and DNA methylation across the life course: a systematic review of studies in humans

Braun KVE*, Mandaviya P*, Franco OH, Nano J, Girschik C, Bramer WM, Muka T, Troup J, van Meurs JBJ, Heil SG, Voortman T. Nutrients and DNA methylation across the life course: a systematic review. Submitted for publication.

*Denotes equal contribution

ABSTRACT

Background and objectives: DNA methylation can be modified by environmental factors, including nutrition. In order to gain more insight in effects of nutrients on DNA methylation, we conducted a systematic review on the relation between nutrients and DNA methylation in humans across the life course.

Methods: The literature search was designed by an experienced biomedical information specialist. Six bibliographic databases (Embase.com, Medline (Ovid), Web-of-Science, PubMed, Cochrane Central and Google Scholar) were searched. We selected studies that examined the association between nutrients (blood levels; dietary intake; or dietary supplements) and DNA methylation (global, site specific, or genome-wide) in humans of any age, with no restrictions on year of publication, language, or study design. Abstract screening, full text selection, and data extraction was performed by two independent reviewers, with a third reviewer available to solve any disagreements.

Results: We identified 3774 references, of which 98 studies met all inclusion criteria. The majority was performed in adult study populations, and folate was the main nutrient of interest. Several candidate gene and epigenome-wide association studies reported differential DNA methylation of CpG sites in response to folate (e.g. IGF2, H19, HOX), fatty acids (e.g. PPRAGC1A, TNFα), and vitamin D (CYP24A1). Some of these observed associations were specific to life course stage (e.g. IGF2 in early life) and tissue (e.g. opposite directions for PPRAGC1A in muscle versus fat tissue).

Conclusions: To date, promising results have been reported in the field of nutrition and DNA methylation in humans at different stages across the life-course; especially for nutrients known to be involved in one-carbon metabolism, including folate, but also others, such as fatty acids and vitamin D. Studies on other nutrients, such as other macronutrients and several minerals are still scare. Further large-scale studies of high quality are needed to expand our understanding on the role of nutrition in DNA methylation and its effects on health and disease.

INTRODUCTION

DNA methylation is prone to modification by environmental factors, including nutrition. The susceptibility of change in DNA methylation in response to nutrition is particularly high during early life.[1] Nevertheless, nutrition has also been reported to be associated with DNA methylation in other stages of the life course, for instance, during adolescence and adulthood.[2]

Nutrients that are known to be involved in DNA methylation through their role in one-carbon metabolism are B-vitamins and methionine. However, also other nutrients such as fatty acids, protein, and vitamin D, are suggested to have an effect on DNA methylation.[3] To date, many studies have been carried out investigating the role of nutrition on DNA methylation, in animal studies as well as human studies. Although evidence from animal studies demonstrates that nutrition has effects on DNA methylation, findings from studies in humans are inconsistent.[4]

In order to gain more insight in effects of several nutrients on DNA methylation across the life-course, a clear overview of the current knowledge is of importance. Identifying which nutrients affect DNA methylation, either globally or at specific CpG sites, will provide insight in the mechanisms that are responsible for the effect of nutrition on several health outcomes. Therefore, we conducted a comprehensive systematic review on the relationship between status and intake of nutrients with DNA methylation in humans across the life course.

METHODS

This systematic review was performed and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.[5]

Literature search

A literature search was designed for six electronic databases by an experienced biomedical information specialist. The search engines Embase.com (Medline and Embase), Medline (Ovid), Cochrane Central, Web-of-Science, PubMed, and Google Scholar were searched from inception until May 10th 2016 (date last searched) to identify published studies that examined the association between nutrients and DNA methylation. The full search strategies of all databases are provided in **Supplement 2.1.1.**

Study selection and inclusion criteria

We selected studies that examined the association between nutrients (blood levels; dietary intake; and/or dietary supplements) and DNA methylation (global, site specific, and/or genome-wide) in humans. We excluded studies that were performed among patients with chronic diseases (e.g. Alzheimer's disease, diabetes, anorexia nervosa, cardiovascular diseases, etc.) and case reports (n<5). No restrictions were set on year of publication, language, or study design. We excluded studies on caloric intake, alcohol intake, glucose levels, triglyceride levels, and cholesterol levels as these were outside the scope of our review. Two reviewers screened the retrieved titles and abstracts and selected eligible studies according to predefined selection criteria independently of each other (**Supplement**

2.1.2). Discrepancies between the two reviewers were resolved through discussion, with an arbitrator available if no consensus was reached. We retrieved full texts for studies that satisfied all selection criteria. These full-text articles were evaluated in detail once more by two investigators against the selection criteria.

Data extraction

A structured database was developed prior to the data extraction. Detailed characteristics of individual studies were extracted including study design, study size, country, characteristics of the study population, and details on exposure and outcome assessment. In addition, we extracted information on covariate adjustments, and conclusions. The association of each nutrient with each methylation measure, either global or genome-wide site-specific or gene-specific, was extracted separately to report each specific analysis. Gene-specific studies examining multiple loci were considered per gene separately.

Quality analysis

The quality of included studies was evaluated by two reviewers using a predefined scoring system. This quality score (QS) was previously developed for its use in systematic reviews and meta-analyses including studies with various study designs.[6] A score of 0, 1 or 2 points was allocated to each of the following five items: 1) study design; 2) size of the population for analysis; 3) quality of the methods used for exposure assessment or appropriate blinding of an intervention; 4) quality of the methods used for outcome assessment; and 5) adjustment for potential confounders or adequate randomization of an intervention. The combined scores resulted in a total QS between 0 and 10 points, with 10 representing the highest quality. Details on the QS are presented in **Supplement 2.1.3**.

3 RESULTS

Characteristics of the included studies

From the literature search we identified 3,774 unique references, of which 3,523 were excluded after screening of title and abstract based on the selection criteria. Of the 251 remaining references, full-texts were retrieved and reviewed, of which 98 studies met all criteria and were included in this systematic review (**Figure 2.1.1**). Of the included studies, 25 studies investigated the association between maternal nutrition and offspring DNA methylation, nine studies were carried out during infancy, childhood, or adolescence, and 70 studies were performed in adults (**Table 2.1.1, Figure 2.1.2**). Most studies examined associations between nutrients and gene-specific DNA methylation, of which the majority focused on folate as nutrient of interest. Summaries of the findings of all included studies are presented below per nutrient or nutrient groups and per life stage (**Tables 2.1.2-2.1.18**). Detailed information of population characteristics, DNA methylation assessment, confounder adjustment, and description of results is included in the online supplementary material.

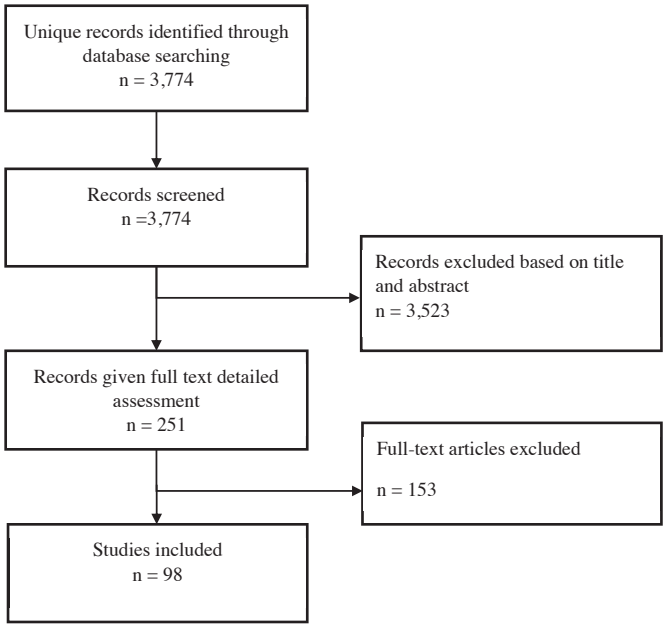


Figure 2.1.1. Flow chart of included studies

Vitamin A, α-carotene, and β-carotene

Vitamin A in infancy, childhood, and adolescence and DNA methylation

Perng et al. (QS: 6) showed that higher plasma vitamin A levels in children aged 5 to 12 years were associated with global DNA hypomethylation in blood.[7]

Vitamin A in adulthood and DNA methylation

Piyathilake et al. (QS: 6) observed no association between plasma vitamin A levels and global DNA methylation in either PBMCs or cervical cells.[8] Bollati et al (QS: 6) investigated the association of intake of α-carotene, β-carotene, and retinol with methylation of CD14, Et-1, HERV-w, iNOS and TNFα in blood. They observed an association of higher intake of β-carotene with hypermethylation of HERV-w and higher intake of β-carotene and retinol was associated with hypomethylation of TNFα. However, no significant association was observed for the other genes.[9] Stidley et al (QS: 6) did not find an association of α-carotene, β-carotene, and retinol intake with methylation index of genes (p16, MGMT, DAPK, RASSF1A, PAX5α, PAX5β, GATA4, and GATA5) in sputum.[10]

Vitamin B1

Vitamin B1 in adulthood and DNA methylation

Marques-Rocha et al. (QS: 5) showed that higher vitamin B1 intake was associated with global DNA hypomethylation in blood.[11]

Table 2.1.1. Number of studies per category of life course stage, nutrient group, and DNA methylation type

	Maternal nutrition & offspring DNA methylation				Infant, child, & adolescent nutrition & DNA methylation				Adult nutrition & DNA methylation				Total
	Total	Global	Gene-specific	Genome-wide	Total	Global	Gene-specific	Genome-wide	Total	Global	Gene-specific	Genome-wide	
Micronutrients													
Vitamin A, α-carotene, and β-carotene	0	0	0	0	1	1	0	0	3	1	2	0	4
Vitamin B1	0	0	0	0	0	0	0	0	1	1	0	0	1
Vitamin B2	1	0	1	0	0	0	0	0	5	5	0	0	6
Vitamin B3	1	0	1	0	0	0	0	0	1	1	0	0	2
Vitamin B6	1	0	1	0	0	0	0	0	11	11	1	0	12
Folate	16	6	7	4	3	3	1	0	43	34	12	1	59
Vitamin B12	4	2	2	0	2	1	1	0	22	18	7	0	26
Vitamin C	0	0	0	0	0	0	0	0	5	1	4	0	5
Vitamin D	3	0	2	1	3	1	2	1	5	3	3	0	10
Vitamin E	0	0	0	0	0	0	0	0	4	1	3	0	4
Choline & betaine	2	2	1	0	0	0	0	0	3	3	1	0	3
Minerals and trace elements	0	0	0	0	1	1	0	0	7	6	2	0	8
Combined nutrients	1	0	0	1	0	0	0	0	6	3	2	1	7
Bioactive compounds	0	0	0	0	0	0	0	0	6	1	5	0	6
Macronutrients													
Fat and fatty acids	5	2	3	1	2	0	0	2	18	4	10	5	25
Carbohydrates & fiber	2	1	1	0	0	0	0	0	6	4	3	0	8
Protein & amino acids	2	1	1	0	0	0	0	0	7	5	3	0	9
Total	25	9	13	6	9	4	4	3	70	42	26	7	98

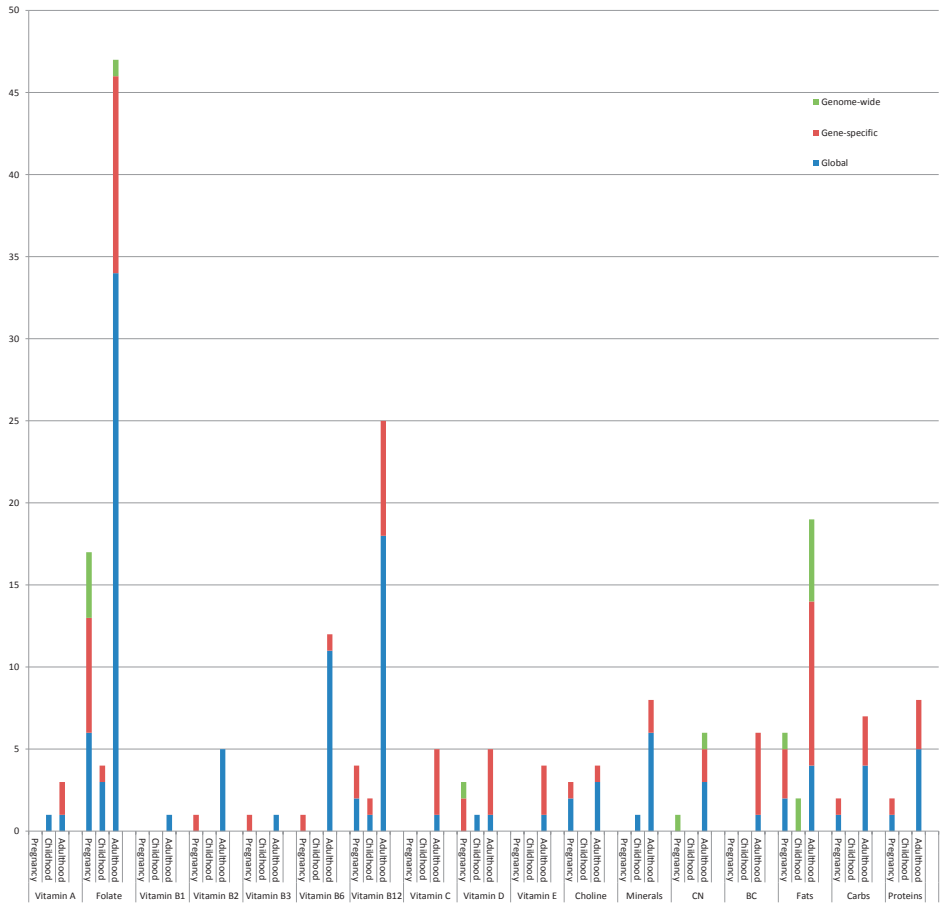


Figure 2.1.2. Distribution of studies (N=98) per category of life course stage, nutrient group, and DNA methylation outcome type. Choline; includes choline and betaine, Minerals; include minerals and trace elements, CN; combined nutrients, BC; bioactive compounds, Carbs; carbohydrates.

Vitamin B2

Maternal vitamin B2 and offspring DNA methylation

Azzi et al. (QS: 6) showed that higher vitamin B2 intake during pregnancy was associated with hypermethylation of ZAC1 in cord blood.[12]

Vitamin B2 in adulthood and DNA methylation

Five studies (QS: 3-5) investigated the association between vitamin B2 intake and global DNA methylation in blood or colonic tissue,[11, 13-16] of which only one study found that higher vitamin B2 intake was associated with DNA hypermethylation in blood.[11] In addition, Figueiredo et al. (QS: 4) found no association between vitamin B2 levels in plasma with global methylation in colonic tissue.[13] Zhang et al. (QS: 5) showed that vitamin B2 intake was not associated with methylation of IL-6 in blood.[16]

Vitamin B3

Maternal vitamin B3 and offspring DNA methylation

Azzi et al. (QS: 6) showed that vitamin B3 intake during pregnancy was not associated with methylation of ZAC1 in cord blood.[12]

Vitamin B3 in adulthood and DNA methylation

Marques-Rocha et al. (QS: 5) showed that higher vitamin B3 intake was associated with global DNA hypomethylation in blood.[11]

Vitamin B6

Maternal vitamin B6 and offspring DNA methylation

Azzi et al. (QS: 6) reported that vitamin B6 intake during pregnancy was not associated with methylation of ZAC1 in cord blood.[12]

Vitamin B6 in adulthood and DNA methylation

Eight studies (QS: 3-5) investigated the association between vitamin B6 intake with global DNA methylation in blood or colonic tissue,[7, 13-19] and only one study found an association and reported that higher vitamin B6 intake was associated with global hypomethylation in blood.[17] Two other studies (QS: 4 and 6) investigated the association of vitamin B6 levels in plasma and venous blood with global DNA methylation in blood or colonic tissue, but did not find any association.[13, 20] An intervention study by Hübner et al. (QS: 3), showed that vitamin B6 supplementation had no effect on global DNA methylation in blood.[21] Zhang et al. (QS: 5) showed that vitamin B6 intake was not associated with methylation of IL-6 in blood.[16]

Folate

Maternal folate and offspring DNA methylation

Sixteen unique studies investigated the association between maternal folate and offspring DNA methylation, six at a global level, seven gene-specific, and four at a genome-wide level. Of the four studies investigating global DNA methylation in cord blood (QS: 2-6), one found that an increased dietary folate as well as RBC folate were associated with global hypomethylation,[22] whereas the other three studies reported no significant association with folate.[23-25] In contrast, studies investigating DNA methylation in other fetal tissues, including placenta, brain and heart, showed that maternal folate levels were associated with global hypermethylation.[26, 27] Four studies investigated the association between levels or intake of maternal folate and *IGF2* methylation in offspring (QS: 5-7). Three of these studies measured methylation in cord blood. In these studies, higher folate intake was shown to be associated with *IGF2* hypermethylation,[22] and no association was observed between serum folate and DNA methylation.[28] RBC folate showed association with *IGF2* hypermethylation in one study,[29] and no association in another study.[22] One observational study found *IGF2* hypermethylation in offspring at 12-18 months in mothers who used folate supplements during pregnancy compared to those of

mothers who did not.[30] Furthermore, Hoyo et al. (QS: 5 and 7) found an association of maternal folate supplements and erythrocytes with *H19* hypomethylation in cord blood.[29, 31] Maternal intake and erythrocyte folate were found to be associated with PEG3 hypomethylation in cord blood,[22, 29] whereas no association was found with maternal RBC folate (QS: 5-7).[22] Furthermore, maternal erythrocyte folate was also associated with hypermethylation of *PEG1/MEST* and hypomethylation of *PEG10/SGCE* in cord blood (QS: 7).[29] Maternal erythrocyte folate was associated with hypomethylation of *MEG3*, *MEG3-IG* (intergenic), *NNAT* and *PLAGL1* promoters in cord blood (QS:7).[29] Van Mil et al. showed that supplement use and venous blood levels of folate were associated with *NR3C1* hypermethylation, and higher folate supplement use was associated with 5-HTT hypomethylation in cord blood (QS: 6 and 7).[32] Furthermore, no associations were observed between maternal folate intake or RBC and *SNRPN* or *ZAC1* methylation in cord blood (QS: 5 and 6).[12, 22]

Of the four studies investigating maternal folate levels and epigenome-wide DNA methylation in cord blood, three used Illumina 450K arrays and identified several CpG sites. Amarasekera et al. (QS: 4, N=23) found seven differentially methylated regions, of which ZFP57 was validated using Sequenom EpiTyper platform.[33] Joubert et al. (QS: 8, N=1988) identified 443 CpG sites annotated to 320 genes, of which some novel genes included *APC2*, *GRM8*, *SLC16A12*, *OPCML*, *PRPH*, *LHX1*, *KLK4* and *PRSS21*.[34] Gonseth et al. (QS: 6, N=347) found that maternal folate intake was associated with three CpG sites annotated to genes *TFAP2A*, *STX11* and *CYS1*.[35] The fourth EWAS (QS: 6, N=200) did not find associations between maternal folate levels and genome-wide DNA methylation.[36]

Folate in infancy, childhood, and adolescence and DNA methylation

Three unique studies investigated the association between folate levels and global DNA methylation during infancy or childhood, of which one also examined gene-specific DNA methylation. Two studies (QS:4 and 5) examined the association between folate levels and global DNA methylation in cord blood: Haggerty et al. showed that RBC folate was associated with global hypomethylation whereas Fryer et al. observed no association.[22, 24] Perng et al. (QS: 4) studied the association between erythrocyte folate and global methylation in children aged 5-12 years, but found no association.[7] Haggerty et al. (QS: 6) also examined the association between RBC folate and DNA methylation of *IGF2*, *PEG3* and *SNRPN* in cord blood. In line with their findings for maternal folate and offspring DNA methylation, they found that RBC folate at birth was also associated with *IGF2* hypermethylation, *PEG3* hypomethylation and no difference in *SNRPN* methylation.[22]

Folate in adulthood and DNA methylation

There were 41 unique studies that investigated the association between folate and adult DNA methylation. Thirty-three studies examined DNA methylation at a global level, 12 gene-specific, and one at a genome-wide level. In observational studies that measured DNA methylation in blood, folate intake was associated with DNA hypomethylation in one study (QS: 5),[14] DNA hypermethylation in three studies (QS: 3-5),[16, 18, 37] but no difference in methylation was observed in majority of the studies (QS: 1-6).[15, 17, 19, 23, 38, 39] Blood levels of folate were associated with DNA hypomethylation in one study (QS: 6),[40] whereas DNA hypermethylation in six studies (QS: 4-6).[20, 41-45] A few studies showed no difference in methylation (QS: 3-6).[19, 46-48] Overall, dietary intake and levels of folate in these observational studies tend to be more often associated with global DNA hypermethylation than hypomethylation.

In contrast, intervention studies with folate supplements showed an effect of folate on DNA hypomethylation in three studies (QS: 2-8),[49-51] DNA hypermethylation in two studies (QS: 3),[52, 53] and no effect on methylation in three studies (QS: 3-8).[21, 54, 55] The direction of effect in these intervention studies were contradictory to those shown in the observational studies. In studies that measured DNA methylation in colonic tissue, folate intake was associated with global DNA hypomethylation on one side of the colon (QS: 3),[13] whereas folate in blood levels[13, 56-58] and colonic tissue[56] showed no association with DNA methylation (QS: 4-6). Furthermore, Llanos et al. measured DNA methylation in breast tissue and observed no association with folate levels from plasma and breast tissue (QS: 6).[59] In an intervention study, Aarabi et al. found a significant association between folate supplements and DNA hypomethylation (QS: 4).[60]

Two studies investigated the association between levels or intake of folate and methylation in *IGF2* and *H19*. In contrast to the studies that examined the same genes in maternal-offspring,[22, 29-31] Hanks et al. (QS: 7) did not find an association between folate in serum, RBC and colonic tissue with *IGF2* methylation in adults.[56] In addition, in an intervention study by Aarabi et al. (QS: 5), no effect was observed between folate supplements and methylation in *H19*.[60]

Higher folate intake was associated with promotor hypomethylation of TNFα and methylation index of genes (p16, MGMT, DAPK, RASSF1A, PAX5α, PAX5β, GATA4, and GATA5) in blood and sputum, respectively. However, no association was observed with promoter TLR2, CD14, Et-1, HERV-w, iNOS and IL-6 methylation in blood (QS: 5-6).[9, 16, 61] Dhillon et al. (QS: 5) showed no significant association between folate intake and GSTM1 methylation in serum. However, when analysis was stratified for MTHFR genotype, low folate intake was associated with lower methylation of GSTM1 methylation for CT and TT group only.[62] Two studies investigated the association of levels of folate in serum, RBC, plasma, and colonic tissue with gene-specific methylation in colonic tissue or rectal mucosa (QS: 6-7). Hypermethylation of MYOD, SFRP1, SFRP2 and methylation index of genes (HPP1, APC, SFRP1, SFRP2, SOX17, WIF1, ESR1, MYOD, N33) was observed with higher folate levels. However, no significant associations were observed for methylation of individual genes: APC, MYOD1, MLH1, N33, SOX17 and/or ESR1.[56, 58] Ottini et al. (QS: 5) reported that higher folate levels in plasma were associated with hypomethylation of methylation index of genes (p16, FHIT, RAR, CDH1, DAPK1, hTERT, RASSF1A, MGMT, BRCA1 and PALB2).[63]

Some studies observed associations for nutrient levels in a tissue specific manner. Tapp et al. (QS: 6) investigated the association of folate levels with HPP1 and WIF1 methylation in rectal mucosa. They observed hypermethylation of HPP1 and WIF1 in association with plasma folate levels, but not with RBC folate levels.[58] Hanks et al. (QS: 7) examined the association between folate levels and MGMT methylation in colonic tissue. They observed hypomethylation of MGMT in association with serum folate levels, but not with folate levels in RBC and colonic tissue.[56] In women with HPV, plasma folate levels were associated with HPV 16 hypermethylation (promoter) in blood (QS: 5).[64] In healthy women, plasma and breast folate levels were associated with hypomethylation of $p16^{INK4a}$ (QS: 7).[65] In intervention studies, folate supplementation had no effect on ESR1 and MLH1 methylation in colonic mucosa, and DLK1/GTL2, MEST, SNRPN, PLAGL1 and KCNQ1OT1 methylation in sperm (QS: 5 and 7).[60, 66] Song et al. (QS: 4) performed an EWAS of folate levels in breast tissue in women, and found two differentially methylated CpG sites. One CpG site near JAG2 was hypomethylated and another CpG site near DNAJC2 was hypermethylated.[67]

Vitamin B12

Maternal vitamin B12 and offspring DNA methylation

Two studies investigated the association of intake and serum levels of vitamin B12 with global DNA methylation in cord blood (QS: 5 and 6). Higher serum levels of vitamin B12 were associated with hypermethylation,[25] whereas no difference in methylation was observed with higher vitamin B12 intake.[23] For the gene-specific studies, Ba et al. (QS: 6) showed that higher vitamin B12 levels in serum were associated with IGF2 promoter hypomethylation in cord blood.[28] Azzi et al. (QS: 6) showed that higher vitamin B12 intake during pregnancy was associated with hypermethylation of ZAC1 in cord blood.[12]

Vitamin B12 in infancy, childhood, and adolescence and DNA methylation

In children aged 5 to 12 years, no association was observed between plasma vitamin B12 levels and global DNA methylation in blood (QS: 4).[7] In cord blood, McKay et al. (QS: 7) showed that higher serum vitamin B12 levels were associated with hypomethylation of IGFBP3.[25]

Vitamin B12 in adulthood and DNA methylation

Eight studies (QS: 3-6) found no association between vitamin B12 intake and global DNA methylation.[14-19, 23, 38] Also for vitamin B12 levels in plasma or serum, no associations were found with global DNA methylation in blood, colonic tissue or rectal mucosa in eight studies (QS: 3-6).[13, 19, 20, 46-48, 57, 58] However, in one study (QS: 6), plasma vitamin B12 levels were associated with global hypermethylation in blot clots.[68] An intervention study (QS: 3) showed that vitamin B12 supplements had no effect on global DNA methylation in blood.[69] For gene-specific studies, vitamin B12 intake was not associated with methylation of CD14, Et-1, HERV-w, iNOS, TNFα, IL-6, and TLR2 in blood (QS: 5-6).[9, 16, 61] In addition, Tapp et al (QS: 6) found no association between vitamin B12 levels in plasma and methylation of HPP1, APC, SFRP1, SFRP2, SOX17, WIF1, ESR1, MYOD, N33 and methylation index of these genes.[58] Al-Ghnaniem et al. (QS: 6) observed that vitamin B12 levels in venous blood were associated with hypomethylation of ERα in colonic mucosa.[70] Furthermore, plasma vitamin B12 levels were associated with hypermethylation of HPV 16 (QS: 5) and no difference in methylation index of genes (p16, FHIT, RAR, CDH1, DAPK1, hTERT, RASSF1A, MGMT, BRCA1 and PALB2) (QS: 3).[63, 64]

Vitamin C

Vitamin C in adulthood and DNA methylation

Piyathilake et al. (QS: 6) investigated the association between plasma vitamin C levels and global DNA methylation in PBMCs and cervical cells, but did not find a significant association.[8] Two studies (QS: 6) investigated the association between vitamin C intake and gene-specific promoter DNA methylation. One study found that higher intake of vitamin C was associated with hypomethylation of PON1 in venous blood,[71] whereas the other study did not find an association of vitamin C intake with methylation index of genes (p16, MGMT, DAPK, RASSF1A, PAX5α, PAX5β, GATA4, and GATA5) in sputum.[10] Furthermore, Bollati et al (QS: 6) investigated the association of intake of ascorbic acid

with methylation of CD14, Et-1, HERV-w, iNOS and TNFα in blood, but no significant association was observed.[9] Piyathilake et al. (QS: 5) found no association between vitamin C levels in plasma and HPV 16 methylation in blood.[64]

Vitamin D

Vitamin D and offspring DNA methylation

One study (QS: 5) found that higher maternal vitamin D levels in plasma was associated with hypomethylation of RXRA in umbilical cord tissue,[72] whereas another study (QS: 4) did not find an association between maternal vitamin D levels in serum and CYP24A1 methylation in placenta.[73] Mozhui et al (QS: 6) conducted an EWAS of vitamin D levels, but found no significant CpG site associations in cord blood.[36]

Vitamin D in infancy, childhood, and adolescence and DNA methylation

For vitamin D, Zhu et al. (QS: 7) found that plasma vitamin D levels as well as an intervention with vitamin D supplementation were associated with global hypermethylation in blood.[74] In neonates, Novakovic et al (QS: 4) found no association between serum vitamin D levels and CYP24A1 methylation in placenta.[73] Zhu et al. (QS: 4) investigated the associated between plasma vitamin D levels and gene-specific methylation in blood among children around 16 years of age. Higher vitamin D levels were found to be associated with hypomethylation of CYP2R1, hypermethylation of CYP24A1, and differential methylation in opposite directions of two CpG sites annotated to DHCR7.[75] In an EWAS of vitamin D levels, Zhu et al. (QS: 5) identified two differentially methylated CpG sites: higher vitamin D levels were associated with hypomethylation of DOI3 and hypermethylation of MAPRE2.[75]

Vitamin D in adulthood and DNA methylation

Three studies, of which two were cross-sectional (QS: 5) and one intervention (QS: 3) found no effect of vitamin D supplementation or vitamin D levels in plasma or serum on global DNA methylation in blood or rectal mucosa.[21, 58, 76] In gene-specific studies, Bollati et al (QS: 6) found no association between vitamin D intake and promoter methylation of CD14, Et-1, HERV-w, iNOS and TNFα.[9] Tapp et al. (QS: 6) investigated the association between plasma vitamin D levels and gene-specific methylation of HPP1, APC, SFRP1, SFRP2, SOX17, WIF1, ESR1, MYOD and N33 in rectal mucosa. They found that higher vitamin D levels were associated with hypomethylation of index of these genes and also individual genes such as APC, WIF1 and MYOD. No association was found for the other individual genes.[58] Furthermore, Ashktorab et al. (QS: 7) found no significant association between serum vitamin D levels and DKK1 methylation in blood.[77]

Vitamin E

Vitamin E in adulthood and DNA methylation

Piyathilake et al. (QS: 6) observed that higher levels of plasma vitamin E were associated with global hypomethylation in PBMCs, but not with global methylation in cervical cells.[8] In a gene-specific study, Stidley et al (QS: 6) did not find an association of vitamin E intake with methylation index of

any of the studied genes (p16, MGMT, DAPK, RASSF1A, PAX5α, PAX5β, GATA4, and GATA5) in sputum.[10]

Choline and Betaine

Maternal choline and betaine and offspring DNA methylation

Boeke et al. (QS: 6) showed that choline and betaine intake during pregnancy were associated with global DNA hypomethylation in cord blood.[23] In contrast, Jiang et al. (QS: 6), in an intervention study, reported that higher supplemental choline intake during pregnancy was associated with global DNA hypermethylation in placenta, but not in cord blood.[78] For gene-specific studies, an intervention study by Jiang et al. (QS: 5) showed that higher supplemental choline intake during pregnancy was associated with hypermethylation of promoter CRH and NR3C1 in placenta. In contrast to placental tissue, in cord blood promoter CRH and NR3C1 were hypomethylation in response to a higher supplemental choline intake. No effect of choline supplementation was observed in methylation of GNAS-AS1, IGF2, IL-10, or LEP in placenta or cord blood.[78]

Choline and betaine in adulthood and DNA methylation

Shin et al. (QS: 3) showed that choline supplementation was associated with global DNA hypermethylation in blood, but only for subjects with the MTHFR 677CC genotype.[79] However, another intervention study (QS: 4) showed no effect of choline supplementation on global DNA methylation in blood.[78] In addition, in a cross-sectional study by Boeke et al. (QS: 6) choline and betaine intake was not associated with global DNA methylation in blood in pregnant women.[23] In an intervention study by Jiang et al. (QS: 5) a higher supplemental choline intake was not associated with methylation of promoter CRH and NR3C1 in blood among pregnant women.[78]

Minerals and trace elements

Minerals and trace elements in infancy, childhood, and adolescence and DNA methylation

In a cross-sectional study among children aged 5 to 12 years, Perng et al (QS: 4) investigated the association of plasma levels of ferritin and serum levels of zinc with global DNA methylation in blood, but no significant associations were observed.[7]

Minerals and trace elements in adulthood and DNA methylation

Marques-Rocha et al. (QS: 5) showed that higher intakes of both copper and iron were associated with global DNA hypermethylation in blood.[11] In contrast, Gomes et al (QS: 3) reported that higher intakes of magnesium were associated with DNA hypomethylation in blood.[17] Two studies (QS: 3 and 5) examined associations between zinc intake and global DNA methylation in blood, but both observed no significant associations.[17, 38] Furthermore, Tapp et al. (QS: 5) reported that higher plasma levels of selenium were associated with global DNA hypermethylation in rectal mucosa, but only in women.[58] McChelland et al. (QS: 6) reported that higher serum levels of phosphate were associated

with global DNA hypomethylation in blood.[80] In an intervention study by Hubner et al. (QS: 3) no effect was observed of calcium supplementation on global DNA methylation in blood.[21]

For gene-specific studies, Shimazu et al. (QS: 6) investigated the associations of salt intake with methylation of miR-124a-3, EMX1 and NKX6-1 in gastric mucosa, but observed no significant associations.[81] Tapp et al. (QS: 6) examined associations of plasma selenium levels with methylation of HPP1, APC, SFRP1, SFRP2, SOX17, WIF1, ESR1, MYOD, N33 and an index of these genes. They observed that higher levels of selenium were associated with hypermethylation of N33, but only in males. No significant associations were observed for the other genes or index.[58]

Combined nutrients

Maternal combined vitamins and minerals and offspring DNA methylation

In an intervention study, Khulan et al. examined the effect of maternal micronutrient supplementation on genome-wide DNA methylation in cord blood and peripheral blood during infancy (QS:6).[82] In cord blood 14 CpG sites were identified in males and 21 CpG sites in females, with no overlap between the two. In addition, in infant blood 108 CpG sites were identified in males and 106 CpG sites in females, with only 5 (XIRP1, MEOX1, GNA11, C1orf54, KRTAP21-1) in agreement between male and female infants. Within each sex a significant proportion of the identified CpG sites in cord bloods were also observed in infants (males 7/14: MEOX1, SPAG4L, SPATA22, NRN1L, C1orf54, PIP, PTPN20B; females 8/21: HSPC176, XIRP1, BPIL1, MGMT, CHIT1, C14orf152, KLRC2, DEFB123).

Combined vitamins and minerals in adulthood and DNA methylation

Two studies investigated the association between folate and vitamin B12 combined with global DNA methylation. Piyathilake et al. (QS: 6) reported that vitamin B12 and folate plasma was associated with global hypermethylation in PBMCs, but not in cervical cells.[8] However, in an intervention study by Fenech et al. (QS: 6) vitamin B12 and folate supplements had no effect on global DNA methylation.[46] Pusceddu et al. (QS:5) studied the effect of vitamin B12, vitamin B6, folate, vitamin D and calcium supplementation on global DNA methylation After intervention LINE-1 hypermethylation was observed.[83] For gene-specific studies, Stidley et al. (QS: 6) observed that intake of multivitamins was associated with hypomethylation of methylation index (p16, MGMT, DAPK, RASSF1A, PAX5α, PAX5β, GATA4, and GATA5).[10] In an intervention study by Zhou et al. (QS: 5), vitamin D and calcium supplementation led to hypomethylation of CYP24A1, but had no effect on CYP2R1, CYP27A1, and CYP27B1.[84] Kok et al. studied the effect of vitamin B12 and folate supplementations on epigenome-wide DNA methylation (QS: 8). After intervention 162 CpG sites were identified as compared to baseline. Comparisons of folic acid and vitamin B12 versus placebo revealed one CpG sites (cg19380919) and 6 differentially methylated regions (DMRs), with pronounced changes in DIRAS3, ARMC8, and NODAL. In addition, serum levels of folate and vitamin B12 were related to DNA methylation of 173 and 425 regions, respectively, including several members of the developmental HOX genes.[85]

Bioactive compounds

Bioactive compounds in adulthood and DNA methylation

Bollati et al (QS: 6) investigated the association of intake carotenoids, polyphenols, and flavonoids with methylation of CD14, Et-1, HERV-w, iNOS and TNFα in blood. They observed an association of higher intake of carotenoids with hypermethylation of HERV-w and with hypomethylation of TNFα. However, no significant association was observed for the other genes.[9] De la Iglesia et al. (QS: 6) observed that that lycopene intake was associated with PON1 hypomethylation.[71] Zhong et al. (QS: 6) reported that flavonoid intake was associated with TLR2 hypomethylation.[61] Stidley et al (QS: 6) did not find an association of carotene, lutein, zeaxanthin, and lycopene intake with methylation index of genes (p16, MGMT, DAPK, RASSF1A, PAX5α, PAX5β, GATA4, and GATA5) in sputum.[10]

Fat and fatty acids

Maternal fat and fatty acids and offspring DNA methylation

Rerkasem et al. (QS: 3) found no association between total fat intake during pregnancy and global DNA methylation in blood of children around 20 years of age.[86] Godfrey et al. (QS: 3) also found no association between total fat intake during pregnancy and methylation of RXRA and eNOS in cord blood.[87] For individual fatty acids, an intervention study by Lee et al. (QS: 8) found that ω-3 PUFA supplementation during pregnancy had an effect on global hypermethylation in cord blood.[88] In the same study, Lee et al. (QS: 8) investigated gene-specific DNA methylation and found that ω-3 PUFA supplementation during pregnancy had an effect on hypermethylation of IGF2, but no association was found for methylation of IFNγ, TNFα, IL-13, GATA3, STAT3, IL-10 and FOXP3 in cord blood.[88] In an intervention by Amarasekera et al. (QS: 8), no significant effects of ω-3 PUFA were found on genome-wide DNA methylation in cord blood.[89]

Fat and fatty acids in infancy, childhood, and adolescence and DNA methylation

Voisin et al. (QS: 5) conducted an EWAS and identified four, 130, 158 and 16 CpG sites in blood significantly associated with intake of total fat, (MUFA+PUFA)/SFA, MUFA/SFA and PUFA/SFA, respectively, in children around the age of 10 years.[90] On the other hand, in an intervention study by Lind et al. (QS: 8) found no epigenome-wide CpG site associations of ω-3 PUFA supplementation as fish oil.[91]

Fat and fatty acids in adulthood and DNA methylation

Total fat: Four studies (QS: 3-4) investigated the association between total fat intake and global DNA methylation in blood, of which two found that higher fat intake was associated with global DNA hypomethylation,[17, 19] whereas the other two studies did not find an association.[15, 16] Gómez-Uriz et al. (QS: 4) found that total fat intake was associated with TNFα hypomethylation in blood.[92] However, Bollati et al. and Zhang et al (QS: 5 and 6) investigated the association of intake of total fat with methylation of CD14, Et-1, HERV-w, iNOS, TNFα and IL-6 in blood, but no significant association was observed.[9, 16] In addition, Stidley et al. (QS: 6) found no association of total fat as well as animal fat intake with methylation index of genes (p16, MGMT, DAPK, RASSF1A, PAX5α, PAX5β, GATA4,

and GATA5) in sputum.[10] Brøns and Gillberg et al. (QS: 6) investigated the effect of high fat overfeeding in healthy men on promoter PPARGC1A methylation.[93, 94] They found that high fat overfeeding in men with normal birthweight was associated with PPARGC1A hypomethylation in muscle tissue,[93] whereas high fat overfeeding in men with low birthweight was associated with PPARGC1A hypermethylation in subcutaneous adipose tissue.[94] In an EWAS by Irvin et al. (QS: 8) fatty acids were not associated with epigenome-wide DNA methylation after correcting for multiple testing.[69] In an intervention study among healthy men, Jacobsen and Gillberg et al (QS: 6) investigated the effect of high fat overfeeding on epigenome-wide DNA methylation. They observed that high fat overfeeding affected methylation at 652 CpG sites (including in CDK5, IGFBP5 and SLC2A4) in subcutaneous adipose tissue.[95] Furthermore, change in DNA methylation of 7,909 CpG sites (including, DNM2, MGMT, SLC2A3/GLUT3, MRC1 and ACAT2) in skeletal muscle biopsies was observed in response to high fat overfeeding. Within the same study, it was shown that these changes in genome-wide DNA methylation were more pronounced in participants with a low birth weight, compared to those with a normal birthweight.[96]

Individual fatty acids: Zhang et al. (QS: 4) found no association between PUFA intake and global DNA methylation in blood.[15, 16] For methylation measured in blood, one study found that higher ω-6 PUFA intake was associated with hypomethylation of TNFα,[97] and another study found that higher PUFA intake was associated with hypermethylation of CLOCK.[98] However, Zhang et al (QS: 5) and Bollati et al (QS: 6) observed no association of intake of total PUFA and ω-3 PUFA with methylation of CD14, Et-1, HERV-w, iNOS, TNFα and IL-6 in blood.[9, 16] Furthermore, Ma et al. (QS: 7) found that higher ω-3 PUFA levels in erythrocytes were associated with hypomethylation of IL-6 promoter.[99] In another study by Ma et al. (QS: 7), higher levels of erythrocyte EPA were associated with hypomethylation of ABCA1 promoter. In addition, higher erythrocyte ALA was associated with hypomethylation of APOE for the CC genotype, whereas higher erythrocyte ALA was associated with hypermethylation of APOE for the AA genotype.[100] In an EWAS by Aslibekyan et al. (QS:6), higher levels of ω-3 PUFA in erythrocytes were associated with 27 CpG sites (including NAV1, CCL17, ACTA2/FAS and AHRR) in blood. Zhang et al. (QS: 4) found no association between MUFA intake and global DNA methylation in blood.[15, 16] Milagro et al. (QS: 4) found that higher intake of MUFA was associated with hypomethylation of CLOCK in blood.[98] However, Zhang et al. (QS: 5) and Bollati et al (QS: 6) investigated the association of intake of MUFA with methylation of CD14, Et-1, HERV-w, iNOS, TNFα and IL-6 in blood, but no significant association was observed.[9, 16] Zhang et al. (QS: 4) observed that higher SFA intake was associated with global DNA hypomethylation when methylation was measured using pyrosequencing,[15] whereas no association was observed when methylation was measured using PCR technique.[16] Zhang et al. (QS: 5) and Bollati et al (QS: 6) investigated the association of intake of SFA with methylation of CD14, Et-1, HERV-w, iNOS, TNFα and IL-6 in blood, but no significant association was observed.[9, 16]

Carbohydrates and fiber

Maternal carbohydrates and fiber and offspring DNA methylation

Rerkasem et al. (QS: 3) investigated the association between intake of carbohydrates during pregnancy and global DNA methylation in their offspring around the age of 20 years, but observed no significant

association.[86] Godfrey et al. (QS: 3) explored the association between carbohydrate intake during pregnancy and gene-specific promoter DNA methylation in cord tissue. They observed that higher intake of carbohydrate was associated with RXRA hypomethylation, but no significant association was observed with eNOS methylation.[87]

Carbohydrates and fiber in adulthood and DNA methylation

Four studies investigated the association between carbohydrate intake and global DNA methylation in blood. One study (QS:3) reported that higher intake of carbohydrate was associated with global DNA hypomethylation,[17] whereas the other three studies (QS:4) observed no significant associations.[15, 16, 19] Three studies (QS: 4-6) were conducted on the association of carbohydrate and fiber intake and gene-specific DNA methylation in blood, of which one observed that carbohydrate intake was associated with hypermethylation of BMAL1,[101] but in the other two studies no significant associations were observed with CD14, Et-1, HERV-w, iNOS, TNFα, and IL-6.[9, 16]

Protein and amino acids

Maternal protein and amino acids and offspring DNA methylation

Rerkasem et al. (QS: 3) observed no associations between intake of protein during pregnancy and global DNA methylation in their offspring around the age of 20 years.[86] Godfrey et al. (QS: 3) found no association of protein intake during pregnancy with promoter methylation of RXRA or eNOS in cord tissue.[87]

Protein and amino acids in adulthood and DNA methylation

Three studies (QS: 3-4) investigated the association between protein intake and global DNA methylation in blood, but no significant associations were observed.[15-17] In addition, the four studies (QS: 3-4) that explored associations between methionine intake and global DNA methylation in blood showed no significant associations.[15, 16, 18, 38] Two studies (QS: 5 and 6) investigated associations of protein intake and methylation of CD14, Et-1, HERV-w, iNOS, TNFα, and IL-6 in blood, but reported no significant associations.[9, 16] Similarly, two studies (QS: 5 and 6) that examined the associations of methionine intake with IL-6 and TLR2 promoter methylation in blood showed no significant associations.[16, 61]

DISCUSSION

This is the first systematic review on nutrients and DNA methylation in humans. This review included studies examining associations between all nutrients with different types of methylation outcomes in all life-course stages. The majority of the studies performed on this topic investigated the association of folate, other B-vitamins, or fatty acids with DNA methylation, whereas studies for other nutrients are scarce or even absent, such as for vitamin K. The evidence so far shows that folate and fatty acids are associated with DNA methylation on a global, gene-specific, and genome-wide level. Several candidate gene and epigenome-wide association studies reported differential DNA methylation of CpG

sites in response to folate (e.g. IGF2, H19, HOX), fatty acids (e.g. PPRAGC1A, TNFα), and vitamin D (CYP24A1). Some of these observed associations were specific to life course stage (e.g. IGF2 in early life) and tissue (e.g. opposite directions for PPRAGC1A in muscle vs. fat tissue). However, some of the directions of these associations are inconsistent across the different studies. Reasons that may explain these discrepancies in findings include the differences in the type of tissue used to measure nutrient levels as well as DNA methylation and differences across life course stages.

DNA methylation across the life course

Age affects methylation patterns via transcriptional regulation.[102] Evidence from our systematic review suggests age-specific differences in folate associations with global DNA methylation in blood. In early life, higher folate intakes and levels were associated with global DNA hypomethylation,[22] whereas during adulthood higher folate intakes and levels were associated with DNA hypermethylation.[16, 18, 20, 37, 41-45] This pattern could not be observed for the other nutrients due to the limited number of studies across the different life course stages or null associations, but similar patterns for other nutrients could be possible. Indeed, some studies on gene-specific DNA methylation observed similar age-specific differences. Methylation of some genes in blood may be specifically prone to environmental influences during early life. For example, evidence from our systematic literature review shows that during pregnancy, at birth and during childhood, higher folate intake and levels were associated with IGF2 hypermethylation.[22, 29, 30] However, when this association was studied in adults, no significant associations were observed.[56] Possibly, methylation of this gene is no longer susceptible to changes during adulthood. Similarly, higher choline intake during pregnancy was associated with CRH and NR3C1 hypomethylation in cord blood, but no association was found within pregnant women.[78] Several studies included in our review suggested that IGF2 methylation could also be affected by vitamin B12 and PUFAs.[28, 103] Higher folate during pregnancy and at birth was associated with PEG3 hypomethylation in cord blood.[22, 29] Due to the absence of studies in adults, we cannot determine age-related associations in relation to these nutrients. Future studies should be conducted in order to establish the life course patterns of these nutrients on DNA methylation.

DNA methylation across nutrients

We were interested in exploring how different nutrients may affect methylation of the same genes, because this can be indicative of which genes are susceptible to nutritional influences in general. As it is difficult to compare findings across studies due to the different population structure, lab techniques, batch effects and analytical methods used, we compared the associations of different nutrients with methylation of the same genes within the same study to discard the bias caused by heterogeneity. From these studies that explored associations between multiple individual nutrients and gene-specific methylation for the same genes and same tissues, we observed that genes such as CD14, eNOS, ESR1, Et-1, IL-6, iNOS, N33 and SOX17 were not susceptible to any nutritional influence.[9, 16, 58, 87] Methylation of some other genes, such as APC, HERV-w, HPP1, RXRA, TLR2, SFRP1, SFRP2 and methylation index of genes (p16, FHIT, RAR, CDH1, DAPK1, hTERT, RASSF1A, MGMT, BRCA1 and PALB2) showed susceptibility to only one nutrient.[9, 58, 61, 87] This suggests that methylation of these genes could be responsive to one particular nutrient due to a certain pathway in which they are involved. However, for certain other genes, such as CLOCK, HPV 16, MYOD, PON1,

TNFα, WIF1, ZAC1, and methylation indexes of genes (HPP1, APC, SFRP1, SFRP2, SOX17, WIF1, ESR1, MYOD and N33; p16, MGMT, DAPK, RASSF1A, PAX5α, PAX5β, GATA4 and GATA5), methylation was susceptible to multiple nutrients.[9, 10, 12, 58, 64, 71, 98] Of these genes, HPV 16 and ZAC1 were hypermethylated in association with folate, vitamin B2 or vitamin B12, but not other nutrients,[12, 64] suggesting a possible role in one-carbon metabolism. Overall, these findings suggest that methylation of some genes are susceptible to nutritional influences in general, whereas methylation of other genes are only responsive to specific nutrients.

DNA methylation across nutrient tissues

The tissue used to measure nutrient exposure might contribute to the nutrient-methylation association. For certain types of tissues it may be easier to detect associations of nutrients with DNA methylation, due to a better reflection of long-term status of these nutrients.[104] Within studies that investigated the association between folate from different tissues or intake and same gene-specific methylation using same methylation tissue, we observed that some genes (H19 and $p16^{INK4a}$,[29, 31, 65] MYOD, NR3C1_a, SFRP1, SFRP2 and methylation index (HPP1, APC, SFRP1, SFRP2, SOX17, WIF1, ESR1, MYOD, N33)) showed differential methylation in association with folate regardless of the source: for plasma, RBC and dietary folate.[32, 58] However, methylation of HPP1, MGMT, PEG3, and WIF1 consistently showed no association with RBC folate, but showed differential methylation in association with plasma, serum or dietary folate within the same studies.[22, 56, 58] Even though RBC folate has shown associations that are in line with other folate tissues in most studies, these four studies have shown discrepancies in their findings, which might be explained by the difference in reflection of long-term folate status between RBC and other tissue of folate.[105]

DNA methylation across methylation tissues

In addition to levels from different tissues or methods used to measure intake of nutrients, tissue-specificity in DNA methylation is an important factor to take into account when interpreting findings of the studies included in this review. A few studies explored associations of specific nutrients with gene-specific DNA methylation measured in different tissues within the same subjects. For example, an intervention study by Jiang et al. showed opposite directions of association between choline intake during pregnancy and CRH or NR3C1 methylation in placenta versus cord blood, which shows that associations may differ between tissues.[78] Another intervention study showed tissue-specific PPARGC1A methylation in muscle biopsy and subcutaneous adipose tissue in opposite directions in association with high-fat overfeeding diet.[93, 94]

DNA methylation of replicable genes

Replication of results across different study populations is of utmost importance in order to confirm findings. In this systematic review, we identified different studies examining same nutrient and gene-specific methylation associations within the same stage of the life course. Different measures of folate during pregnancy were associated with hypomethylation of PEG3 and hypermethylation of IGF2 in offspring, and these findings were replicated across different studies.[22, 29, 30] Methylation of TNF-α was not replicable across studies, possibly due to differences in population characteristics. One study

explored this association in normal-weight young women, whereas the other study explored this association in overweight older adults, which may explain the non-replicability.[9, 97]

Genome-wide DNA methylation

In this systematic review, we identified five EWASs of folate, of which four studies were conducted in relation to maternal nutrition and offspring DNA methylation,[33-36] and one study in relation to nutrition in adulthood and DNA methylation.[67] We investigated whether any genes of which methylation showed significant associations in gene-specific studies in relation to folate, were also found to be genome-wide significant in these EWASs. In line with the gene-specific study of Hoyo et al., who found an association of maternal folate supplements and erythrocytes with *H19* hypomethylation in cord blood,[29, 31] Joubert et al. in their EWAS of maternal plasma folate also found genome-wide significant hypomethylation of H19.[34] Furthermore, this EWAS found hypomethylation of SNRPN which showed null association in relation to folate intake in a gene-specific study.[22] Two EWASs of vitamin D levels were conducted by Mozhui et al. [36] and Zhu et al.,[75] but the former found null association and the later showed no comparison in genes with previous gene-specific studies.

Methodological considerations

In this systematic review, a comprehensive overview is provided on the current evidence on associations between nutrients and DNA methylation in humans of all ages and during all life stages. We included studies examining nutrient intake from diet and supplements, as well as nutrient levels measured in blood or other tissues. Furthermore, we included studies using different tissues to measure DNA methylation, which enabled us to provide more insight on tissue-specific DNA methylation. The literature was designed by an experienced biomedical information specialist and multiple search engines were used, limiting the probability of undetected references. Furthermore, the selection process was performed by two independent reviewers. In order to objectively measure study quality, we used a quality score based on a predefined scoring system developed for systematic reviews. Study quality was determined by five study characteristics: study design, sample size, quality of exposure assessment, quality of outcome assessment, and confounder adjustment.

Overall, most studies are of medium to high quality. A majority of studies (28%) have a quality score 6 and 77% of the studies were of the quality score 5 or above. However, many studies did not properly adjust for confounding. Even though age, sex, and lifestyle covariates were often taken into account, other factors important in DNA methylation studies, such as leukocyte proportion, were rarely included in the models. Lack of adjustment for these covariates may have led to spurious findings. As the majority of the included studies were of observational nature, causal inference from these studies is not possible. Only a few studies tested for effect modification by sex of genetic variation. This should be considered in future studies, as some studies show that associations may be different for males and females or different genotypes, such as MTHFR. There were a few EWASs included in our review, enabling the possibility to identify novel CpG sites in relation to certain nutrients. Nevertheless, with this type of approach caution must be taken when interpreting the results. As EWAS follows a hypothesis-free approach, significant findings could be the result of false positives caused by multiple testing. To solve this issue, several statistical methods could be used, such as Bonferroni correction or false detection rate (FDR). Furthermore, in these type of studies that use a hypothesis-

free approach, there is a need for replication. Therefore, to ensure that findings are true-positives, a discovery and a replication cohort should be used. Even though almost all EWAS included in this review used a proper method for multiple testing, not all studies included a replication cohort in their study. Therefore, findings from these studies still need to be replicated in other cohorts in order to confirm that novel CpG sites identified in EWAS are true positives. On the other hand, the EWAS approach is also prone to type II error, or false-negatives. Usually, this is caused by lack of power due to a small sample size. Out of the included studies that used EWAS, only five had a sample size of >100 participants. Possibly, potential associations between nutrients and genome-wide DNA methylation may have been overlooked in studies with a small sample size.

Conclusions

To date, promising results have been reported in the field of nutrition and DNA methylation in humans at different stages across the life-course. In particular, nutrients known to be involved in one-carbon metabolism, such as folate, have been shown to be related to DNA methylation at a global, gene-specific and genome-wide level. In addition, also other nutrients, such as fatty acids, are suggested to be involved in DNA methylation. However, further large-scale studies of high quality are needed to expand our understanding on the role of nutrition in DNA methylation and its effects on health and disease.

Table 2.1.2. Study characteristics and results: Vitamin A, α-carotene, and β-carotene.

Nutritional exposure	First author, year	Study design	N	Country	Exposure assesment	Genes	Tissue	Direction results	QS
CHILDHOOD									
GLOBAL									
Vitamin A	Perng, 2012	Cross-sectional	568	Colombia	Plasma	NA	Blood	Hypo	6
ADULTHOOD									
GLOBAL									
Vitamin A	Piyathilake, 2011	Cross-sectional	376	USA	Plasma	NA	PBMC	No	6
Vitamin A	Piyathilake, 2011	Cross-sectional	376	USA	Plasma	NA	Cervical cells	No	6
GENE-SPECIFIC									
Vitamin A (Retinol)	Bollati, 2014	Cross-sectional	165	Italy	FFQ	CD14	Blood	No	6
Vitamin A (Retinol)	Bollati, 2014	Cross-sectional	165	Italy	FFQ	Et-1	Blood	No	6
Vitamin A (Retinol)	Bollati, 2014	Cross-sectional	165	Italy	FFQ	HERV-w	Blood	No	6
Vitamin A (Retinol)	Bollati, 2014	Cross-sectional	165	Italy	FFQ	iNOS	Blood	No	6
Vitamin A (Retinol)	Bollati, 2014	Cross-sectional	165	Italy	FFQ	TNFα	Blood	Hypo	6
Vitamin A (Retinol)	Stidley, 2010	Cross-sectional	1101	USA	FFQ	Index (p16, MGMT, DAPK, RASSF1A, PAX5α, PAX5β, GATA4, and GATA5)	Sputum	No	6
Carotenoids (α-carotene)	Bollati, 2014	Cross-sectional	165	Italy	FFQ	CD14	Blood	No	6
Carotenoids (α-carotene)	Bollati, 2014	Cross-sectional	165	Italy	FFQ	Et-1	Blood	No	6
Carotenoids (α-carotene)	Bollati, 2014	Cross-sectional	165	Italy	FFQ	HERV-w	Blood	No	6
Carotenoids (α-carotene)	Bollati, 2014	Cross-sectional	165	Italy	FFQ	iNOS	Blood	No	6
Carotenoids (α-carotene)	Bollati, 2014	Cross-sectional	165	Italy	FFQ	TNFα	Blood	No	6
Carotenoids (α-carotene)	Stidley, 2010	Cross-sectional	1101	USA	FFQ	Index (p16, MGMT, DAPK, RASSF1A, PAX5α, PAX5β, GATA4, and GATA5)	Sputum	No	6

Table 2.1.2. Study characteristics and results: Vitamin A, α-carotene, and β-carotene. (continued)

Nutritional exposure	First author, year	Study design	N	Country	Exposure assesment	Genes	Tissue	Direction results	QS
Carotenoids (β-carotene)	Bollati, 2014	Cross-sectional	165	Italy	FFQ	CD14	Blood	No	6
Carotenoids (β-carotene)	Bollati, 2014	Cross-sectional	165	Italy	FFQ	Et-1	Blood	No	6
Carotenoids (β-carotene)	Bollati, 2014	Cross-sectional	165	Italy	FFQ	HERV-w	Blood	Hyper	6
Carotenoids (β-carotene)	Bollati, 2014	Cross-sectional	165	Italy	FFQ	iNOS	Blood	No	6
Carotenoids (β-carotene)	Bollati, 2014	Cross-sectional	165	Italy	FFQ	TNFα	Blood	Hypo	6
Carotenoids (β-carotene)	Stidley, 2010	Cross-sectional	1101	USA	FFQ	Index (p16, MGMT, DAPK, RASSF1A, PAX5α, PAX5β, GATA4, and GATA5)	Sputum	No	6

Table 2.1.3. Study characteristics and results: Vitamin B1.

First author, year	Study design	N	Country	Exposure assesment	Genes	Tissue	Direction results	QS
ADULTHOOD								
GLOBAL								
Marques-Rocha, 2016	Cross-sectional	156	Brazil	72h food record	NA	WBCs	Hypo	5

Table 2.1.4. Study characteristics and results: Vitamin B2.

First author, year	Study design	N	Country	Exposure assesment	Genes	Tissue	Direction results	QS
MATERNAL OFFSPRING RELATIONSHIP								
GENE-SPECIFIC								
Azzi, 2014	Cross-sectional	2002	France	FFQ & 24h recall	ZAC1	Cord blood	Hyper	6
ADULTHOOD								
GLOBAL								
Marques-Rocha, 2016	Cross-sectional	156	Brazil	72h food record	NA	WBCs	Hyper	5
Ono, 2012	Cross-sectional	384	Japan	FFQ & 24h recall	NA	Blood	No	5
Zhang 2012	Cross-sectional	165	USA	FFQ	NA	WBCs	No	4
Zhang, 2011	Cross-sectional	149	USA	FFQ	NA	Peripheral blood leukocytes	No	4
Figueiredo, 2009	Cross-sectional	388	North America	FFQ	NA	biopsie from right and left colon	NA (Quartiles)	3
Figueiredo, 2009	Cross-sectional	388	North America	Plasma	NA	biopsie from right and left colon	NA (Quartiles)	4

Table 2.1.5. Study characteristics and results: Vitamin B3.

First author, year	Study design	N	Country	Exposure assesment	Genes	Tissue	Direction results	QS
MATERNAL OFFSPRING RELATIONSHIP								
GENE-SPECIFIC								
Azzi, 2014	Cross-sectional	2002	France	FFQ & 24h recall	ZAC1	Cord blood	No	6
ADULTHOOD								
GLOBAL								
Marques-Rocha, 2016	Cross-sectional	156	Brazil	72h food record	NA	WBCs	Hypo	5

Table 2.1.6. Study characteristics and results: Vitamin B6.

First author, year	Study design	N	Country	Exposure assesment	Genes	Tissue	Direction results	QS
MATERNAL OFFSPRING RELATIONSHIP								
GENE-SPECIFIC								
Azzi, 2014	Cross-sectional	2002	France	FFQ & 24h recall	ZAC1	Cord blood	No	6
ADULTHOOD								
GLOBAL								
Gomes, 2012	Cross-sectional	126	Brazil	24h recall on 3 different days	NA	Venous blood	Hypo	3
Ono, 2012	Cross-sectional	384	Japan	FFQ & 24h recall	NA	Blood	No	5
Huang, 2012	Cross-sectional	493	USA	FFQ	NA	Blood	No	3
Perng, 2014	Cross-sectional	987	USA	FFQ	NA	Blood	No	3
Zhang, 2011	Cross-sectional	149	USA	FFQ	NA	Peripheral blood leukocytes	No	4
Ulrich, 2012	Cross-sectional	173	USA	FFQ	NA	Lymphocytes	No	4
Zhang 2012	Cross-sectional	165	USA	FFQ	NA	WBCs	No	4
Figueiredo, 2009	Cross-sectional	388	North America	FFQ	NA	biopsie from right and left colon	NA (Quartiles)	3
Figueiredo, 2009	Cross-sectional	388	North America	Plasma	NA	biopsie from right and left colon	NA (Quartiles)	4
Friso, 2002	Cross-sectional	292	Italy	Venous blood	NA	Whole blood	No	6
Hübner, 2013	Intervention (non-blinded)	50	Germany	Supplements	NA	Blood	No	3
GENE-SPECIFIC								
Zhang 2012	Cross-sectional	165	USA	FFQ	IL-6	WBCs	No	5

Table 2.1.7. Study characteristics and results: Folate.

First author, year	Study design	N	Country	Exposure assesment	Genes	Methylation Tissue	Direction results	QS
MATERNAL OFFSPRING RELATIONSHIP								
GLOBAL								
Fryer, 2009	Cross-sectional	24	UK	Questionnaire (supplements)	NA	Cord blood	No	2
Boeke, 2012	Cross-sectional	534	USA	FFQ	NA	Cord blood	No	6
Haggarty, 2013	Cross-sectional	913	UK	FFQ	NA	Cord blood	Hypo	4
Haggarty, 2013	Cross-sectional	913	UK	RBC	NA	Cord blood	Hypo	5
McKay, 2012	Cross-sectional	197	UK	RBC	NA	Cord blood	No	5
Chang, 2011	Cross-sectional	20	China	Blood	NA	Brain, skin, heart, kidney, lung & liver	Hyper	2
Park, 2005	Cross-sectional	107	Korea	Serum	NA	Placenta	Hyper	6
GENE-SPECIFIC								
Hoyo, 2011	Cross-sectional	438	USA	Questionnaire	H19	Cord blood	Hypo	5
Azzi, 2014	Cross-sectional	2002	France	FFQ & 24h recall	ZAC1	Cord blood	No	6
Azzi, 2014	Cross-sectional	2002	France	FFQ & 24h recall (supplements)	ZAC1	Cord blood	No	6
van Mil, 2014	Cross-sectional	463	Netherlands	FFQ (supplements)	NR3C1_a	Cord blood	Hyper	6
van Mil, 2014	Cross-sectional	463	Netherlands	FFQ (supplements)	5HTT_b	Cord blood	Hypo	6
Steegers-Theunissen, 2009	Cross-sectional	120	Netherlands	Questionnaire (supplements)	IGF2	Whole blood	Hyper	5
Haggarty, 2013	Cross-sectional	913	UK	FFQ	IGF2	Cord blood	Hyper	5
Haggarty, 2013	Cross-sectional	913	UK	FFQ	PEG3	Cord blood	Hypo	5
Haggarty, 2013	Cross-sectional	913	UK	FFQ	SNRPN	Cord blood	No	5
Haggarty, 2013	Cross-sectional	913	UK	RBC	IGF2	Cord blood	No	6
Haggarty, 2013	Cross-sectional	913	UK	RBC	PEG3	Cord blood	No	6

Table 2.1.7. Study characteristics and results: Folate. (continued)

First author, year	Study design	N	Country	Exposure assesment	Genes	Methylation Tissue	Direction results	QS
Haggarty, 2013	Cross-sectional	913	UK	RBC	SNRPN	Cord blood	No	6
Hoyo, 2014	Cross-sectional	496	USA	RBC	MEG3 (Promoter)	Cord blood	Hypo	7
Hoyo, 2014	Cross-sectional	496	USA	RBC	NNAT (Promoter)	Cord blood	Hypo	7
Hoyo, 2014	Cross-sectional	496	USA	RBC	PEG10/SGCE (Promoter)	Cord blood	Hypo	7
Hoyo, 2014	Cross-sectional	496	USA	RBC	MEG3-IG (Promoter)	Cord blood	Hypo	7
Hoyo, 2014	Cross-sectional	496	USA	RBC	PLAGL1 (Promoter)	Cord blood	Hypo	7
Hoyo, 2014	Cross-sectional	496	USA	RBC	PEG3 (Promoter)	Cord blood	Hypo	7
Hoyo, 2014	Cross-sectional	496	USA	RBC	PEG1/MEST (Promoter)	Cord blood	Hyper	7
Hoyo, 2014	Cross-sectional	496	USA	RBC	H19	Cord blood	Hypo	7
Hoyo, 2014	Cross-sectional	496	USA	RBC	IGF2 (Promoter)	Cord blood	Hyper	7
Ba, 2011	Cross-sectional	99	China	Serum	IGF2 (Promoter)	Cord blood	No	6
van Mil, 2014	Cross-sectional	463	Netherlands	Venous blood	NR3C1_a	Cord blood	Hyper	7
GENOME-WIDE								
Gonseth, 2015	Cross-sectional	347	USA	FFQ	NA	Neonatal blood	Hyper/ Hypo	6
Amarasekera, 2014 - FASEB	Cross-sectional	23 infants	Australia	Serum	NA	Cord blood	Hyper/ Hypo	4

Table 2.1.7. Study characteristics and results: Folate. (continued)

First author, year	Study design	N	Country	Exposure assesment	Genes	Methylation Tissue	Direction results	QS
Joubert, 2016	Cross-sectional	1988	Netherlands and Norway	Plasma	NA	Cord blood	Hyper/ Hypo	8
Mozhui, 2015	Cross-sectional	200	USA	Plasma/Serum	NA	Cord blood from buffy coats	No	6
CHILDHOOD								
GLOBAL								
Fryer, 2009	Cross-sectional	24	UK	Cord serum	NA	Cord blood	No	4
Haggarty, 2013	Cross-sectional	913	UK	RBC	NA	Cord blood	Hypo	5
Perng, 2012	Cross-sectional	568	Colombia	RBC	NA	Blood	No	4
GENE-SPECIFIC								
Haggarty, 2013	Cross-sectional	913	UK	RBC	IGF2	Cord blood	Hyper	6
Haggarty, 2013	Cross-sectional	913	UK	RBC	PEG3	Cord blood	Hypo	6
Haggarty, 2013	Cross-sectional	913	UK	RBC	SNRPN	Cord blood	No	6
ADULTHOOD								
GLOBAL								
Ono, 2012	Cross-sectional	384	Japan	FFQ &24h recall	NA	Blood	Hypo	5
Protiva, 2011	Cross-sectional	20	USA	3d questionnaires & 24h recalls	NA	Blood	No	1
Gomes, 2012	Cross-sectional	126	Brazil	24h recall on 3 d	NA	Venous blood	No	3
Huang, 2012	Cross-sectional	493	USA	FFQ (supplements + diet + fortified)	NA	Blood	No	3
Huang, 2012	Cross-sectional	493	USA	FFQ (DFEs)	NA	Blood	No	3
Huang, 2012	Cross-sectional	493	USA	FFQ (Natural)	NA	Blood	Hyper	3
Boeke, 2012	Cross-sectional	534	USA	FFQ	NA	Blood	No	6
Perng, 2014	Cross-sectional	987	USA	FFQ	NA	Blood	No	3

Table 2.1.7. Study characteristics and results: Folate. (continued)

First author, year	Study design	N	Country	Exposure assesment	Genes	Methylation Tissue	Direction results	QS
Agodi, 2015	Cross-sectional	177	Italy	FFQ	NA	Whole blood	Hyper	5
Zhang 2012	Cross-sectional	165	US	FFQ	NA	WBC	No	4
Zhang 2012	Cross-sectional	165	US	FFQ	NA	WBC	Hyper	5
Zhang 2012	Cross-sectional	165	US	FFQ	NA	WBC	No	4
Zhang, 2011	Cross-sectional	149	US	FFQ	NA	Peripheral blood leukocytes	No	4
Zhang, 2011	Cross-sectional	149	US	FFQ	NA	Peripheral blood leukocytes	No	4
Zhang, 2011	Cross-sectional	149	US	FFQ	NA	Peripheral blood leukocytes	No	4
Ulrich, 2012	Cross-sectional	173	US	FFQ	NA	Lymphocytes	No	4
Figueiredo, 2009	Cross-sectional	388	North America	FFQ	NA	Colonic biopsie	NA (Quartiles)	3
Badiga, 2014	Cross-sectional	325	USA	Plasma	NA	Blood	Hypo	6
Pilsner, 2007	Cross-sectional	294	Bangladesh	Plasma	NA	Blood	Hyper	6
Piyathilake, 2013	Cross-sectional	470	USA	Plasma	NA	Blood	Hyper	4
Gadgil, 2014	Cross-sectional	49	India	Plasma	NA	Blood	No	3
Narayanan, 2004	Cross-sectional	408	Scotland	Plasma	NA	Venous blood	No	4
Tapp, 2013	Cross-sectional	185	UK	Plasma	NA	Rectal mucosa	No	5
Llanos, 2015 - Epigenetics	Cross-sectional	121	USA	Plasma	NA	Breast tissue	No	6
Figueiredo, 2009	Cross-sectional	388	North America	Plasma	NA	Colonic biopsie	NA (Quartiles)	4
Hanks, 2013	Cross-sectional	336	UK	Serum	NA	Colonic tissue biopsies	No	6
Pufulete, 2005	Cross-sectional	68	UK	Serum	NA	Colonic mucosa	No	5

Table 2.1.7. Study characteristics and results: Folate. (continued)

First author, year	Study design	N	Country	Exposure assesment	Genes	Methylation Tissue	Direction results	QS
Ulrich, 2012	Cross-sectional	173	US	Serum	NA	Lymphocytes	No	5
Wang, 2012	Cross-sectional	115	China	Serum	NA	Whole blood	Hyper	6
Fenech, 1998	Cross-sectional	106	Australia	RBC	NA	Blood	No	6
Bae, 2014	Cross-sectional	408	USA	RBC (Prefortification period)	NA	Blood	Hyper	6
Bae, 2014	Cross-sectional	408	USA	RBC (Postfortification period)	NA	Blood	Hyper	6
Figueiredo, 2009	Cross-sectional	388	North America	RBC	NA	Colonic biopsie	NA (Quartiles)	4
Hanks, 2013	Cross-sectional	336	UK	RBC	NA	Colonic tissue biopsies	No	6
Pufulete, 2005	Cross-sectional	68	UK	RBC	NA	Colonic mucosa	No	5
Tapp, 2013	Cross-sectional	185	UK	RBC	NA	Rectal mucosa	No	5
Friso, 2002	Cross-sectional	292	Italy	Venous blood	NA	Whole blood	Hyper	6
Friso, 2005	Cross-sectional	198	Italy	Venous blood	NA	Whole blood	Hyper	5
Hanks, 2013	Cross-sectional	336	UK	Colonic tissue	NA	Colonic tissue biopsies	No	6
Llanos, 2015 - Epigenetics	Cross-sectional	121	USA	Breast tissue	NA	Breast tissue	No	6
Jacob, 1998	Intervention (non-blinded)	8	USA	Restricted diet	NA	Blood	Hypo	2
Rampersaud, 2000	Intervention (non-blinded)	33	USA	Diet	NA	Leukocytes	Hyper	3
Hübner, 2013	Intervention (non-blinded)	50	Germany	Supplementation	NA	Blood	No	3
Axume, 2007	Intervention (non-blinded)	43	USA	7 w of folate restriction with 135 µg DFEs/day followed by 7 weeks of folate treatment with 400 or 800 µg DFEs/day	NA	Blood	Hypo	5

Table 2.1.7. Study characteristics and results: Folate. (continued)

First author, year	Study design	N	Country	Exposure assesment	Genes	Methylation Tissue	Direction results	QS
Aarabi, 2015	Intervention (non-blinded)	28	Canada	Folic acid (5mg/day) for 6 m	NA	Sperm	Hypo	4
Shelnutt, 2004	Intervention (non-blinded)	41	USA	Folate depletion, followed by repletion (7 w, 115 mcg DFE/d; 7 w, 400 mcg DFE/d)	NA	Leukocytes	Hyper	3
Jung, 2011	Intervention (double-blinded)	216	Netherlands	FFQ; folic acid (0.8 mg/d) or placebo	NA	Blood	No	8
Crider, 2011	Intervention (double-blinded)	1108	China	Folic acid supplementation (100, 400, 4000 µg per day) for 6 m	NA	Coagulated and uncoagulated blood	Hypo	8
Basten, 2006	Intervention (double-blinded)	61	UK	Plasma	NA	Whole blood	No	5
Basten, 2006	Intervention (double-blinded)	61	UK	RBC	NA	Whole blood	No	5
Basten, 2006	Intervention (double-blinded)	61	UK	Lymphocytes	NA	Whole blood	No	5
GENE-SPECIFIC								
Bollati, 2014	Cross-sectional	165	Italy	FFQ	CD14 (Promoter)	Blood	No	6
Bollati, 2014	Cross-sectional	165	Italy	FFQ	Et-1 (Promoter)	Blood	No	6
Bollati, 2014	Cross-sectional	165	Italy	FFQ	HERV-w (Promoter)	Blood	No	6
Bollati, 2014	Cross-sectional	165	Italy	FFQ	iNOS (Promoter)	Blood	No	6
Bollati, 2014	Cross-sectional	165	Italy	FFQ	TNFα (Promoter)	Blood	Hypo	6

Table 2.1.7. Study characteristics and results: Folate. (continued)

First author, year	Study design	N	Country	Exposure assesment	Genes	Methylation Tissue	Direction results	QS
Zhong, 2015	Cross-sectional	573	US	FFQ	TLR2 (Promoter)	Whole blood	No	6
Zhang 2012	Cross-sectional	165	US	FFQ	IL-6 (Promoter)	WBC	No	5
Zhang 2012	Cross-sectional	165	US	FFQ	IL-6 (Promoter)	WBC	No	5
Zhang 2012	Cross-sectional	165	US	FFQ	IL-6 (Promoter)	WBC	No	5
Stidley, 2010	Cross-sectional	1101	USA	FFQ	Index (p16, MGMT, DAPK, RASSF1A, PAX5α, PAX5β, GATA4, and GATA5) (Promoter)	Sputum	Hypo	6
Dhillon, 2007	Cross-sectional	379	India	Questionnaire	GSTM1 (Promoter)	Semen	Hypo	5
Ottini, 2015	Cross-sectional	21	Italy	Plasma	Index (p16, FHIT, RAR, CDH1, DAPK1, hTERT, RASSF1A, MGMT, BRCA1 and PALB2) (Promoter)	Blood	Hypo	5

Table 2.1.7. Study characteristics and results: Folate. (continued)

First author, year	Study design	N	Country	Exposure assesment	Genes	Methylation Tissue	Direction results	QS
Piyathilake, 2014	Cross-sectional	315	USA	Plasma	HPV 16 (Promoter)	Blood	Hyper	5
Llanos, 2015 - Carcinogenesis	Cross-sectional	138	USA	Plasma	$p16^{INK4a}$ (Promoter)	Breast tissue	Hypo	7
Tapp, 2013	Cross-sectional	185	UK	Plasma	Index (HPP1, APC, SFRP1, SFRP2, SOX17, WIF1, ESR1, MYOD, N33)	Rectal mucosa	Hyper	6
Tapp, 2013	Cross-sectional	185	UK	Plasma	HPP1	Rectal mucosa	Hyper	6
Tapp, 2013	Cross-sectional	185	UK	Plasma	APC	Rectal mucosa	No	6
Tapp, 2013	Cross-sectional	185	UK	Plasma	SFRP1	Rectal mucosa	Hyper	6
Tapp, 2013	Cross-sectional	185	UK	Plasma	SFRP2	Rectal mucosa	Hyper	6
Tapp, 2013	Cross-sectional	185	UK	Plasma	SOX17	Rectal mucosa	No	6
Tapp, 2013	Cross-sectional	185	UK	Plasma	WIF1	Rectal mucosa	Hyper	6
Tapp, 2013	Cross-sectional	185	UK	Plasma	ESR1	Rectal mucosa	No	6
Tapp, 2013	Cross-sectional	185	UK	Plasma	MYOD	Rectal mucosa	Hyper	6
Tapp, 2013	Cross-sectional	185	UK	Plasma	N33	Rectal mucosa	No	6
Hanks, 2013	Cross-sectional	336	UK	Serum	MLH1 (Promoter)	Colonic tissue biopsies	No	7
Hanks, 2013	Cross-sectional	336	UK	Serum	MGMT (Promoter)	Colonic tissue biopsies	Hypo	7
Hanks, 2013	Cross-sectional	336	UK	Serum	APC (Promoter)	Colonic tissue biopsies	No	7

Table 2.1.7. Study characteristics and results: Folate. (continued)

First author, year	Study design	N	Country	Exposure assesment	Genes	Methylation Tissue	Direction results	QS
Hanks, 2013	Cross-sectional	336	UK	Serum	IGF2 (Promoter)	Colonic tissue biopsies	No	7
Hanks, 2013	Cross-sectional	336	UK	Serum	MYOD1 (Promoter)	Colonic tissue biopsies	No	7
Hanks, 2013	Cross-sectional	336	UK	Serum	N33 (Promoter)	Colonic tissue biopsies	No	7
Hanks, 2013	Cross-sectional	336	UK	Serum	ESR1 (Promoter)	Colonic tissue biopsies	No	7
Hanks, 2013	Cross-sectional	336	UK	RBC	MLH1 (Promoter)	Colonic tissue biopsies	No	7
Hanks, 2013	Cross-sectional	336	UK	RBC	MGMT (Promoter)	Colonic tissue biopsies	No	7
Hanks, 2013	Cross-sectional	336	UK	RBC	APC (Promoter)	Colonic tissue biopsies	No	7
Hanks, 2013	Cross-sectional	336	UK	RBC	IGF2 (Promoter)	Colonic tissue biopsies	No	7
Hanks, 2013	Cross-sectional	336	UK	RBC	MYOD1 (Promoter)	Colonic tissue biopsies	No	7
Hanks, 2013	Cross-sectional	336	UK	RBC	N33 (Promoter)	Colonic tissue biopsies	No	7
Hanks, 2013	Cross-sectional	336	UK	RBC	ESR1 (Promoter)	Colonic tissue biopsies	No	7

Table 2.1.7. Study characteristics and results: Folate. (continued)

First author, year	Study design	N	Country	Exposure assesment	Genes	Methylation Tissue	Direction results	QS
Tapp, 2013	Cross-sectional	185	UK	RBC	Index (HPP1, APC, SFRP1, SFRP2, SOX17, WIF1, ESR1, MYOD, N33)	Rectal mucosa	Hyper	6
Tapp, 2013	Cross-sectional	185	UK	RBC	HPP1	Rectal mucosa	No	6
Tapp, 2013	Cross-sectional	185	UK	RBC	APC	Rectal mucosa	No	6
Tapp, 2013	Cross-sectional	185	UK	RBC	SFRP1	Rectal mucosa	Hyper	6
Tapp, 2013	Cross-sectional	185	UK	RBC	SFRP2	Rectal mucosa	Hyper	6
Tapp, 2013	Cross-sectional	185	UK	RBC	SOX17	Rectal mucosa	No	6
Tapp, 2013	Cross-sectional	185	UK	RBC	WIF1	Rectal mucosa	No	6
Tapp, 2013	Cross-sectional	185	UK	RBC	ESR1	Rectal mucosa	No	6
Tapp, 2013	Cross-sectional	185	UK	RBC	MYOD	Rectal mucosa	Hyper	6
Tapp, 2013	Cross-sectional	185	UK	RBC	N33	Rectal mucosa	No	6
Hanks, 2013	Cross-sectional	336	UK	Colonic tissue	MLH1 (Promoter)	Colonic tissue biopsies	No	7
Hanks, 2013	Cross-sectional	336	UK	Colonic tissue	MGMT (Promoter)	Colonic tissue biopsies	No	7
Hanks, 2013	Cross-sectional	336	UK	Colonic tissue	APC (Promoter)	Colonic tissue biopsies	No	7
Hanks, 2013	Cross-sectional	336	UK	Colonic tissue	IGF2 (Promoter)	Colonic tissue biopsies	No	7
Hanks, 2013	Cross-sectional	336	UK	Colonic tissue	MYOD1 (Promoter)	Colonic tissue biopsies	No	7

Table 2.1.7. Study characteristics and results: Folate. (continued)

First author, year	Study design	N	Country	Exposure assesment	Genes	Methylation Tissue	Direction results	QS
Hanks, 2013	Cross-sectional	336	UK	Colonic tissue	N33 (Promoter)	Colonic tissue biopsies	No	7
Hanks, 2013	Cross-sectional	336	UK	Colonic tissue	ESR1 (Promoter)	Colonic tissue biopsies	No	7
Llanos, 2015 - Carcinogenesis	Cross-sectional	138	USA	Breast tissue	$p16^{INK4a}$ (Promoter)	Breast tissue	Hypo	7
Aarabi, 2015	Intervention (non-blinded)	30	Canada	Folic acid (5mg/d) for 6 m	H19	Sperm	No	5
Aarabi, 2015	Intervention (non-blinded)	30	Canada	Folic acid (5mg/d) for 6 m	DLK1/GTL2	Sperm	No	5
Aarabi, 2015	Intervention (non-blinded)	30	Canada	Folic acid (5mg/d) for 6 m	MEST	Sperm	No	5
Aarabi, 2015	Intervention (non-blinded)	30	Canada	Folic acid (5mg/d) for 6 m	SNRPN	Sperm	No	5
Aarabi, 2015	Intervention (non-blinded)	30	Canada	Folic acid (5mg/d) for 6 m	PLAGL1	Sperm	No	5
Aarabi, 2015	Intervention (non-blinded)	30	Canada	Folic acid (5mg/d) for 6 m	KCNQ1OT1	Sperm	No	5
Al-Ghnaniem Abbadi, 2013	Intervention (double-blinded)	29	UK	Folic acid (400 µg/d) or placebo for 10 w	ESR1 (Promoter)	Normal-appearing colonic mucosa	No	7
Al-Ghnaniem Abbadi, 2013	Intervention (double-blinded)	29	UK	Folic acid (400 µg/d) or placebo for 10 w	MLH1 (Promoter)	Normal-appearing colonic mucosa	No	7
GENOME-WIDE								
Song, 2016	Cross-sectional	81	USA	Breast tissue	NA	Breast tissue	Hyper/ Hypo	4

Table 2.1.8. Study characteristics and results: Vitamin B12.

First author, year	Study design	N	Country	Exposure assesment	Genes	Tissue	Direction results	QS
MATERNAL OFFSPRING RELATIONSHIP								
GLOBAL								
Boeke, 2012	Cross-sectional	534	USA	FFQ	NA	Cord blood	No	6
McKay, 2012	Cross-sectional	158	UK	Serum	NA	Cord blood	Hyper	5
GENE-SPECIFIC								
Azzi, 2014	Cross-sectional	2002	France	FFQ & 24h recall	ZAC1	Cord blood	Hyper	6
Ba, 2011	Cross-sectional	99	China	Serum	IGF2	Cord blood	Hypo	6
CHILDHOOD								
GLOBAL								
Perng, 2012	Cross-sectional	568	Colombia	Plasma	NA	Blood	No	4
GENE-SPECIFIC								
McKay, 2012	Cross-sectional	158	UK	Serum cord blood	IGFBP3	Cord blood	Hypo	7
ADULTHOOD								
GLOBAL								
Gomes, 2012	Cross-sectional	126	Brazil	24h recall (3 d)	NA	Venous blood	No	3
Ono, 2012	Cross-sectional	384	Japan	FFQ &24 h recall	NA	Blood	No	5
Huang, 2012	Cross-sectional	493	USA	FFQ	NA	Blood	No	3
Perng, 2014	Cross-sectional	987	USA	FFQ	NA	Blood	No	5
Boeke, 2012	Cross-sectional	534	USA	FFQ	NA	Blood	No	6
Ulrich, 2012	Cross-sectional	173	USA	FFQ	NA	Lymphocytes	No	4
Zhang 2012	Cross-sectional	165	USA	FFQ	NA	WBCs	No	4

Table 2.1.8. Study characteristics and results: Vitamin B12. (continued)

First author, year	Study design	N	Country	Exposure assesment	Genes	Tissue	Direction results	QS
Zhang, 2011	Cross-sectional	149	USA	FFQ	NA	Peripheral blood leukocytes	No	4
Figueiredo, 2009	Cross-sectional	388	North America	Plasma	NA	Biopsie from right and left colon	NA (Quartiles)	4
Narayanan, 2004	Cross-sectional	408	Scotland	Plasma	NA	Venous blood	No	4
Quinlivan, 2013	Cross-sectional	376	China	Plasma	NA	Coagulated Blood clots vs. uncoagulated EDTA-Blood cells	Hyper	4
Quinlivan, 2013	Cross-sectional	376	China	Plasma	NA	Blood clots	Hyper	6
Gadgil, 2014	Cross-sectional	49	India	Plasma	NA	Blood	No	3
Tapp, 2013	Cross-sectional	185	UK	Plasma	NA	Rectal mucosa	No	5
Fenech, 1998	Cross-sectional	106	Australia	Serum	NA	Blood	No	6
Friso, 2002	Cross-sectional	292	Italy	venous blood	NA	Whole blood	No	6
Ulrich, 2012	Cross-sectional	173	USA	Serum	NA	Lymphocytes	No	5
Pufulete, 2005	Cross-sectional	68	UK	Serum	NA	Colonic mucosa	No	3
Hübner, 2013	Intervention (non-blinded)	50	Germany	Supplements	NA	Blood	No	3
GENE-SPECIFIC								
Bollati, 2014	Cross-sectional	165	Italy	FFQ	CD14	Blood	No	6
Bollati, 2014	Cross-sectional	165	Italy	FFQ	Et-1	Blood	No	6
Bollati, 2014	Cross-sectional	165	Italy	FFQ	HERV-w	Blood	No	6
Bollati, 2014	Cross-sectional	165	Italy	FFQ	iNOS	Blood	No	6

Table 2.1.8. Study characteristics and results: Vitamin B12. (continued)

First author, year	Study design	N	Country	Exposure assesment	Genes	Tissue	Direction results	QS
Bollati, 2014	Cross-sectional	165	Italy	FFQ	TNFα	Blood	No	6
Zhang 2012	Cross-sectional	165	USA	FFQ	IL-6	WBCs	No	5
Zhong, 2015	Cross-sectional	573	USA	FFQ	TLR2	Whole blood	No	6
Piyathilake, 2014	Cross-sectional	315	USA	Plasma	HPV 16	Blood	Hyper	5
Ottini, 2015	Cross-sectional	22	Italy	Plasma	Index (p16, FHIT, RAR, CDH1, DAPK1, hTERT, RASSF1A, MGMT, BRCA1 and PALB2)	Blood	No	3
Tapp, 2013	Cross-sectional	185	UK	Plasma	Index (HPP1, APC, SFRP1, SFRP2, SOX17, WIF1, ESR1, MYOD, N33)	Rectal mucosa	No	6
Tapp, 2013	Cross-sectional	185	UK	Plasma	HPP1	Rectal mucosa	No	6
Tapp, 2013	Cross-sectional	185	UK	Plasma	APC	Rectal mucosa	No	6
Tapp, 2013	Cross-sectional	185	UK	Plasma	SFRP1	Rectal mucosa	No	6
Tapp, 2013	Cross-sectional	185	UK	Plasma	SFRP2	Rectal mucosa	No	6
Tapp, 2013	Cross-sectional	185	UK	Plasma	SOX17	Rectal mucosa	No	6
Tapp, 2013	Cross-sectional	185	UK	Plasma	WIF1	Rectal mucosa	No	6
Tapp, 2013	Cross-sectional	185	UK	Plasma	ESR1	Rectal mucosa	No	6
Tapp, 2013	Cross-sectional	185	UK	Plasma	MYOD	Rectal mucosa	No	6
Tapp, 2013	Cross-sectional	185	UK	Plasma	N33	Rectal mucosa	No	6
Al-Ghnaniem, 2007	Cross-sectional	73	UK	Venous blood	ERα	Normal-appearing colonic mucosa	Hypo	6

Table 2.1.9. Study characteristics and results: Vitamin C.

First author, year	Study design	N	Country	Exposure assesment	Genes	Tissue	Direction results	QS
ADULTHOOD								
GLOBAL								
Piyathilake, 2011	Cross-sectional	376	USA	Plasma	NA	PBMC	No	6
Piyathilake, 2011	Cross-sectional	376	USA	Plasma	NA	Cervical cells	No	6
GENE-SPECIFIC								
Bollati, 2014	Cross-sectional	165	Italy	FFQ	CD14	Blood	No	6
Bollati, 2014	Cross-sectional	165	Italy	FFQ	Et-1	Blood	No	6
Bollati, 2014	Cross-sectional	165	Italy	FFQ	HERV-w	Blood	No	6
Bollati, 2014	Cross-sectional	165	Italy	FFQ	iNOS	Blood	No	6
Bollati, 2014	Cross-sectional	165	Italy	FFQ	TNFα	Blood	No	6
Stidley, 2010	Cross-sectional	1101	USA	FFQ	Index (p16, MGMT, DAPK, RASSF1A, PAX5α, PAX5β, GATA4, and GATA5)	Sputum	No	6
Piyathilake, 2014	Cross-sectional	315	USA	Plasma	HPV 16	Blood	No	5
de la Iglesia, 2014	Intervention (non-blinded)	47	Spain	48 h weighed food record	PON1	Venous blood	Hypo	6

Table 2.1.10. Study characteristics and results: Vitamin D

First author, year	Study design	N	Country	Exposure assesment	Genes	Tissue	Direction results	QS
MATERNAL OFFSPRING RELATIONSHIP								
GENE-SPECIFIC								
Novakovic, 2016	Cross-sectional	68	Canada	Serum	CYP24A1	Placenta	No	4
Harvey, 2014	Cross-sectional	230	UK	Plasma	RXRA	umbilical cord	Hypo	5
GENOME-WIDE								
Mozhui, 2015	Cross-sectional	147	USA	Plasma/ Serum	NA	Cord blood from buffy coats	No	6
CHILDHOOD								
GLOBAL								
Zhu, 2016	Cross-sectional	454	USA	Plasma	NA	Peripherial Blood	Hyper	6
Zhu, 2016	Intervention (double-blinded)	58	USA	Supplements	NA	Peripherial Blood	Hyper	7
GENE-SPECIFIC								
Zhu, 2013	Cross-sectional	22	USA	Plasma	DHCR7	Whole blood leukocytes	Opposite	4
Zhu, 2013	Cross-sectional	22	USA	Plasma	CYP2R1	Whole blood leukocytes	Hypo	4
Zhu, 2013	Cross-sectional	22	USA	Plasma	CYP24A1	Whole blood leukocytes	Hyper	4
Novakovic, 2016	Cross-sectional	68	Canada	Serum	CYP24A1	Placenta	No	4
GENOME-WIDE								
Zhu, 2013	Cross-sectional	22	USA	Plasma	NA	Whole blood leukocytes	Hyper/ Hypo	5
ADULTHOOD								
GLOBAL								
Tapp, 2013	Cross-sectional	185	UK	Plasma	NA	Rectal mucosa	No	5
Nair-Shalliker, 2014	Cross-sectional	208	Australia	Serum	NA	Lymphocytes	No	5
Hübner, 2013	Intervention (non-blinded)	50	Germany	Supplements	NA	Blood	No	3
GENE-SPECIFIC								
Bollati, 2014	Cross-sectional	165	Italy	FFQ	CD14	Blood	No	6
Bollati, 2014	Cross-sectional	165	Italy	FFQ	Et-1	Blood	No	6

Table 2.1.10. Study characteristics and results: Vitamin D (continued)

First author, year	Study design	N	Country	Exposure assesment	Genes	Tissue	Direction results	QS
Bollati, 2014	Cross-sectional	165	Italy	FFQ	HERV-w	Blood	No	6
Bollati, 2014	Cross-sectional	165	Italy	FFQ	iNOS	Blood	No	6
Bollati, 2014	Cross-sectional	165	Italy	FFQ	TNFα	Blood	No	6
Ashktorab, 2011	Cross-sectional	187	USA	Serum	DKK1	Whole blood	No	7
Tapp, 2013	Cross-sectional	185	UK	Plasma	Index (HPP1, APC, SFRP1, SFRP2, SOX17, WIF1, ESR1, MYOD, N33)	Rectal mucosa	Hypo	6
Tapp, 2013	Cross-sectional	185	UK	Plasma	HPP1	Rectal mucosa	No	6
Tapp, 2013	Cross-sectional	185	UK	Plasma	APC	Rectal mucosa	Hypo	6
Tapp, 2013	Cross-sectional	185	UK	Plasma	SFRP1	Rectal mucosa	No	6
Tapp, 2013	Cross-sectional	185	UK	Plasma	SFRP2	Rectal mucosa	No	6
Tapp, 2013	Cross-sectional	185	UK	Plasma	SOX17	Rectal mucosa	No	6
Tapp, 2013	Cross-sectional	185	UK	Plasma	WIF1	Rectal mucosa	Hypo	6
Tapp, 2013	Cross-sectional	185	UK	Plasma	ESR1	Rectal mucosa	No	6
Tapp, 2013	Cross-sectional	185	UK	Plasma	MYOD	Rectal mucosa	Hypo	6
Tapp, 2013	Cross-sectional	185	UK	Plasma	N33	Rectal mucosa	No	6

Table 2.1.11. Study characteristics and results: Vitamin E.

First author, year	Study design	N	Country	Exposure assesment	Genes	Tissue	Direction results	QS
ADULTHOOD								
GLOBAL								
Piyathilake, 2011	Cross-sectional	376	USA	Plasma	NA	PBMC	Hypo	6
Piyathilake, 2011	Cross-sectional	376	USA	Plasma	NA	Cervical cells	No	6
GENE-SPECIFIC								
Bollati, 2014	Cross-sectional	165	Italy	FFQ	CD14	Blood	No	6
Bollati, 2014	Cross-sectional	165	Italy	FFQ	Et-1	Blood	No	6
Bollati, 2014	Cross-sectional	165	Italy	FFQ	HERV-w	Blood	No	6
Bollati, 2014	Cross-sectional	165	Italy	FFQ	iNOS	Blood	No	6
Bollati, 2014	Cross-sectional	165	Italy	FFQ	TNFα	Blood	No	6
de la Iglesia, 2014	Intervention (non-blinded)	47	Spain	48 h weighed food record	PON1	Venous blood	Hypo	6
Stidley, 2010	Cross-sectional	1101	USA	FFQ	Index (p16, MGMT, DAPK, RASSF1A, PAX5α, PAX5β, GATA4, and GATA5)	Sputum	No	6

Table 2.1.12. Study characteristics and results: Choline and betaine

Nutritional exposure	First author, year	Study design	N	Country	Exposure assesment	Genes	Tissue	Direction results	QS
MATERNAL OFFSPRING RELATIONSHIP									
GLOBAL									
Choline	Boeke, 2012	Cross-sectional	534	USA	FFQ	NA	Cord blood	Hypo	6
Choline	Jiang, 2012	Intervention (non-blinded)	24	USA	12-wk controlled feeding study	NA	Placenta	Hyper	4
Choline	Jiang, 2012	Intervention (non-blinded)	24	USA	12-wk controlled feeding study	NA	Cord blood	No	4
Betaine	Boeke, 2012	Cross-sectional	534	USA	FFQ	NA	Cord blood	Hypo	6
GENE-SPECIFIC									

Table 2.1.12. Study characteristics and results: Choline and betaine (continued)

Nutritional exposure	First author, year	Study design	N	Country	Exposure assesment	Genes	Tissue	Direction results	QS
Choline	Jiang, 2012	Intervention (non-blinded)	24	USA	12-wk controlled feeding study	CRH	Cord blood	Hypo	5
Choline	Jiang, 2012	Intervention (non-blinded)	24	USA	12-wk controlled feeding study	NR3C1	Cord blood	Hypo	5
Choline	Jiang, 2012	Intervention (non-blinded)	24	USA	12-wk controlled feeding study	GNAS-AS1	Cord blood	No	5
Choline	Jiang, 2012	Intervention (non-blinded)	24	USA	12-wk controlled feeding study	IGF2	Cord blood	No	5
Choline	Jiang, 2012	Intervention (non-blinded)	24	USA	12-wk controlled feeding study	IL10	Cord blood	No	5
Choline	Jiang, 2012	Intervention (non-blinded)	24	USA	12-wk controlled feeding study	LEP	Cord blood	No	5
Choline	Jiang, 2012	Intervention (non-blinded)	24	USA	12-wk controlled feeding study	CRH	Placenta	Hyper	5
Choline	Jiang, 2012	Intervention (non-blinded)	24	USA	12-wk controlled feeding study	NR3C1	Placenta	Hyper	5
Choline	Jiang, 2012	Intervention (non-blinded)	24	USA	12-wk controlled feeding study	GNAS-AS1	Placenta	No	5
Choline	Jiang, 2012	Intervention (non-blinded)	24	USA	12-wk controlled feeding study	IGF2	Placenta	No	5
Choline	Jiang, 2012	Intervention (non-blinded)	24	USA	12-wk controlled feeding study	IL10	Placenta	No	5

Table 2.1.12. Study characteristics and results: Choline and betaine (continued)

Nutritional exposure	First author, year	Study design	N	Country	Exposure assesment	Genes	Tissue	Direction results	QS
Choline	Jiang, 2012	Intervention (non-blinded)	24	USA	12-wk controlled feeding study	LEP	Placenta	No	5
ADULTHOOD									
GLOBAL									
Choline	Boeke, 2012	Cross-sectional	534	USA	FFQ	NA	Blood	No	6
Choline	Jiang, 2012	Intervention (non-blinded)	24	USA	12-wk controlled feeding study	NA	Blood	No	4
Choline	Shin, 2010	Intervention (non-blinded)	60	USA	Intervention (12-wk of 300, 550, 1100, or 2200 mg/d)	NA	Leukocytes	Hyper	3
Betaine	Boeke, 2012	Cross-sectional	534	USA	FFQ	NA	Blood	No	6
GENE-SPECIFIC									
Choline	Jiang, 2012	Intervention (non-blinded)	24	USA	12-wk controlled feeding study	CRH	Blood	No	5
Choline	Jiang, 2012	Intervention (non-blinded)	24	USA	12-wk controlled feeding study	NR3C1	Blood	No	5

Table 2.1.13. Study characteristics and results: Minerals and trace elements

Nutritional exposure	First author, year	Study design	N	Country	Exposure assesment	Genes	Tissue	Direction results	QS
CHILDHOOD									
GLOBAL									
Iron (Ferritin)	Perng, 2012	Cross-sectional	568	Colombia	Plasma	NA	Blood	No	4
Zinc	Perng, 2012	Cross-sectional	568	Colombia	Serum	NA	Blood	No	4
ADULTHOOD									
GLOBAL									
Calcium	Hübner, 2013	Intervention (non-blinded)	50	Germany	Supplements	NA	Blood	No	3
Copper	Marques-Rocha, 2016	Cross-sectional	156	Brazil	72h food record	NA	WBCs	Hyper	5
Iron	Marques-Rocha, 2016	Cross-sectional	156	Brazil	72h food record	NA	WBCs	Hyper	5
Magnesium	Gomes, 2012	Cross-sectional	126	Brazil	24h recall o(3 d)	NA	Venous blood	Hypo	3
Phosphate	McChelland, 2016	Cross-sectional	666	UK	Serum	NA	serum Blood	Hypo	6
Selenium	Tapp, 2013	Cross-sectional	185	UK	Plasma	NA	Rectal mucosa	Hyper	5
Zinc	Perng, 2014	Cross-sectional	987	USA	FFQ	NA	Blood	No	5
Zinc	Gomes, 2012	Cross-sectional	126	Brazil	24h recall (3 d)	NA	Venous blood	No	3
GENE-SPECIFIC									
Sodium (Salt)	Shimazu, T2015	Cross-sectional	281	Japan	FFQ	miR-124a-3	Gastric mucosa	No	6
Sodium (Salt)	Shimazu, T2015	Cross-sectional	281	Japan	FFQ	EMX1	Gastric mucosa	No	6
Sodium (Salt)	Shimazu, T2015	Cross-sectional	281	Japan	FFQ	NKX6-1	Gastric mucosa	No	6
Selenium	Tapp, 2013	Cross-sectional	185	UK	Plasma	HPP1	Rectal mucosa	No	6
Selenium	Tapp, 2013	Cross-sectional	185	UK	Plasma	APC	Rectal mucosa	No	6

Table 2.1.13. Study characteristics and results: Minerals and trace elements (continued)

Nutritional exposure	First author, year	Study design	N	Country	Exposure assesment	Genes	Tissue	Direction results	QS
Selenium	Tapp, 2013	Cross-sectional	185	UK	Plasma	SFRP1	Rectal mucosa	No	6
Selenium	Tapp, 2013	Cross-sectional	185	UK	Plasma	SFRP2	Rectal mucosa	No	6
Selenium	Tapp, 2013	Cross-sectional	185	UK	Plasma	SOX17	Rectal mucosa	No	6
Selenium	Tapp, 2013	Cross-sectional	185	UK	Plasma	WIF1	Rectal mucosa	No	6
Selenium	Tapp, 2013	Cross-sectional	185	UK	Plasma	ESR1	Rectal mucosa	No	6
Selenium	Tapp, 2013	Cross-sectional	185	UK	Plasma	MYOD	Rectal mucosa	No	6
Selenium	Tapp, 2013	Cross-sectional	185	UK	Plasma	N33	Rectal mucosa	No	6
Selenium	Tapp, 2013	Cross-sectional	185	UK	Plasma	Index (HPP1, APC, SFRP1, SFRP2, SOX17, WIF1, ESR1, MYOD, N33)	Rectal mucosa	No	6

Table 2.1.14. Study characteristics and results: Combined nutrients

Nutritional exposure	First author, year	Study design	N	Country	Exposure assesment	Genes	Tissue	Direction results	QS
MATERNAL OFFSPRING RELATIONSHIP									
GENOME-WIDE									
Multivitamin	Khulan, 2012	Intervention (double-blinded)	59	Gambia	Supplements	NA	Cord blood	Hyper/Hypo	6
Multivitamin	Khulan, 2012	Intervention (double-blinded)	25	Gambia	Supplements	NA	Peripheral Blood from infants (about 9 months old)	Hyper/Hypo	6

Table 2.1.14. Study characteristics and results: Combined nutrients (continued)

Nutritional exposure	First author, year	Study design	N	Country	Exposure assesment	Genes	Tissue	Direction results	QS
ADULTHOOD									
GLOBAL									
Folate & Vitamin B12	Piyathilake, 2011	Cross-sectional	376	USA	Plasma	NA	PBMC	Hyper	6
Folate & Vitamin B12	Piyathilake, 2011	Cross-sectional	376	USA	Plasma	NA	Cervical cells	No	6
Vitamin B6, B12, Folate, Vitamin D & Calcium vs. (Control: Vitamin D & Calcium)	Pusceddu, 2016	Intervention (double-blinded)	65	Germany	Supplements; Serum after 1 y intervention	NA	Blood	Hyper/Hypo	5
Vitamin D & Calcium	Pusceddu, 2016	Intervention (double-blinded)	65	Germany	Serum	NA	Blood	Hypo	5
Folate & Vitamin B12	Fenech, 1998	Intervention (double-blinded)	64	Australia	Supplements	NA	Blood	No	6
GENE-SPECIFIC									
Multivitamimin	Stidley, 2010	Cross-sectional	1101	USA	FFQ	Index (p16, MGMT, DAPK, RASSF1A, PAX5α, PAX5β, GATA4, and GATA5)	Sputum	Hypo	6
Calcium & Vitamin D	Zhou, 2014	Intervention (non-blinded)	446	USA	Supplements (Vitamin D (1100 IU/d) & calcium (1400-1500 mg/d))	CYP24A1	Serum	Hypo	5

Table 2.1.14. Study characteristics and results: Combined nutrients (continued)

Nutritional exposure	First author, year	Study design	N	Country	Exposure assesment	Genes	Tissue	Direction results	QS
Calcium & Vitamin D	Zhou, 2014	Intervention (non-blinded)	446	USA	Supplements (Vitamin D (1100 IU/d) & calcium (1400-1500 mg/d))	CYP2R1	Serum	No	5
Calcium & Vitamin D	Zhou, 2014	Intervention (non-blinded)	446	USA	Supplements (Vitamin D (1100 IU/d) & calcium (1400-1500 mg/d))	CYP27A1	Serum	No	5
Calcium & Vitamin D	Zhou, 2014	Intervention (non-blinded)	446	USA	Supplements (Vitamin D (1100 IU/d) & calcium (1400-1500 mg/d))	CYP27B1	Serum	No	5
GENOME-WIDE									
Folate & Vitamin B12	Kok, 2015	Intervention (double-blinded)	92	Netherlands	Supplements; Serum measured after 2 y intervention	NA	Buffy coats	Hyper/Hypo	8
Folate & Vitamin B12	Kok, 2015	Intervention (double-blinded)	92	Netherlands	Supplements; Serum measured after 2 y intervention	NA	Buffy coats	Hyper/Hypo	8

Table 2.1.15. Study characteristics and results: Bioactive compounds

Nutritional exposure	First author, year	Study design	N	Country	Exposure assesment	Genes	Tissue	Direction results	QS
ADULTHOOD									
GLOBAL									
Carotene	Piyathilake, 2011	Cross-sectional	376	USA	Plasma	NA	PBMC	No	6
Carotene	Piyathilake, 2011	Cross-sectional	376	USA	Plasma	NA	Cervical cells	No	6
GENE-SPECIFIC									
Carotenoids	Bollati, 2014	Cross-sectional	165	Italy	FFQ	CD14	Blood	No	6
Carotenoids	Bollati, 2014	Cross-sectional	165	Italy	FFQ	Et-1	Blood	No	6
Carotenoids	Bollati, 2014	Cross-sectional	165	Italy	FFQ	HERV-w	Blood	Hyper	6
Carotenoids	Bollati, 2014	Cross-sectional	165	Italy	FFQ	iNOS	Blood	No	6
Carotenoids	Bollati, 2014	Cross-sectional	165	Italy	FFQ	TNFα	Blood	Hypo	6
Carotene	Piyathilake, 2014	Cross-sectional	315	USA	Plasma	HPV 16	Blood	No	5
Carotene	Stidley, 2010	Cross-sectional	1101	USA	FFQ	Index (p16, MGMT, DAPK, RASSF1A, PAX5α, PAX5β, GATA4, and GATA5)	Sputum	No	6
Lycopene	Stidley, 2010	Cross-sectional	1101	USA	FFQ	Index (p16, MGMT, DAPK, RASSF1A, PAX5α, PAX5β, GATA4, and GATA5)	Sputum	No	6
Lycopene	de la Iglesia, 2014	Intervention (non-blinded)	47	Spain	48h weighed food record	PON1	Venous blood	Hypo	6
Flavonoids	Zhong, 2015	Cross-sectional	573	USA	FFQ	TLR2	Whole blood	Hypo	6

Table 2.1.15. Study characteristics and results: Bioactive compounds (continued)

Nutritional exposure	First author, year	Study design	N	Country	Exposure assesment	Genes	Tissue	Direction results	QS
Polyphenols & Flavonoids	Bollati, 2014	Cross-sectional	165	Italy	FFQ	CD14	Blood	No	6
Polyphenols & Flavonoids	Bollati, 2014	Cross-sectional	165	Italy	FFQ	Et-1	Blood	No	6
Polyphenols & Flavonoids	Bollati, 2014	Cross-sectional	165	Italy	FFQ	HERV-w	Blood	No	6
Polyphenols & Flavonoids	Bollati, 2014	Cross-sectional	165	Italy	FFQ	iNOS	Blood	No	6
Polyphenols & Flavonoids	Bollati, 2014	Cross-sectional	165	Italy	FFQ	TNFα	Blood	No	6
Lutein & zeaxanthin	Stidley, 2010	Cross-sectional	1101	USA	FFQ	Index (p16, MGMT, DAPK, RASSF1A, PAX5α, PAX5β, GATA4, and GATA5)	Sputum	No	6

Table 2.1.16. Study characteristics and results: Fat and fatty acids

Nutritional exposure	First author, year	Study design	N	Country	Exposure assesment	Genes	Tissue	Direction results	QS
MATERNAL OFFSPRING RELATIONSHIP									
GLOBAL									
Fat	Rerkasem, 2015	Cross-sectional	249	Thailand	FFQ & 24h recall	NA	Venous blood samples in children	No	3
PUFA n-3 [docosahexaenoic acid (DHA)]	Lee, 2013	Intervention (double-blinded)	261	Mexico	Supplements	NA	Cord blood mononuclear cells	Hyper	8
GENE-SPECIFIC									
Fat	Godfrey, 2011	Cross-sectional	78	UK	FFQ	RXRA	Cord tissue	No	3
Fat	Godfrey, 2011	Cross-sectional	78	UK	FFQ	eNOS	Cord tissue	No	3
PUFA n-3 [docosahexaenoic acid (DHA)]	Lee, 2013	Intervention (double-blinded)	261	Mexico	Supplements	IFNγ	Cord blood mononuclear cells	No	9
PUFA n-3 [docosahexaenoic acid (DHA)]	Lee, 2013	Intervention (double-blinded)	261	Mexico	Supplements	TNF-α	Cord blood mononuclear cells	No	9
PUFA n-3 [docosahexaenoic acid (DHA)]	Lee, 2013	Intervention (double-blinded)	261	Mexico	Supplements	IL13	Cord blood mononuclear cells	No	9
PUFA n-3 [docosahexaenoic acid (DHA)]	Lee, 2013	Intervention (double-blinded)	261	Mexico	Supplements	GATA3	Cord blood mononuclear cells	No	9
PUFA n-3 [docosahexaenoic acid (DHA)]	Lee, 2013	Intervention (double-blinded)	261	Mexico	Supplements	STAT3	Cord blood mononuclear cells	No	9

Table 2.1.16. Study characteristics and results: Fat and fatty acids (continued)

Nutritional exposure	First author, year	Study design	N	Country	Exposure assesment	Genes	Tissue	Direction results	QS
PUFA n-3 [docosahexaenoic acid (DHA)]	Lee, 2013	Intervention (double-blinded)	261	Mexico	Supplements	IL10	Cord blood mononuclear cells	No	9
PUFA n-3 [docosahexaenoic acid (DHA)]	Lee, 2013	Intervention (double-blinded)	261	Mexico	Supplements	FOXP3	Cord blood mononuclear cells	No	9
PUFA n-3 [docosahexaenoic acid (DHA)]	Lee, 2014	Intervention (double-blinded)	261	Mexico	Supplements	IGF2	Cord blood mononuclear cells	Hyper	9
GENOME-WIDE									
PUFA n-3	Amarasekera, 2014 - Epigenetics	Intervention (double-blinded)	70	Australia	Intervention; 3.7 g of fish oil (with 56.0% as DHA and 27.7% as EPA) or placebo in capsules daily from 20 weeks of gestation until delivery	NA	Cord blood	No	8
CHILDHOOD									
GENOME-WIDE									
Fat	Voisin. 2014	Cross-sectional	69	Greece	24h recalls (3 d)	NA	Peripheral whole blood	Hyper/ Hypo	5
MUFA/SFA	Voisin. 2014	Cross-sectional	69	Greece	24h recalls (3 d)	NA	Peripheral whole blood	Hyper/ Hypo	5
PUFA/SFA	Voisin. 2014	Cross-sectional	69	Greece	24h recalls (3 d)	NA	Peripheral whole blood	Hyper/ Hypo	5

Table 2.1.16. Study characteristics and results: Fat and fatty acids (continued)

Nutritional exposure	First author, year	Study design	N	Country	Exposure assesment	Genes	Tissue	Direction results	QS
(MUFA+PUFA)/SFA	Voisin. 2014	Cross-sectional	69	Greece	24h recalls (3 d)	NA	Peripheral whole blood	Hyper/ Hypo	5
PUFA n-3	Lind, 2015	Intervention (double-blinded)	12	Denmark	Intervention (Fish oil)	NA	Buffy coats	No	8
ADULTHOOD									
GLOBAL									
Fat	Zhang 2012	Cross-sectional	165	USA	FFQ	NA	WBCs	No	4
Fat	Zhang, 2011	Cross-sectional	149	USA	FFQ	NA	Peripheral blood leukocytes	No	4
Fat	Ulrich, 2012	Cross-sectional	173	USA	FFQ	NA	Lymphocytes	Hypo	4
Fat (Lipids)	Gomes, 2012	Cross-sectional	126	Brazil	24h recall o(3 d)	NA	Venous blood	Hypo	3
MUFA	Zhang 2012	Cross-sectional	165	USA	FFQ	NA	WBCs	No	4
MUFA	Zhang, 2011	Cross-sectional	149	USA	FFQ	NA	Peripheral blood leukocytes	No	4
PUFA	Zhang 2012	Cross-sectional	165	USA	FFQ	NA	WBCs	No	4
PUFA	Zhang, 2011	Cross-sectional	149	USA	FFQ	NA	Peripheral blood leukocytes	No	4
SFA	Zhang 2012	Cross-sectional	165	USA	FFQ	NA	WBCs	No	4
SFA	Zhang, 2011	Cross-sectional	149	USA	FFQ	NA	Peripheral blood leukocytes	Hypo	4
GENE-SPECIFIC									

Table 2.1.16. Study characteristics and results: Fat and fatty acids (continued)

Nutritional exposure	First author, year	Study design	N	Country	Exposure assesment	Genes	Tissue	Direction results	QS
Fat (Animal)	Stidley, 2010	Cross-sectional	1101	USA	FFQ	Index (p16, MGMT, DAPK, RASSF1A, PAX5α, PAX5β, GATA4, and GATA5)	Sputum	No	6
Fat	Zhang 2012	Cross-sectional	165	USA	FFQ	IL-6	WBCs	No	5
Fatty acids (circulating ALA)	Ma, 2016 - AJCN	Cross-sectional	991	USA	RBC	APOE	$CD4^+$ T cells isolated from buffy coats	NA	7
Fatty acids (circulating EPA)	Ma, 2016 - AJCN	Cross-sectional	991	USA	RBC	ABCA1	$CD4^+$ T cells isolated from buffy coats	NA	7
Fat (Lipids)	Bollati, 2014	Cross-sectional	165	Italy	FFQ	CD14	Blood	No	6
Fat (Lipids)	Bollati, 2014	Cross-sectional	165	Italy	FFQ	Et-1	Blood	No	6
Fat (Lipids)	Bollati, 2014	Cross-sectional	165	Italy	FFQ	HERV-w	Blood	No	6
Fat (Lipids)	Bollati, 2014	Cross-sectional	165	Italy	FFQ	iNOS	Blood	No	6
Fat (Lipids)	Bollati, 2014	Cross-sectional	165	Italy	FFQ	TNFα	Blood	No	6
Fat (Lipids)	Gómez-Uriz, 2014	Cross-sectional	12	Spain	FFQ	TNF-α	Whole blood	Hypo	4
MUFA	Zhang 2012	Cross-sectional	165	USA	FFQ	IL-6	WBCs	No	5
MUFA	Bollati, 2014	Cross-sectional	165	Italy	FFQ	CD14	Blood	No	6
MUFA	Bollati, 2014	Cross-sectional	165	Italy	FFQ	Et-1	Blood	No	6
MUFA	Bollati, 2014	Cross-sectional	165	Italy	FFQ	HERV-w	Blood	No	6
MUFA	Bollati, 2014	Cross-sectional	165	Italy	FFQ	iNOS	Blood	No	6
MUFA	Bollati, 2014	Cross-sectional	165	Italy	FFQ	TNFα	Blood	No	6

Table 2.1.16. Study characteristics and results: Fat and fatty acids (continued)

Nutritional exposure	First author, year	Study design	N	Country	Exposure assesment	Genes	Tissue	Direction results	QS
MUFA	Milagro, 2012	Cross-sectional	60	Spain	24h recall	CLOCK	WBCs	Hypo	4
PUFA	Zhang 2012	Cross-sectional	165	USA	FFQ	IL-6	WBCs	No	5
PUFA	Bollati, 2014	Cross-sectional	165	Italy	FFQ	CD14	Blood	No	6
PUFA	Bollati, 2014	Cross-sectional	165	Italy	FFQ	Et-1	Blood	No	6
PUFA	Bollati, 2014	Cross-sectional	165	Italy	FFQ	HERV-w	Blood	No	6
PUFA	Bollati, 2014	Cross-sectional	165	Italy	FFQ	iNOS	Blood	No	6
PUFA	Bollati, 2014	Cross-sectional	165	Italy	FFQ	TNFα	Blood	No	6
PUFA	Milagro, 2012	Cross-sectional	60	Spain	24h recall	CLOCK	WBCs	Hyper	4
PUFA n-3	Bollati, 2014	Cross-sectional	165	Italy	FFQ	CD14	Blood	No	6
PUFA n-3	Bollati, 2014	Cross-sectional	165	Italy	FFQ	Et-1	Blood	No	6
PUFA n-3	Bollati, 2014	Cross-sectional	165	Italy	FFQ	HERV-w	Blood	No	6
PUFA n-3	Bollati, 2014	Cross-sectional	165	Italy	FFQ	iNOS	Blood	No	6
PUFA n-3	Bollati, 2014	Cross-sectional	165	Italy	FFQ	TNFα	Blood	No	6
PUFA n-3	Ma, 2016 - MNFR	Cross-sectional	848	USA	RBC	IL6	$CD4^+$ T cells isolated from buffy coats	Hypo	7
PUFA n-6	Hermsdorff, 2013	Cross-sectional	40	Spain	FFQ	TNFα	Blood	Hypo	3
Fat (Saturated)	Zhang 2012	Cross-sectional	165	USA	FFQ	IL-6	WBCs	No	5
SFA	Bollati, 2014	Cross-sectional	165	Italy	FFQ	CD14	Blood	No	6
SFA	Bollati, 2014	Cross-sectional	165	Italy	FFQ	Et-1	Blood	No	6
SFA	Bollati, 2014	Cross-sectional	165	Italy	FFQ	HERV-w	Blood	No	6
SFA	Bollati, 2014	Cross-sectional	165	Italy	FFQ	iNOS	Blood	No	6
SFA	Bollati, 2014	Cross-sectional	165	Italy	FFQ	TNFα	Blood	No	6

Table 2.1.16. Study characteristics and results: Fat and fatty acids (continued)

Nutritional exposure	First author, year	Study design	N	Country	Exposure assesment	Genes	Tissue	Direction results	QS
Fat	Stidley, 2010	Cross-sectional	1101	USA	FFQ	Index (p16, MGMT, DAPK, RASSF1A, PAX5α, PAX5β, GATA4, and GATA5)	Sputum	No	6
Fat	Brøns, 2010	Intervention (non-blinded)	46	Denmark	Diet (3-d control diet including 30% fat & 5-d high-fat overfeeding diet containing 50% extra calories & 60% fat)	PPARGC1A	Muscle biopsie	Hypo	6
Fat	Gillberg, 2014	Intervention (non-blinded)	45	Denmark	Diet (5-d high-fat overfeeding diet (60E% fat, 50% extra calories) and a weight maintaining control diet (30E% fat); 6-8 wk of wash-out)	PPARGC1A	Subcutaneous adipose tissue	Hyper	6
GENOME-WIDE									
PUFA n-3	Aslibekyan, 2014	Cross-sectional	185	USA	RBCs	NA	Blood	Hyper/ Hypo	6

Table 2.1.16. Study characteristics and results: Fat and fatty acids (continued)

Nutritional exposure	First author, year	Study design	N	Country	Exposure assesment	Genes	Tissue	Direction results	QS
Fat	Jacobsen, 2012	Intervention (non-blinded)	21	Denmark	Diet (Restricted high fat diet for 5 d)	NA	Skeletal Muscle biopsies	Hyper/ Hypo	6
Fat	Jacobsen, 2014	Intervention (non-blinded)	40	Denmark	Diet (Restricted high fat diet for 5 d)	NA	Skeletal Muscle biopsies	Hyper/ Hypo	6
Fat	Gillberg, 2016	Intervention (non-blinded)	34	Denmark	Diet (5-d high-fat overfeeding diet (60E% fat, 50% extra calories) & a weight maintaining control diet (30E% fat); 6–8 wk of wash-out + 7 wk of overfeeding	NA	Subcutaneous adipose tissue	Hyper/ Hypo	6
Fatty acids	Irvin, 2014	Intervention (non-blinded)	1048	USA	Diet (Restricted)	NA	Blood	No	8

Table 2.1.17. Study characteristics and results: Carbohydrates and fiber

Nutritional exposure	First author, year	Study design	N	Country	Exposure assesment	Genes	Tissue	Direction results	QS
MATERNAL OFFSPRING RELATIONSHIP									
GLOBAL									
Carbohydrates	Rerkasem, 2015	Cross-sectional	249	Thailand	FFQ & 24h recall	NA	Venous blood samples in children	No	3
GENE-SPECIFIC									
Carbohydrates	Godfrey, 2011	Cross-sectional	78	UK	FFQ	RXRA	Cord tissue	Hypo	3
Carbohydrates	Godfrey, 2011	Cross-sectional	78	UK	FFQ	eNOS	Cord tissue	No	3
ADULTHOOD									
GLOBAL									
Carbohydrates	Gomes, 2012	Cross-sectional	126	Brazil	24h recall (3 d)	NA	Venous blood	Hypo	3
Carbohydrates	Zhang, 2011	Cross-sectional	149	USA	FFQ	NA	Peripheral blood leukocytes	No	4
Carbohydrates	Zhang 2012	Cross-sectional	165	USA	FFQ	NA	WBCs	No	4
Carbohydrates	Ulrich, 2012	Cross-sectional	173	USA	FFQ	NA	Lymphocytes	No	4
GENE-SPECIFIC									
Carbohydrates	Samblas, 2016	Cross-sectional	61	Spain	Questionnaire	BMAL1	Whole blood	Hyper	4
Carbohydrates	Bollati, 2014	Cross-sectional	165	Italy	FFQ	CD14	Blood	No	6
Carbohydrates	Bollati, 2014	Cross-sectional	165	Italy	FFQ	Et-1	Blood	No	6
Carbohydrates	Bollati, 2014	Cross-sectional	165	Italy	FFQ	HERV-w	Blood	No	6
Carbohydrates	Bollati, 2014	Cross-sectional	165	Italy	FFQ	iNOS	Blood	No	6
Carbohydrates	Bollati, 2014	Cross-sectional	165	Italy	FFQ	TNFα	Blood	No	6
Carbohydrates	Zhang 2012	Cross-sectional	165	USA	FFQ	IL-6	WBCs	No	5
Fiber	Bollati, 2014	Cross-sectional	165	Italy	FFQ	CD14	Blood	No	6
Fiber	Bollati, 2014	Cross-sectional	165	Italy	FFQ	Et-1	Blood	No	6
Fiber	Bollati, 2014	Cross-sectional	165	Italy	FFQ	HERV-w	Blood	No	6
Fiber	Bollati, 2014	Cross-sectional	165	Italy	FFQ	iNOS	Blood	No	6
Fiber	Bollati, 2014	Cross-sectional	165	Italy	FFQ	TNFα	Blood	No	6

Table 2.1.18. Study characteristics and results: Protein and amino acids

Nutritional exposure	First author, year	Study design	N	Country	Exposure assesment	Genes	Tissue	Direction results	QS
MATERNAL OFFSPRING RELATIONSHIP									
GLOBAL									
Protein	Rerkasem, 2015	Cross-sectional	249	Thailand	FFQ & 24h recall	NA	Venous blood samples in children	No	3
GENE-SPECIFIC									
Protein	Godfrey, 2011	Cross-sectional	78	UK	FFQ	RXRA	Cord tissue	No	3
Protein	Godfrey, 2011	Cross-sectional	78	UK	FFQ	eNOS	Cord tissue	No	3
ADULTHOOD									
GLOBAL									
Protein	Gomes, 2012	Cross-sectional	126	Brazil	24h recall (3 d)	NA	Venous blood	No	3
Protein	Zhang 2012	Cross-sectional	165	USA	FFQ	NA	WBCs	No	4
Protein	Zhang, 2011	Cross-sectional	149	USA	FFQ	NA	Peripheral blood leukocytes	No	4
Methionine	Huang, 2012	Cross-sectional	493	USA	FFQ	NA	Blood	No	3
Methionine	Perng, 2014	Cross-sectional	987	USA	FFQ	NA	Blood sample	No	3
Methionine	Zhang 2012	Cross-sectional	165	USA	FFQ	NA	WBCs	No	4
Methionine	Zhang, 2011	Cross-sectional	149	USA	FFQ	NA	Peripheral blood leukocytes	No	4
GENE-SPECIFIC									
Protein	Bollati, 2014	Cross-sectional	165	Italy	FFQ	CD14	Blood	No	6
Protein	Bollati, 2014	Cross-sectional	165	Italy	FFQ	Et-1	Blood	No	6
Protein	Bollati, 2014	Cross-sectional	165	Italy	FFQ	HERV-w	Blood	No	6
Protein	Bollati, 2014	Cross-sectional	165	Italy	FFQ	iNOS	Blood	No	6
Protein	Bollati, 2014	Cross-sectional	165	Italy	FFQ	TNFα	Blood	No	6
Protein	Zhang 2012	Cross-sectional	165	USA	FFQ	IL-6	WBCs	No	5
Methionine	Zhang 2012	Cross-sectional	165	USA	FFQ	IL-6	WBCs	No	5
Methionine	Zhong, 2015	Cross-sectional	573	USA	FFQ	TLR2	Whole blood	No	6

REFERENCES

1. Richmond RC, Timpson NJ, Sorensen TI. Exploring possible epigenetic mediation of early-life environmental exposures on adiposity and obesity development. Int J Epidemiol. 2015;44(4):1191-8.
2. Friso S, Udali S, De Santis D, Choi SW. One-carbon metabolism and epigenetics. Mol Aspects Med. 2017;54:28-36.
3. Milagro FI, Mansego ML, De Miguel C, Martínez JA. Dietary factors, epigenetic modifications and obesity outcomes: Progresses and perspectives. Mol Asp Med. 2013;34(4):782-812.
4. Anderson OS, Sant KE, Dolinoy DC. Nutrition and epigenetics: An interplay of dietary methyl donors, one-carbon metabolism and DNA methylation. J Nutr Biochem. 2012;23(8):853-9.
5. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. Br Med J 2009;339:b2700.
6. Voortman T, Vitezova A, Bramer WM, Ars CL, Bautista PK, Buitrago-Lopez A, et al. Effects of protein intake on blood pressure, insulin sensitivity, and blood lipids in children: a systematic review. Br J Nutr. 2015;113(3):In press.
7. Perng W, Rozek LS, Mora-Plazas M, Duchin O, Marin C, Forero Y, et al. Micronutrient status and global DNA methylation in school-age children. Epigenetics. 2012;7(10):1133-41.
8. Piyathilake CJ, Macaluso M, Alvarez RD, Chen M, Badiga S, Siddiqui NR, et al. A higher degree of LINE-1 methylation in peripheral blood mononuclear cells, a one-carbon nutrient related epigenetic alteration, is associated with a lower risk of developing cervical intraepithelial neoplasia. Nutrition. 2011;27(5):513-9.
9. Bollati V, Favero C, Albetti B, Tarantini L, Moroni A, Byun HM, et al. Nutrients intake is associated with DNA methylation of candidate inflammatory genes in a population of obese subjects. Nutrients. 2014; 6(10):4625-39.
10. Stidley CA, Picchi MA, Leng S, Willink R, Crowell RE, Flores KG, et al. Multivitamins, folate, and green vegetables protect against gene promoter methylation in the aerodigestive tract of smokers. Cancer Res. 2010;70(2):568-74.
11. Marques-Rocha JL, Milagro FI, Mansego ML, Mourão DM, Martínez JA, Bressan J. LINE-1 methylation is positively associated with healthier lifestyle but inversely related to body fat mass in healthy young individuals. Epigenetics. 2016;11(1):49-60.
12. Azzi S, Sas TCJ, Koudou Y, Le Bouc Y, Souberbielle JC, Dargent-Molina P, et al. Degree of methylation of ZAC1 (PLAGL1) is associated with prenatal and post-natal growth in healthy infants of the EDEN mother child cohort. Epigenetics. 2014;9(3):338-45.
13. Figueiredo JC, Grau MV, Wallace K, Levine AJ, Shen L, Hamdan R, et al. Global DNA hypomethylation (LINE-1) in the normal colon and lifestyle characteristics and dietary and genetic factors. Cancer Epidemiol Biomarkers Prev. 2009;18(4):1041-9.
14. Ono H, Iwasaki M, Kuchiba A, Kasuga Y, Yokoyama S, Onuma H, et al. Association of dietary and genetic factors related to one-carbon metabolism with global methylation level of leukocyte DNA. Cancer Sci. 2012;103(12):2159-64.
15. Zhang FF, Morabia A, Carroll J, Gonzalez K, Fulda K, Kaur M, et al. Dietary patterns are associated with levels of global genomic DNA methylation in a cancer-free population. J Nutr. 2011;141(6):1165-71.
16. Zhang FF, Santella RM, Wolff M, Kappil MA, Markowitz SB, Morabia A. White blood cell global methylation and IL-6 promoter methylation in association with diet and lifestyle risk factors in a cancer-free population. Epigenetics. 2012;7(6):606-14.

17. Gomes MVM, Toffoli LV, Arruda DW, Soldera LM, Pelosi GG, Neves-Souza RD, et al. Age-Related Changes in the Global DNA Methylation Profile of Leukocytes Are Linked to Nutrition but Are Not Associated with the MTHFR C677T Genotype or to Functional Capacities. PLoS ONE. 2012;7(12).
18. Huang WY, Su LJ, Hayes RB, Moore LE, Katki HA, Berndt SI, et al. Prospective study of genomic hypomethylation of leukocyte DNA and colorectal cancer risk. Cancer Epidemiol Biomarkers Prev. 2012; 21(11):2014-21.
19. Ulrich CM, Toriola AT, Koepl LM, Sandifer T, Poole EM, Duggan C, et al. Metabolic, hormonal and immunological associations with global DNA methylation among postmenopausal women. Epigenetics. 2012;7(9):1020-8.
20. Friso S, Choi SW, Girelli D, Mason JB. A common mutation in the 5, 10-methylenetetrahydrofolate reductase gene affects genomic DNA methylation through an interaction with folate status. Proceedings of the …. 2002.
21. Hubner U, Geisel J, Kirsch SH, Kruse V, Bodis M, Klein C, et al. Effect of 1 year B and D vitamin supplementation on LINE-1 repetitive element methylation in older subjects. Clin Chem Lab Med. 2013;51(3):649-55.
22. Haggarty P, Hoad G, Campbell DM, Horgan GW, Piyathilake C, McNeill G. Folate in pregnancy and imprinted gene and repeat element methylation in the offspring. Am J Clin Nutr. 2013;97(1):94-9.
23. Boeke CE, Baccarelli A, Kleinman KP, Burris HH, Litonjua AA, Rifas-Shiman SL, et al. Gestational intake of methyl donors and global LINE-1 DNA methylation in maternal and cord blood: Prospective results from a folate-replete population. Epigenetics. 2012;7(3):253-60.
24. Fryer AA, Nafee TM, Ismail KM, Carroll WD, Emes RD, Farrell WE. LINE-1 DNA methylation is inversely correlated with cord plasma homocysteine in man: a preliminary study. Epigenetics. 2009; 4(6):394-8.
25. McKay JA, Groom A, Potter C, Coneyworth LJ, Ford D. Genetic and non-genetic influences during pregnancy on infant global and site specific DNA methylation: role for folate gene variants and vitamin B 12: journals.plos.org; 2012.
26. Chang H, Zhang T, Zhang Z, Bao R, Fu C. Tissue-specific distribution of aberrant DNA methylation associated with maternal low-folate status in human neural tube defects. The Journal of nutritional …. 2011.
27. Park BH, Kim YJ, Park JS, Lee HY, Ha EH, Min JW, et al. Folate and homocysteine levels during pregnancy affect DNA methylation in human placenta. J Prev Med Pub Health. 2005;38(4):437-42.
28. Ba Y, Yu H, Liu F, Geng X, Zhu C, Zhu Q, et al. Relationship of folate, vitamin B12 and methylation of insulin-like growth factor-II in maternal and cord blood. Eur J Clin Nutr. 2011;65(4):480-5.
29. Hoyo C, Daltveit AK, Iversen E, Benjamin-Neelon SE, Fuemmeler B, Schildkraut J, et al. Erythrocyte folate concentrations, CpG methylation at genomically imprinted domains, and birth weight in a multiethnic newborn cohort. Epigenetics. 2014;9(8):1120-30.
30. Steegers-Theunissen RP, Obermann-Borst SA. Periconceptional maternal folic acid use of 400 μg per day is related to increased methylation of the IGF2 gene in the very young child: journals.plos.org; 2009.
31. Hoyo C, Murtha AP, Schildkraut JM, Jirtle R, Demark-Wahnefried W, Forman MR, et al. Methylation variation at IGF2 differentially methylated regions and maternal folic acid use before and during pregnancy. Epigenetics. 2011;6(7):928-36.
32. Van Mil NH, Bouwl-Both MI, Stolk L, Verbiest MMPJ, Hofman A, Jaddoe VWV, et al. Determinants of maternal pregnancy one-carbon metabolism and newborn human DNA methylation profiles. Reproduction. 2014;148(6):581-92.

33. Amarasekera M, Martino D, Ashley S, Harb H, Kesper D, Strickland D, et al. Genome-wide DNA methylation profiling identifies a folate-sensitive region of differential methylation upstream of ZFP57-imprinting regulator in humans. FASEB J. 2014;28(9):4068-76.
34. Joubert BR, Den Dekker HT, Felix JF, Bohlin J, Ligthart S, Beckett E, et al. Maternal plasma folate impacts differential DNA methylation in an epigenome-wide meta-analysis of newborns. Nat Commun. 2016;7.
35. Gonseth S, Roy R, Houseman EA, de Smith AJ, Zhou M, Lee ST, et al. Periconceptional folate consumption is associated with neonatal DNA methylation modifications in neural crest regulatory and cancer development genes. Epigenetics. 2015;10(12):1166-76.
36. Mozhui K, Smith AK, Tylavsky FA. Ancestry Dependent DNA Methylation and Influence of Maternal Nutrition. PLoS ONE. 2015;10(3).
37. Agodi A, Barchitta M, Quattrocchi A, Maugeri A, Canto C, Marchese AE, et al. Low fruit consumption and folate deficiency are associated with LINE-1 hypomethylation in women of a cancer-free population. Genes Nutr. 2015;10(5).
38. Perng W, Villamor E, Shroff MR, Nettleton JA, Pilsner JR, Liu Y, et al. Dietary intake, plasma homocysteine, and repetitive element DNA methylation in the Multi-Ethnic Study of Atherosclerosis (MESA). Nutr Metab Cardiovasc Dis. 2014;24(6):614-22.
39. Protiva P, Mason JB, Liu Z, Hopkins ME, Nelson C, Marshall JR, et al. Altered folate availability modifies the molecular environment of the human colorectum: Implications for colorectal carcinogenesis. Cancer Prev Res. 2011;4(4):530-43.
40. Badiga S, Johanning GL, Macaluso M, Azuero A, Chambers MM, Siddiqui NR, et al. A lower degree of PBMC L1 methylation in women with lower folate status may explain the MTHFR C677T polymorphism associated higher risk of CIN in the US post folic acid fortification era. PLoS ONE. 2014;9(10).
41. Bae S, Ulrich CM, Bailey LB, Malysheva O, Brown EC, Maneval DR, et al. Impact of folic acid fortification on global DNA methylation and one-carbon biomarkers in the Women's Health Initiative Observational Study cohort. Epigenetics. 2014;9(3):396-403.
42. Friso S, Girelli D, Trabetti E, Olivieri O, Guarini P, Pignatti PF, et al. The MTHFR 1298A>C polymorphism and genomic DNA methylation in human lymphocytes. Cancer Epidemiol Biomarkers Prev. 2005;14(4):938-43.
43. Pilsner JR, Liu X, Ahsan H, Ilievski V, Slavkovich V, Levy D, et al. Genomic methylation of peripheral blood leukocyte DNA: Influences of arsenic and folate in Bangladeshi adults. Am J Clin Nutr. 2007; 86(4):1179-86.
44. Piyathilake CJ, Badiga S, Alvarez RD, Partridge EE, Johanning GL. A Lower Degree of PBMC L1 Methylation Is Associated with Excess Body Weight and Higher HOMA-IR in the Presence of Lower Concentrations of Plasma Folate. PLoS ONE. 2013;8(1).
45. Wang TC, Song YS, Wang H, Zhang J, Yu SF, Gu YE, et al. Oxidative DNA damage and global DNA hypomethylation are related to folate deficiency in chromate manufacturing workers. J Hazard Mater. 2012;213-214:440-6.
46. Fenech M, Aitken C, Rinaldi J. Folate, vitamin B12, homocysteine status and DNA damage in young Australian adults. Carcinogenesis. 1998;19(7):1163-71.
47. Gadgil MS, Joshi KS, Naik SS, Pandit AN, Otiv SR, Patwardhan BK. Association of homocysteine with global DNA methylation in vegetarian Indian pregnant women and neonatal birth anthropometrics. J Matern -Fetal Neonatal Med. 2014;27(17):1749-53.
48. Narayanan S, McConnell J, Little J, Sharp L, Piyathilake CJ, Powers H, et al. Associations between two common variants C677T and A1298C in the methylenetetrahydrofolate reductase gene and of measures

of folate metabolism and DNA stability (strand breaks, misincorporated uracil, and DNA methylation status) in human lymphocytes in vivo. Cancer Epidemiol Biomarkers Prev. 2004;13(9):1436-43.

49. Axume J, Smith SS, Pogribny IP, Moriarty DJ, Caudill MA. The methylenetetrahydrofolate reductase 677TT genotype and folate intake interact to lower global leukocyte DNA methylation in young Mexican American women. Nutr Res. 2007;27(1):13-7.
50. Crider KS, Quinlivan EP, Berry RJ, Hao L, Li Z, Maneval D, et al. Genomic DNA methylation changes in response to folic acid supplementation in a population-based intervention study among women of reproductive age. PLoS ONE. 2011;6(12).
51. Jacob RA, Gretz DM, Taylor PC, James SJ. Moderate folate depletion increases plasma homocysteine and decreases lymphocyte DNA methylation in postmenopausal women. … Journal of nutrition. 1998.
52. Shelnutt KP, Kauwell GPA, Gregory JF. Methylenetetrahydrofolate reductase 677C→ T polymorphism affects DNA methylation in response to controlled folate intake in young women. The Journal of nutritional …. 2004.
53. Rampersaud GC, Kauwell GPA, Hutson AD, Cerda JJ, Bailey LB. Genomic DNA methylation decreases in response to moderate folate depletion in elderly women. Am J Clin Nutr. 2000;72(4):998-1003.
54. Jung AY, Smulders Y, Verhoef P, Kok FJ, Blom H, Kok RM, et al. No effect of folic acid supplementation on global DNA methylation in men and women with moderately elevated homocysteine. PLoS ONE. 2011;6(9).
55. Basten GP, Duthie SJ, Pirie L, Vaughan N, Hill MH, Powers HJ. Sensitivity of markers of DNA stability and DNA repair activity to folate supplementation in healthy volunteers. Br J Cancer. 2006;94(12): 1942-7.
56. Hanks J, Ayed I, Kukreja N, Rogers C, Harris J, Gheorghiu A, et al. The association between mthfr 677C>T genotype and folate status and genomic and gene-Specific dna methylation in the colon of individuals without colorectal neoplasia1-4. Am J Clin Nutr. 2013;98(6):1564-74.
57. Pufulete M, Al-Ghnaniem R, Rennie JA, Appleby P, Harris N, Gout S, et al. Influence of folate status on genomic DNA methylation in colonic mucosa of subjects without colorectal adenoma or cancer. Br J Cancer. 2005;92(5):838-42.
58. Tapp HS, Commane DM, Bradburn DM, Arasaradnam R, Mathers JC, Johnson IT, et al. Nutritional factors and gender influence age-related DNA methylation in the human rectal mucosa. Aging Cell. 2013;12(1):148-55.
59. Llanos AAM, Marian C, Brasky TM, Dumitrescu RG, Liu Z, Mason JB, et al. Associations between genetic variation in onecarbon metabolism and LINE-1 DNA methylation in histologically normal breast tissues. Epigenetics. 2015;10(8):727-35.
60. Aarabi M, San Gabrie MC, Chan D, Behan NA, Caron M, Pastinen T, et al. High-dose folic acid supplementation alters the human sperm methylome and is influenced by the MTHFR C677T polymorphism. Hum Mol Genet. 2015;24(22):6301-13.
61. Zhong J, Colicino E, Lin X, Mehta A, Kloog I, Zanobetti A, et al. Cardiac autonomic dysfunction: Particulate air pollution effects are modulated by epigenetic immunoregulation of Toll-like receptor 2 and dietary flavonoid intake. J Am Heart Assoc. 2015;4(1).
62. Dhillon VS, Shahid M, Husain SA. Associations of MTHFR DNMT3b 4977 bp deletion in mtDNA and GSTM1 deletion, and aberrant CpG island hypermethylation of GSTM1 in non-obstructive infertility in Indian men. Mol Hum Reprod. 2007;13(4):213-22.
63. Ottini L, Rizzolo P, Siniscalchi E, Zijno A, Silvestri V, Crebelli R, et al. Gene promoter methylation and DNA repair capacity in monozygotic twins with discordant smoking habits. Mutat Res Genet Toxicol Environ Mutagen. 2015;779:57-64.

64. Piyathilake CJ, Macaluso M, Chambers MM, Badiga S, Siddiqui NR, Bell WC, et al. Folate and vitamin B12 may play a critical role in lowering the HPV 16 methylation - Associated risk of developing higher grades of CIN. Cancer Prev Res. 2014;7(11):1128-37.
65. Llanos AA, Dumitrescu RG, Brasky TM, Liu Z, Mason JB, Marian C, et al. Relationships among folate, alcohol consumption, gene variants in one-carbon metabolism and p16INK4a methylation and expression in healthy breast tissues. Carcinogenesis. 2015;36(1):60-7.
66. Al-Ghnaniem Abbadi R, Emery P, Pufulete M. Short-Term Folate Supplementation in Physiological Doses Has No Effect on ESR1 and MLH1 Methylation in Colonic Mucosa of Individuals with Adenoma. J Nutrigenet Nutrigenomics. 2013;5(6):327-38.
67. Song MA, Brasky TM, Marian C, Weng DY, Taslim C, Llanos AA, et al. Genetic variation in one-carbon metabolism in relation to genome-wide DNA methylation in breast tissue from healthy women. Carcinogenesis. 2016;37(5):471-80.
68. Quinlivan EP, Crider KS, Zhu JH, Maneval DR, Hao L, Li Z, et al. Hypomethylation of Serum Blood Clot DNA, but Not Plasma EDTA-Blood Cell Pellet DNA, from Vitamin B12-Deficient Subjects. PLoS ONE. 2013;8(6).
69. Irvin MR, Zhi D, Aslibekyan S, Claas SA, Absher DM, Ordovas JM, et al. Genomics of post-prandial lipidomic phenotypes in the Genetics of Lipid Lowering Drugs and Diet Network (GOLDN) study. PLoS ONE. 2014;9(6).
70. Al-Ghnaniem R, Peters J, Foresti R, Heaton N, Pufulete M. Methylation of estrogen receptor α and mutL homolog 1 in normal colonic mucosa: Association with folate and vitamin B-12 status in subjects with and without colorectal neoplasia. Am J Clin Nutr. 2007;86(4):1064-72.
71. de la Iglesia R, Mansego ML, Sánchez-Muniz FJ, Angeles Zulet M, Alfredo Martinez J. Arylesterase activity is associated with antioxidant intake and Paraoxonase-1 (PON1) gene methylation in metabolic syndrome patients following an energy restricted diet. EXCLI J. 2014;13:416-26.
72. Harvey NC, Sheppard A, Godfrey KM, McLean C, Garratt E, Ntani G, et al. Childhood bone mineral content is associated with methylation status of the RXRA promoter at birth. J Bone Miner Res. 2014; 29(3):600-7.
73. Novakovic B, Galati JC, Chen A, Morley R, Craig JM, Saffery R. Maternal vitamin D predominates over genetic factors in determining neonatal circulating vitamin D concentrations. Am J Clin Nutr. 2012; 96(1):188-95.
74. Zhu H, Bhagatwala J, Huang Y, Pollock NK, Parikh S, Raed A, et al. Race/ethnicity-specific association of Vitamin D and global DNA methylation: Cross-sectional and interventional findings. PLoS ONE. 2016;11(4).
75. Zhu H, Wang X, Shi H, Su S, Harshfield GA, Gutin B, et al. A genome-wide methylation study of severe vitamin d deficiency in African American adolescents. J Pediatr. 2013;162(5):1004-9.e1.
76. Nair-Shalliker V, Dhillon V, Clements M, Armstrong BK, Fenech M. The association between personal sun exposure, serum vitamin D and global methylation in human lymphocytes in a population of healthy adults in South Australia. Mutat Res Fundam Mol Mech Mutagen. 2014;765:6-10.
77. Ashktorab H, Nguza B, Fatemi M, Nouraie M, Smoot DT, Schäffer AA, et al. Case-control study of vitamin D, dickkopf homolog 1 (DKK1) gene methylation, VDR gene polymorphism and the risk of colon adenoma in African Americans. PLoS ONE. 2011;6(10).
78. Jiang X, Yan J, West AA, Perry CA, Malysheva OV, Devapatla S, et al. Maternal choline intake alters the epigenetic state of fetal cortisol-regulating genes in humans. FASEB J. 2012;26(8):3563-74.
79. Shin W, Yan J, Abratte CM, Vermeylen F, Caudill MA. Choline intake exceeding current dietary recommendations preserves markers of cellular methylation in a genetic subgroup of folate-compromised men. J Nutr. 2010;140(5):975-80.

80. McClelland R, Christensen K, Mohammed S, McGuinness D, Cooney J, Bakshi A, et al. Accelerated ageing and renal dysfunction links lower socioeconomic status and dietary phosphate intake. Aging (Albany NY). 2016.
81. Shimazu T, Asada K, Charvat H, Kusano C, Otake Y, Kakugawa Y, et al. Association of gastric cancer risk factors with DNA methylation levels in gastric mucosa of healthy Japanese: a cross-sectional study. Carcinogenesis. 2015;36(11):1291-8.
82. Khulan B, Cooper WN, Skinner BM, Bauer J, Owens S, Prentice AM, et al. Periconceptional maternal micronutrient supplementation is associated with widespread gender related changes in the epigenome: A study of a unique resource in the Gambia. Hum Mol Genet. 2012;21(9):2086-101.
83. Pusceddu I, Herrmann M, Kirsch SH, Werner C, Hübner U, Bodis M, et al. Prospective study of telomere length and LINE-1 methylation in peripheral blood cells: the role of B vitamins supplementation. Eur J Nutr. 2015.
84. Zhou Y, Zhao LJ, Xu X, Ye A, Travers-Gustafson D, Zhou B, et al. DNA methylation levels of CYP2R1 and CYP24A1 predict vitamin D response variation. J Steroid Biochem Mol Biol. 2014;144(PART A): 207-14.
85. Kok DEG, Dhonukshe-Rutten RAM, Lute C, Heil SG, Uitterlinden AG, van der Velde N, et al. The effects of long-term daily folic acid and vitamin B-12 supplementation on genome-wide DNA methylation in elderly subjects. Clinical Epigenetics. 2015;7.
86. Rerkasem K, Rattanatanyong P, Rerkasem A, Wongthanee A, Rungruengthanakit K, Mangklabruks A, et al. Higher Alu methylation levels in catch-up growth in twenty-year-old offsprings. PLoS ONE. 2015; 10(3).
87. Godfrey KM, Sheppard A, Gluckman PD, Lillycrop KA, Burdge GC, McLean C, et al. Epigenetic gene promoter methylation at birth is associated with child's later adiposity. Diabetes. 2011;60(5):1528-34.
88. Lee HS, Barraza-Villarreal A, Hernandez-Vargas H, Sly PD, Biessy C, Ramakrishnan U, et al. Modulation of DNA methylation states and infant immune system by dietary supplementation with v-3 PUFA during pregnancy in an intervention study. Am J Clin Nutr. 2013;98(2):480-7.
89. Amarasekera M, Noakes P, Strickland D, Saffery R, Martino DJ, Prescott SL. Epigenome-wide analysis of neonatal CD4+ T-cell DNA methylation sites potentially affected by maternal fish oil supplementation. Epigenetics. 2014;9(12):1570-6.
90. Voisin S, Almén MS, Moschonis G, Chrousos GP, Manios Y, Schiöth HB. Dietary fat quality impacts genome-wide DNA methylation patterns in a cross-sectional study of Greek preadolescents. Eur J Hum Genet. 2015;23(5):654-62.
91. Lind MV, Martino D, Harsløf LBS, Kyjovska ZO, Kristensen M, Lauritzen L. Genome-wide identification of mononuclear cell DNA methylation sites potentially affected by fish oil supplementation in young infants: A pilot study. Prostaglandins Leukotrienes Essent Fatty Acids. 2015;101:1-7.
92. Gómez-Uriz AM, Goyenechea E, Campión J, De Arce A, Martinez MT, Puchau B, et al. Epigenetic patterns of two gene promoters (TNF-α and PON) in stroke considering obesity condition and dietary intake. J Physiol Biochem. 2014;70(2):603-14.
93. Brøns C, Jacobsen S, Nilsson E, Rönn T, Jensen CB, Storgaard H, et al. Deoxyribonucleic acid methylation and gene expression of PPARGC1A in human muscle is influenced by high-fat overfeeding in a birth-weight-dependent manner. J Clin Endocrinol Metab. 2010;95(6):3048-56.
94. Gillberg L, Jacobsen SC, Rönn T, Brøns C, Vaag A. PPARGC1A DNA methylation in subcutaneous adipose tissue in low birth weight subjects - Impact of 5 days of high-fat overfeeding. Metab Clin Exp. 2014;63(2):263-71.

95. Gillberg L, Perfilyev A, Brøns C, Thomasen M, Grunnet LG, Volkov P, et al. Adipose tissue transcriptomics and epigenomics in low birthweight men and controls: role of high-fat overfeeding. Diabetologia. 2016;59(4):799-812.
96. Jacobsen SC, Gillberg L, Bork-Jensen J, Ribel-Madsen R, Lara E, Calvanese V, et al. Young men with low birthweight exhibit decreased plasticity of genome-wide muscle DNA methylation by high-fat overfeeding. Diabetologia. 2014;57(6):1154-8.
97. Hermsdorff HH, Mansego ML, Campión J, Milagro FI, Zulet MA, Martínez JA. TNF-alpha promoter methylation in peripheral white blood cells: Relationship with circulating TNFα, truncal fat and n-6 PUFA intake in young women. Cytokine. 2013;64(1):265-71.
98. Milagro FI, Gómez-Abellán P, Campión J, Martínez JA, Ordovás JM, Garaulet M. CLOCK, PER2 and BMAL1 DNA methylation: Association with obesity and metabolic syndrome characteristics and monounsaturated fat intake. Chronobiol Int. 2012;29(9):1180-94.
99. Ma Y, Smith CE, Lai CQ, Irvin MR, Parnell LD, Lee YC, et al. The effects of omega-3 polyunsaturated fatty acids and genetic variants on methylation levels of the interleukin-6 gene promoter. Mol Nutr Food Res. 2016;60(2):410-9.
100. Ma Y, Follis JL, Smith CE, Tanaka T, Manichaikul AW, Chu AY, et al. Interaction of methylation-related genetic variants with circulating fatty acids on plasma lipids: A meta-analysis of 7 studies and methylation analysis of 3 studies in the Cohorts for Heart and Aging Research in Genomic Epidemiology consortium. Am J Clin Nutr. 2016;103(2):567-78.
101. Samblas M, Milagro FI, Gomez-Abellan P, Martinez JA, Garaulet M. Methylation on the Circadian Gene BMAL1 Is Associated with the Effects of a Weight Loss Intervention on Serum Lipid Levels. J Biol Rhythms. 2016.
102. Peters MJ, Joehanes R, Pilling LC, Schurmann C, Conneely KN, Powell J, et al. The transcriptional landscape of age in human peripheral blood. Nat Commun. 2015;6:8570.
103. Lee HS, Barraza-Villarreal A, Biessy C, Duarte-Salles T, Sly PD, Ramakrishnan U, et al. Dietary supplementation with polyunsaturated fatty acid during pregnancy modulates DNA methylation at IGF2/H19 imprinted genes and growth of infants. Physiol Genomics. 2014;46(23):851-7.
104. Piyathilake CJ, Robinson CB, Cornwell P. A practical approach to red blood cell folate analysis. Anal Chem Insights. 2007;2:107-10.
105. De Bruyn E, Gulbis B, Cotton F. Serum and red blood cell folate testing for folate deficiency: new features? Eur J Haematol. 2014;92(4):354-9.

SUPPLEMENTARY MATERIAL

Supplement 2.1.1. Details on the search strategy for each of the databases

	References identified	Unique references identified
Embase.com	2337	2294
Medline Ovid	1547	249
Web of science	1714	943
Cochrane	87	19
PubMed publisher	149	133
Google scholar	200	136
Total	**5874**	**3774**

Embase.com 2337

(diet/exp OR 'dietary intake'/exp OR 'food intake'/de OR 'caloric density'/exp OR 'caloric intake'/exp OR 'carbohydrate intake'/exp OR 'diet restriction'/exp OR 'dietary reference intake'/exp OR 'electrolyte intake'/exp OR 'fat intake'/exp OR 'fluid intake'/exp OR 'food deprivation'/exp OR 'mineral intake'/exp OR 'protein intake'/exp OR 'vitamin intake'/exp OR 'vitamin blood level'/exp OR 'diet supplementation'/exp OR 'mineral supplementation'/exp OR 'vitamin supplementation'/exp OR supplementation/de OR 'nutrition'/de OR 'child nutrition'/de OR 'diet therapy'/exp OR 'food intake'/exp OR 'infant nutrition'/exp OR 'maternal nutrition'/de OR nutrient/exp OR 'nutritional status'/exp OR (diet OR nutrition OR ((dietar* OR food* OR beverage* OR alcohol* OR mineral* OR nutrient* OR micronutrient* OR macronutrient* OR vitamin* OR calor* OR energ* OR protein* OR fat OR folic-acid* OR folate* OR vegetabl* OR fruit* OR cholin* OR methionin* OR betain* OR fatty-acid* OR pufa OR pufas OR mufa OR mufas OR safa OR safas OR sfa OR sfas) NEAR/6 (intake* OR ingest* OR supplement* OR consum* OR restrict* OR depriv* OR level OR concentration OR blood OR plasma OR serum OR pattern*)) OR fasting OR ((well OR under) NEXT/1 nourish*) OR ((well OR under) NEXT/1 nourish*)):ab,ti) AND ('DNA methylation'/exp OR (((dna OR 'long interspersed' OR gene OR genes) NEAR/6 (demethylat* OR methylat* OR hypermethylat* OR hypomethylat*))):ab,ti) NOT ([animals]/lim NOT [humans]/lim) AND ('observational study'/exp OR 'cohort analysis'/exp OR 'longitudinal study'/exp OR 'retrospective study'/exp OR 'prospective study'/exp OR 'health survey'/de OR 'health care survey'/de OR 'epidemiological data'/de OR 'case control study'/de OR 'cross-sectional study'/de OR 'correlational study'/de OR 'population research'/de OR 'family study'/de OR 'major clinical study'/de OR 'multicenter study'/de OR 'comparative study'/de OR 'follow up'/de OR 'clinical study'/de OR 'clinical article'/de OR 'clinical trial'/exp OR 'randomization'/exp OR 'intervention study'/de OR 'open study'/de OR 'community trial'/de OR 'review'/exp OR 'systematic review'/exp OR (((observation* OR epidemiolog* OR famil* OR comparativ* OR communit*) NEAR/6 (stud* OR data OR research)) OR cohort* OR longitudinal* OR retrospectiv* OR prospectiv* OR population* OR (national* NEAR/3 (stud* OR survey)) OR (health* NEAR/3 survey*) OR ((case OR cases OR match*) NEAR/3 control*) OR (cross NEXT/1 section*) OR correlation* OR multicenter* OR multi-center* OR follow-up* OR followup* OR clinical* OR trial OR random* OR review* OR meta-analy*):ab,ti) NOT ([Conference Abstract]/lim OR [Letter]/lim OR [Note]/lim OR [Editorial]/lim)

Medline Ovid 1547

(exp Diet/ OR exp "Eating"/ OR "Recommended Dietary Allowances"/ OR "food deprivation"/ OR vitamins/bl OR "Dietary Supplements"/ OR "Nutritional Physiological Phenomena"/ OR exp "Diet Therapy"/ OR "Maternal Nutritional Physiological Phenomena"/ OR "nutritional status"/ OR (diet OR nutrition OR ((dietar* OR food* OR beverage* OR alcohol* OR mineral* OR nutrient* OR micronutrient* OR macronutrient* OR vitamin* OR calor* OR energ* OR protein* OR

fat OR folic-acid* OR folate* OR vegetabl* OR fruit* OR cholin* OR methionin* OR betain* OR fatty-acid* OR pufa OR pufas OR mufa OR mufas OR safa OR safas OR sfa OR sfas) ADJ6 (intake* OR ingest* OR supplement* OR consum* OR restrict* OR depriv* OR level OR concentration OR blood OR plasma OR serum OR pattern*)) OR fasting OR ((well OR under) ADJ nourish*) OR ((well OR under) ADJ nourish*)).ab,ti.) AND ("DNA Methylation"/ OR (((dna OR "long interspersed" OR gene OR genes) ADJ6 (demethylat* OR methylat* OR hypermethylat* OR hypomethylat*))).ab,ti.) NOT (exp animals/ NOT humans/) AND ("observational study"/ OR exp "Epidemiologic Studies"/ OR "health surveys"/ OR "multicenter study"/ OR exp "clinical study"/ OR "Random Allocation"/ OR "review"/ OR (((observation* OR epidemiolog* OR famil* OR comparativ* OR communit*) ADJ6 (stud* OR data OR research)) OR cohort* OR longitudinal* OR retrospectiv* OR prospectiv* OR population* OR (national* ADJ3 (stud* OR survey)) OR (health* ADJ3 survey*) OR ((case OR cases OR match*) ADJ3 control*) OR (cross ADJ section*) OR correlation* OR multicenter* OR multi-center* OR follow-up* OR followup* OR clinical* OR trial OR random* OR review* OR meta-analy*).ab,ti.) NOT (letter OR news OR comment OR editorial OR congresses OR abstracts).pt.

Cochrane 87

((diet OR nutrition OR ((dietar* OR food* OR beverage* OR alcohol* OR mineral* OR nutrient* OR micronutrient* OR macronutrient* OR vitamin* OR calor* OR energ* OR protein* OR fat OR folic-acid* OR folate* OR vegetabl* OR fruit* OR cholin* OR methionin* OR betain* OR fatty-acid* OR pufa OR pufas OR mufa OR mufas OR safa OR safas OR sfa OR sfas) NEAR/6 (intake* OR ingest* OR supplement* OR consum* OR restrict* OR depriv* OR level OR concentration OR blood OR plasma OR serum OR pattern*)) OR fasting OR ((well OR under) NEXT/1 nourish*) OR ((well OR under) NEXT/1 nourish*)):ab,ti) AND ((((dna OR 'long interspersed' OR gene OR genes) NEAR/6 (demethylat* OR methylat* OR hypermethylat* OR hypomethylat*))):ab,ti)

Web of science 1714

TS=(((diet OR nutrition OR ((dietar* OR food* OR beverage* OR alcohol* OR mineral* OR nutrient* OR micronutrient* OR macronutrient* OR vitamin* OR calor* OR energ* OR protein* OR fat OR folic-acid* OR folate* OR vegetabl* OR fruit* OR cholin* OR methionin* OR betain* OR fatty-acid* OR pufa OR pufas OR mufa OR mufas OR safa OR safas OR sfa OR sfas) NEAR/5 (intake* OR ingest* OR supplement* OR consum* OR restrict* OR depriv* OR level OR concentration OR blood OR plasma OR serum OR pattern*)) OR fasting OR ((well OR under) NEAR/1 nourish*) OR ((well OR under) NEAR/1 nourish*))) AND ((((dna OR "long interspersed" OR gene OR genes) NEAR/5 (demethylat* OR methylat* OR hypermethylat* OR hypomethylat*)))) AND ((((observation* OR epidemiolog* OR famil* OR comparativ* OR communit*) NEAR/5 (stud* OR data OR research)) OR cohort* OR longitudinal* OR retrospectiv* OR prospectiv* OR population* OR (national* NEAR/2 (stud* OR survey)) OR (health* NEAR/2 survey*) OR ((case OR cases OR match*) NEAR/2 control*) OR (cross NEAR/1 section*) OR correlation* OR multicenter* OR multi-center* OR follow-up* OR followup* OR clinical* OR trial OR random* OR review* OR meta-analy*)) NOT ((animal* OR rat OR rats OR mouse OR mice OR murine) NOT (human* OR patient*))) AND DT=(article)

PubMed publisher 149

(Diet[mh] OR "Eating"[mh] OR "Recommended Dietary Allowances"[mh] OR "food deprivation"[mh] OR vitamins/bl[mh] OR "Dietary Supplements"[mh] OR "Nutritional Physiological Phenomena"[mh] OR "Diet Therapy"[mh] OR "Maternal Nutritional Physiological Phenomena"[mh] OR "nutritional status"[mh] OR (diet OR nutrition OR ((dietar*[tiab] OR food*[tiab] OR beverage*[tiab] OR alcohol*[tiab] OR mineral*[tiab] OR nutrient*[tiab] OR micronutrient*[tiab] OR macronutrient*[tiab] OR vitamin*[tiab] OR calor*[tiab] OR energ*[tiab] OR protein*[tiab] OR fat OR folic-acid*[tiab] OR folate*[tiab] OR vegetabl*[tiab] OR fruit*[tiab]

OR cholin*[tiab] OR methionin*[tiab] OR betain*[tiab] OR fatty-acid*[tiab] OR pufa OR pufas OR mufa OR mufas OR safa OR safas OR sfa OR sfas) AND (intake*[tiab] OR ingest*[tiab] OR supplement*[tiab] OR consum*[tiab] OR restrict*[tiab] OR depriv*[tiab] OR level OR concentration OR blood OR plasma OR serum OR pattern*[tiab])) OR fasting OR well nourish*[tiab] OR under nourish*[tiab])) AND ("DNA Methylation"[mh] OR (((dna OR "long interspersed" OR gene OR genes) AND (demethylat*[tiab] OR methylat*[tiab] OR hypermethylat*[tiab] OR hypomethylat*[tiab])))) NOT (animals[mh] NOT humans[mh]) AND ("observational study"[mh] OR "Epidemiologic Studies"[mh] OR "health surveys"[mh] OR "multicenter study"[mh] OR "clinical study"[mh] OR "Random Allocation"[mh] OR "review"[mh] OR (((observation*[tiab] OR epidemiolog*[tiab] OR famil*[tiab] OR comparativ*[tiab] OR communit*[tiab]) AND (stud*[tiab] OR data OR research)) OR cohort*[tiab] OR longitudinal*[tiab] OR retrospectiv*[tiab] OR prospectiv*[tiab] OR population*[tiab] OR (national*[tiab] AND (stud*[tiab] OR survey)) OR (health*[tiab] AND survey*[tiab]) OR ((case OR cases OR match*[tiab]) AND control*[tiab]) OR (cross section*[tiab]) OR correlation*[tiab] OR multicenter*[tiab] OR multi-center*[tiab] OR follow-up*[tiab] OR followup*[tiab] OR clinical*[tiab] OR trial OR random*[tiab] OR review*[tiab] OR meta-analy*[tiab])) NOT (letter[pt] OR news[pt] OR comment[pt] OR editorial[pt] OR congresses[pt] OR abstracts[pt]) AND publisher[sb]

Google scholar

diet|nutrition|dietary|nutrients|micronutrients|macronutrients|vitamins|"folic acid"|folate|vegetables|fruit|choline "dna demethylation|methylation|hypermethylation|hypomethylation"

Supplement 2.1.2. Detailed selection criteria

Inclusion criteria	1. Any study design → Including cross-sectional studies, case-control studies, cohort studies or intervention studies. 2. Studies investigating associations of dietary factors with DNA methylation. → Nutrients: Including nutrient levels (in blood or other tissues); intake of nutrients; or dietary supplements. → DNA methylation: Including global DNA methylation, gene specific DNA methylation, or genome-wide DNA methylation. 3. Studies conducted at any age or life stage (e.g., during pregnancy, childhood, adulthood, etc.) 4. No language or date restriction
Exclusion criteria	1. Studies not carried out in humans 2. In vitro studies 3. Studies conducted only in patients with chronic diseases (cancer, diabetes, CHD, etc.) Note: some studies have healthy control groups and report associations within these groups separately, these should be included. 4. Letters, reviews, opinions papers, guidelines, case-reports (n<5), editorials 5. Studies examining only the following exposures: a. eating disorders; b. alcohol dependency/alcoholism ; c. weight loss (unless the weight-loss intervention is solely a restricted diet, not combined with other lifestyle changes). d. Dietary patterns or food groups without examining individual nutrients

Supplement 2.1.3. Quality score

Two independent reviewers evaluated the quality of included studies using the following quality score. The score is composed of 5 items, and each item was allocated 0, 1 or 2 points. This allowed a total score between 0 and 10 points, 10 representing the highest quality. The following items are included in the score:

1. Study design

0 for cross-sectional studies
1 for longitudinal studies (including repeated measurements)
2 for interventional studies

2. Population

Global and candidate gene studies

0 if *n*<30
1 if *n* 30 to 100
2 if *n*>100

EWAS studies

0 if *n*<300
1 if *n* 300 to 1000
2 if *n*>1000

3. Exposure

Observational studies

0 if the study used no appropriate standard diet or nutrient concentration assessment method (see below) or if not reported
1 if the study used food records, a 24h recall or an FFQ
2 if the study used blood levels

Interventional studies

0 if the intervention diet was not described or not blinded
1 if the intervention diet was adequately single-blinded
2 if the intervention is adequately double-blinded

4. **Outcome**

0 Global DNA methylation
1 Gene-specific DNA methylation
2 Genome-wide DNA methylation

5. Adjustments

0 if findings are not controlled for potential confounders*
1 if findings are controlled for at least basic confounders: age and sex, and multiple testing in case of an EWAS
2 if an intervention study is adequately randomized or if findings are controlled for additional confounders (on top of those described above), such as: technical covariates, WBC, BMI, smoking, alcohol, etc.

** 'Controlled for' includes: adjustment for in the statistical analyses (e.g. with multivariable regression); stratification in the analyses (e.g. men and women separately); or restriction or narrow selection criteria of study participants on this covariate.*

Chapter 3

DNA methylation & cardiometabolic health

3.1

The role of DNA methylation in dyslipidaemia: A systematic review.

Braun KVE*, Voortman T*, Dhana K, Troup J, Bramer WM, Troup J, Chowdhury R, Dehghan A, Muka T, Franco OH. The role of DNA methylation in dyslipidaemia: A systematic review. *Progress in Lipid Research. 2016;64:178-91.*

**Denotes equal contribution*

ABSTRACT

Epigenetic mechanisms, including DNA methylation and histone modifications might be involved in the regulation of lipid concentration variability and may thereby contribute to the cardiovascular risk profile. We aimed to systematically review studies investigating the association between epigenetic marks and plasma concentrations of triacylglycerol, total cholesterol, low-density lipoprotein-cholesterol, and high-density lipoprotein-cholesterol. Six medical databases were searched until September 3rd 2015, reference lists were screened, and experts in the field were contacted. Of the initially identified 757 references, 31 articles reporting on 23 unique studies met all our inclusion criteria. These studies included data on 8,027 unique participants. Overall, no consistent associations were observed between global DNA methylation and blood lipids. Candidate gene and epigenome-wide association studies reported epigenetic regulation of several genes to be related with blood lipids, of which results for *ABCG1*, *CPT1A*, *TNNT1*, *MIR33B*, *SREBF1*, and *TNIP* were replicated. To date, no studies have been performed on histone modification in relation to blood lipids. Promising results have been reported in the field of epigenetics and dyslipidaemia, however, further rigorous studies are needed to expand our understanding on the role of epigenetics in regulating human's blood lipid levels and its effects on health and disease.

INTRODUCTION

Dyslipidaemia, defined as decreased high-density lipoprotein-cholesterol (HDL-C), elevated low-density lipoprotein-cholesterol (LDL-C), and/or elevated triacylglycerol (TAG) concentrations,[1, 2] is recognized as a prominent risk factor for cardiovascular disease (CVD)[3] and is associated with many other metabolic disturbances.[4]

Dyslipidaemia is caused by environmental factors, such as an unhealthy diet and lack of exercise.[5] In addition, genetic factors have a strong influence on blood lipid concentrations and lipid metabolism.[6] Genome wide-association studies have identified several common genetic variants that explain up to 12% of the inter-individual variability in plasma blood lipids.[7] Beyond this, there is evidence that epigenetic mechanisms, that have been recognized recently as a potential important link between environmental exposure and disease risk,[8, 9] may account for lipoprotein profile variability and may thereby influence risk of CVD.[10-12] Epigenetics refers to changes in gene activity and expression without alterations in DNA sequence. Epigenetic changes occur during development and differentiation, or in response to environmental factors, such as smoking, diet and physical activity. The best understood epigenetic mechanisms that regulate these non-genetic alterations are attachment of a methyl group to the cytosine-guanine dinucleotide (CpG) of DNA (DNA methylation) and changes in histone proteins associated with DNA (histone modifications). These epigenetic processes occur under the influence of environmental factors and influence packaging and interpreting the genome and thus altering gene expression, and may subsequently affect disease risk.[13] Epigenetics is therefore considered a bridge between genotype and phenotype.

In epidemiological studies, DNA methylation is studied as either methylation at single genes, or as global DNA methylation. The latter refers to global analysis of DNA methylation across the entire genome in a given cell. Global DNA methylation in blood cells has been used as epigenetic biomarkers for predicting disease risk, including cardiovascular disease and type 2 diabetes.[14, 15] In addition to global methylation, previous studies have reported that DNA methylation at specific genes are related to blood lipid concentrations, including methylation at genes that are known to be involved in dyslipidaemia, such as *APOE*, *LPL* and *ABCG1*.[14, 16, 17] It has been suggested that epigenetics is a promising area in science that can help to obtain a better understanding of the aetiology of dyslipidaemia, as well as a promising molecular strategy for disease risk stratification. To date, however, there is no comprehensive assessment of the current evidence for the role of DNA methylation and histone modifications on plasma lipids profile.

Therefore, we conducted a systematic review of all available evidence in humans on the association of DNA methylation and histone modifications with plasma concentrations of triacylglycerol (TAG), total cholesterol (TC), low-density lipoprotein-cholesterol (LDL-C), and high-density lipoprotein-cholesterol (HDL-C). A critical appraisal of current limitations in the field is also presented.

METHODS

Literature search

The search engines Embase.com (Medline and Embase), Medline (Ovid), Web-of-Science, PubMed, Cochrane Central and Google Scholar were searched from inception until September 3rd 2015 (date

last searched) to identify published studies that examined the association between epigenetic marks and blood cholesterol or TAG levels. The search was designed by an experienced biomedical information specialist. The search terms related to the exposure included epigenetics, DNA methylation, histone and CpG, and were combined with the terms TAG, TC, LDL-C, and HDL-C. We did not apply any language restriction, but we restricted the search to studies on humans. The full search strategies of all databases are provided in **Supplement 3.1.1.**. In order to identify potential additional studies, we checked the reference lists of studies included and contacted other experts in the field.

Study selection and inclusion criteria

We selected studies that examined the association between epigenetic marks (global, site specific or genome-wide methylation of DNA, or histone modifications) and plasma lipids (TAG, TC, LDL-C, and HDL-C) in humans. Two independent reviewers screened the retrieved titles and abstracts and selected eligible studies according to the selection criteria (**Supplement 3.1.2.**). Discrepancies between the two reviewers were resolved through discussion, with an arbitrator available if no consensus was reached. We retrieved full texts for studies that satisfied all selection criteria. These full-text articles were evaluated in detail once more by two investigators against the selection criteria.

Data extraction

Data extraction was performed using a predesigned form. Extracted information included study design, characteristics of the study population, location of the study, and level of confounder adjustment. Furthermore, for each study the tissue type and methods used to determine DNA methylation, specific CpG sites (for candidate gene studies), directions of the associations were reported.

RESULTS

After deduplication, we identified 757 potentially relevant citations **(Figure 3.1.1).** Based on the titles and abstracts, full texts of 44 articles were selected for more detailed evaluation. Of those, 31 articles reporting results of 23 unique studies met all our eligibility criteria and were therefore included in the analysis (**Tables 3.1.1-3.1.5**).

Characteristics of the included studies

Detailed characteristics of these studies are summarized in **Table 3.1.1-3.1.5**. All included studies were of cross-sectional design. Combined, the studies included data on 8,027 unique participants. Six studies included participants from China, five studies from Canada, and the rest included participants from USA, Italy, Iran, Egypt, Germany, Japan and Spain (**Table 3.1.1-3.1.5**). Of the 23 unique studies, one study examined solely on LDL-C as outcome, one study examined only HDL-C, whereas the rest of the studies included two or more types of blood lipids.

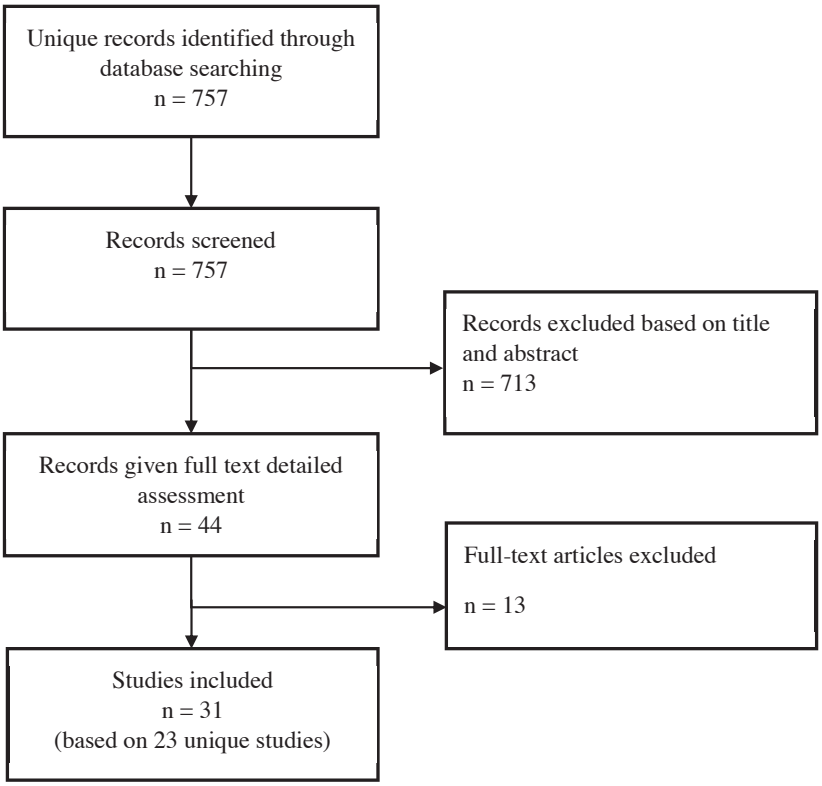


Figure 3.1.1. Flow chart of included studies

Global DNA methylation

Global DNA methylation refers to the overall genome-wide content of methylated cytosines within CpG sites. The majority of CpG sites (about 80%) are found in repetitive sequences (multiple copies of DNA that are normally methylated), such as Alu and long-interspersed nuclear element (LINE-1) and correlate with total genomic methylation content.[18-20] Methylation of these repetitive elements is widely used as a surrogate measure for global DNA methylation.[21, 22] Other methods (e.g., Luminometric Methylation Assay, LUMA and the [^{3}H]-methyl acceptance based method) to assess global genomic DNA methylation are primarily based on the digestion of genomic DNA by restriction enzymes HpaII and MspI.[23]

Global DNA methylation and triacylglycerol

Three studies examined global DNA-methylation assessed at LINE-1 elements or 5hmC methylation in relation to TAG levels (**Table 3.1.1**). One study assessed global DNA methylation in blood and found a positive association with TAG.[24] In line with this, another study reported that 5-hmC methylation levels were positively correlated with TAG blood levels.[25] The third study found no association of LINE-1 methylation in visceral adipose tissue with TAG.[26]

Global DNA methylation and total cholesterol

There were three studies that examined global DNA methylation in relation to TC (**Table 3.1.1**). All three studies assessed global methylation in blood.[24, 27, 28] One study assessed global DNA methylation in LINE-1 elements and found higher levels of global methylation with increasing levels of TC.[24] Similarly, another study reported 5hmC methylation levels to positively correlate with TC.[28] The last study used restriction enzymes HpaII and MspI to assess global methylation and found no association with TC.[27]

Table 3.1.1. Global DNA methylation and plasma lipid levels

Author	Population	Tissue	Methods	Adjustment	Association
TAG					
Turcot V et al. 2012[26]	Canada , n=186, 35.1 ± 7.7 y,	Visceral Adipose tissue	LINE-1 methylation	Age, sex and smoking and waist circumference	No association
Pearce MS et al. 2012[24]	UK, n=228, 49-51 y	PB	LINE-1 methylation	Sex	Positive association
Nicoletti CF et al. 2015[28]	Multiple countries, n=45, 43.1 ± 7.9 y, W.	Blood	5hmC (%)		Positive association
Total Cholesterol					
Pearce MS et al. 2012[24]	UK, n=228, 49-51 y	PB	LINE-1 methylation	Sex	Positive association
Nicoletti CF et al. 2015[28]	Multiple countries, n=45, 43.1 ± 7.9 y, W.	Blood	5hmC (%)		Positive association
Kato S et al. 2012[27]	Japan, n=44, 59 ± 12 y	PB	HpaII/MspI ratio		No association
LDL-C					
Cash HL et al. 2011[29]	USA, n=355, 32.0 ± 3.7 y	Peripheral Lymphocyte	LINE-1	Age, sex, BMI, HDL-C	Inverse association
Pearce MS et al. 2012 [24]	UK, n=228, 49-51 y, M and W	PB	LINE-1	Sex	Positive association
HDL-C					
Cash HL et al. 2011[29]	USA, n=355, 32.0 ± 3.7 y	Peripheral Lymphocyte	LINE-1	Age, sex, BMI, LDL-C	Positive association
Pearce MS et al. 2012[24]	UK, n=228, 49-51 y	PB	LINE-1	Sex	Inverse association
Turcot V et al. 2012[26]	Canada, n=186, 35.1 ± 7.7 y	Visceral Adipose tissue	LINE-1	Age, sex and smoking and waist circumference	No association
Kato S et al. 2012[27]	Japan, n=44, 59 ± 12,	PB	HpaII/MspI ratio		No association
HDL-C:LDL-C ratio					
Pearce MS et al. 2012[24]	UK, n=228, 49-51 y, M and W	PB	LINE-1	Sex	Inverse association

M, men; PB, peripheral blood; W, women; y, years.

Global DNA methylation and low-density lipoprotein-cholesterol

Global DNA methylation assessed at LINE-1 elements in blood in relation to LDL-C was examined in two studies (**Table 3.1.1**). Interestingly, the abovementioned study that found a positive association of global methylation and HDL-C (see section 3.2.1), observed an inverse association between global methylation and LDL-C,[29] whereas one of the other studies that reported an inverse association for HDL-C (see section 3.2.1) found a positive association for LDL-C.[24] The first study reported that the inverse association between global methylation and LDL-C was statistically significant only in women.

Global DNA methylation and high-density lipoprotein-cholesterol

Four studies assessed global DNA methylation in relation to HDL-C (**Table 3.1.1**). Two studies assessed global methylation in LINE-1 elements in blood, but showed contradictory results. One study reported a positive association between global methylation and HDL-C, which after stratification by sex was significant only in men.[29] In contrast, the other study reported an inverse association.[24] Another study assessed global methylation in LINE-1 elements in visceral adipose tissue and reported no significant association with HDL-C.[26] Similarly, a study using restriction enzymes HpaII and MspI found no significant association between global methylation and HDL-C.[27]

Candidate gene studies

Most of the studies on DNA methylation and blood lipids have used a hypotheses-driven, candidate gene approach. With this approach, only epigenetic marks at certain genes are investigated, that have been selected based on knowledge from previous genetic studies or based on knowledge about the biological function of the specific gene.

Methylation of candidate genes and triacylglycerol

There were 15 studies that examined methylation sites in or near known candidate genes for variations in TAG (**Table 3.1.2**). One study assessed DNA methylation in placenta,[30, 31] one in visceral adipose tissue[32] and the remaining studies examined blood cells. Overall, these studies found positive correlations between TAG and overall methylation at *NPC1* and *IGF2*, and methylation at several CpG sites at *FIAM* in blood; and negative correlations with overall DNA methylation at *MCP-1* and methylation at *MMP9*-CpG2, *LIPC*-CpG2, *GCK*-CpG2 and *GCK*-CpG4 in blood cells. Three studies showed sex-specific associations between different methylated regions and TAG: in women, there was a positive association between DNA methylation at *PLA2G7*,[12] or *BCL11A* with TAG,[33] and a negative association of methylation at *ABCG1*-CpG3 with TAG, whereas in men, no associations were found.[10] The latter study also showed that DNA methylation at *LIPC*-CpG2 was negatively correlated with TAG in men but not in women. No associations were found between TAG and DNA methylation at *SCARB1*,[10] *PLTP*,[10] *CETP*,[10] *LPL*,[10] *CD14*,[34] *TCF7L2*,[35] *MTHFR*,[36] *APOE*,[37] *ABCG1*,[38] *GALNT2*,[38] *HMGCR*,[38] *DPP4*,[32] *CD14*,[34] *Et-1*,[34] *HERV-W*,[34] *iNOS*,[34] or *TNF*-α.[34] Finally, one study (reported in two articles[30, 31]) included pregnant women and found a negative correlation between TAG concentrations in cord blood and methylation levels of the *ABCA1* gene in placenta, whereas no association was found with methylation at the *LPL* gene.

Table 3.1.2. Specific gene methylation and triglycerides: gene and genome-wide approaches

Author	Study design	Population/Age range/ Follow-up	Tissue type	methods	Methylation sites/	Adjustments	Main finding
Candidate gene approach							
Wu L et al. 2015[40]	Cross-sectional study	China, n=98, 8-18 y	PBL	MassARRAY platform	*FIAM* gene promoter, 36 CpG sites	Age, gender, BMI	Positive association
Guay SP et al. 2015[42]	Cross-sectional study	Canada, n=50, 53.8 ± 1.5 y, M	PBL	Bisulfite-pyrosequencing	*MMP9*	Age, BMI and glucose levels	Negative association (CpG2)
Guay et al., 2014[10]	Cross-sectional study	Canada, n=98, Men: n=61, 46.3 ±1.7 y, Women: n=37, 36.9±2.0 y, untreated FH patients	Blood leukocytes	Bisulfite pyrosequencing	*Primary: ABCG1*, *LIPC*, *PLTP* and *SCARB1* *secondary*: *ABCA1*, *CETP*, and *LPL*	Age, waist circumference, blood pressure, fasting plasma lipids and glucose levels	Men, Negative association for *LIPC*-CpGA2 Women: Negative asscociation for *ABCG1*-CpGC3 No significant associations were observed for other examined sites.
Jiang et al., 2013[12]	Case-control for CHD	China, n=72: cases: n=36, 62.5 ±5.5 y, controls: n=36, 61.7 ±5.2 y	Blood	Bisulphite pyrosequencing	*PLA2G7* gene promotor	Age, history of smoking, diabetes, hypertension	Women: Positive association Men: No significant association
Guay et al., 2014[67]	Case-control for CAD	Canada, n= 88 men, cases: n=38, controls: 50, median age 61 y, Caucasian , all recruited from patients who underwent heart surgery.	Blood leukocytes	Bis-pyrosequencing	*ABCA1*	Age, CAD status and medication	In older men (≥61 years): positive association In younger men (<61 years): no significant association
Afzali et al. 2013[41]	Case-control study	Iran, n=100, 59.9±12.1 y	PBL	Nested-methylation specific polymerase chain reaction method	*NCP1*		Positive association

Table 3.1.2. Specific gene methylation and triglycerides: gene and genome-wide approaches (continued)

Author	Study design	Population/Age range/ Follow-up	Tissue type	methods	Methylation sites/	Adjustments	Main finding
Bollati et al. 2014[34]	Cross-sectional study	Italy, n=165, 50.3±11.5 y	Whole blood	Pyromark MD System (Pyrosequencing, Inc. Westborough, MA, USA).	*CD14*, *Et-1*, *HERV-W*, *iNOS* and *TNF-α*	Age, sex, BMI, smoking, and % of neutrophils	No significant associations
Canivell et al. 2014[35]	Case-control study for T2D	Spain, n=186, 67.9±10.5 y	Whole blood	Sequenom EpiTYPER system	*TCF7L2*	Age, gender, BMI, physical activity, smoking status and waist circumference	No significant associations
Deodati et al. 2013[44]	Cross-sectional study	Italy, n=85, 11.5 ± 2.2 y	Blood lymphocytes	Methyl-Profiler DNA Methylation qPCR Assay (SABiosciences)	*IGF2*		Positive association
Ghattas et al. 2014[36]	Case-control study of end-stage renal cases	Egypt, n=96, 30-70 y	Peripheral blood	EpiTech Bisulfite conversion kit	*MTHFR*		No significant associations
Houde A.A. et al. 2014[30]	Cross-sectional study	Canada, n=126, 20-39 y, Women	Placenta	Bisulfite pyrosequencing and quantitative real-time PCR.	*LPL* gene, 3 CpG sites	Age, BMI, TAG at first trimester (for HDL), maternal glucose 2-h post oral glucose tolerance test at second trimester (for TAG) and history of gestational diabetes	No significant association
Houde A et al. 2013[31]	Cross-sectional study	Canada, n=100, 20-39 y, Women	Placenta and cord blood	Bisulfite pyrosequencing	*ABCA1* gene promotor, 22 CpG sites	Age, BMI, TAG at first trimester and history of gestational diabetes	Negative association for *ABCA1*-CPG5
Liu, Z. H. et al. 2012[45]	Case-control study for diabetes	China, n=47, 51.46±14.11 y	Peripheral blood	Bisulfite modification	*MCP-1*		Negative association

Table 3.1.2. Specific gene methylation and triglycerides: gene and genome-wide approaches (continued)

Author	Study design	Population/Age range/ Follow-up	Tissue type	methods	Methylation sites/	Adjustments	Main finding
Ma, Y. et al. 2015[37]	Cross-sectional study	United States, n=993, 47.88 y	CD4+ T cells	Bisulfite treatment	*APOE*, 13 CpG sites	Pedigree, sex, center, and the first principal component of cellular purity and population structure	No significant associations
Peng, P. et al. 2014[38]	Case-control study for CHD	China, n=139, 59.35±9.12 y	Peripheral blood	Bisulfite treatment	*ABCG1*, *GALNT2* and *HMGCR*	Age, gender, smoking, lipid level, history of hypertension, and history of diabetes	No significant associations
Tang, L. et al 2014[33]	Cases control study for diabetes	China, n=96, 59.1± 7.6 y	Peripheral blood	Sodium bisulfite	*BCL11A*, 5 CpG sites		Women: Positive association Men: No significant association
Turcot, V. et al 2013[32]	Cross-sectional study	Canada, n=92, 34 ± 0.7 y	Visceral adipose tissue	sodium bisulfite	*DPP4*, CpG94 to CpG102	Age, sex, smoking and waist circumference	No significant association
Xu, L. et al 2014[43]	Cases control study for CHD	China, n=72, age not reported	Peripheral blood	Bisulfite pyrosequencing	*GCK*, 4 CpG sites	Age, history of smoking, diabetes, and hypertension	Negative association for 2CpG sites
Genome-wide approach							
Pfeifferm L et al. 2015[11]	Cross-sectional study	Germany, Discovery set: KORA F4, n=1776, 60.8 ± 8.9 years, Replication set: KORA F3, n=499, 52.9 ± 9.6 years, InCHIANTI, n=472, 71.2 ± 16.0 y, and MuTHER study, n=634, 59.4 ± 9.0 y, women.	WB, adipose tissue and fibroblast	Infinium HumanMethylation27 BeadArray (Illumina)		Age, sex, BMI, smoking, alcohol, lipid-lowering drugs, physical activity, history of myocardial infarction, current hypertension, HbA1c levels, C-reactive protein and white blood cell count	Discovery set: Positive association for 5 CpG sites located in *MIR33B/SREBF1*, *APOA5, ABCG1*, and 2 other CpGs with no gene annotation. Replication: Results for *MIR33B/SREBF1*, *ABCG1*, and 1 CpG without annotation were replicated

Table 3.1.2. Specific gene methylation and triglycerides: gene and genome-wide approaches (continued)

Author	Study design	Population/Age range/ Follow-up	Tissue type	methods	Methylation sites/	Adjustments	Main finding
Irvin et al. 2014 [68]	Cross-sectional study	USA , GOLDN intervention study n=991, 48.8±16 years	CD4+ T cells	Illumina Infinium HumanMethylation450 Beadchip		Age, sex, study site, cell purity, family structure	Four CpGs (cg00574958, cg17058475, cg01082498, and cg09737197) in intron 1 of carnitine palmitoyltransferase 1A (*CPT1A*) were strongly positively associated with TAG (P=1.6×10–26 to 1.5×10–9). DNA methylation at *CPT1A* cg00574958 explained 11.6% and 5.5% of the variation in TAG in the discovery and replication cohorts, respectively.

FH, familial hypercholesterolemia; KORA, Cooperative health research in the Region of Augsburg; MuTHER, Multiple Tissue Human Expression Source; PBL,peripheral blood leukocytes; WB, whole blood; y, year

Methylation of candidate genes and total cholesterol

We identified 14 studies that examined methylation sites in, or near, candidate genes for blood lipids in relation to TC (**Table 3.1.3**). All studies assessed DNA methylation in blood, and one study assessed DNA methylation in both blood and adipose tissue.[39] Overall, these studies found a higher degree of methylation of *FIAM*, *ABCG1*, *PLA2G7*, *NPC1* and *MTHFR* and lower levels of methylation of *DPP4* in the peripheral blood with higher levels of plasma TC. Furthermore, higher methylation of *APOE*-CpG-7, *COL14A1*-CpG2, *LEP* and *ADIPOQ*; and lower methylation levels of *APOE*-CpG1, 2, 10, 12, 13, *ABCG1*-CpGC3 and *TCF7L2*-CpG27 were associated with higher levels of TC. No associations were found between TC and DNA methylation at *IGF2*, *ABCA1*, *MCP-1*, *GALNT2*, *HMGCR* and *BCL11A*. There were two studies investigating the association between DNA methylation at *ABCG1* and TC: one study found an inverse association between *ABCG1*-CpG3 and TC in men,[10] whereas the other study reported no association between overall methylation at this gene and TC in either sex.[38]

Methylation of candidate genes and low-density lipoprotein-cholesterol

Fourteen studies investigated methylation sites in or near known candidate genes for blood lipids in relation to LDL-C (**Table 3.1.4**). Ten studies used blood cells to assess DNA methylation, one study examined visceral adipose tissue, one study used placenta, and two studies used combinations of these tissues. These studies found that higher overall methylation at *ABCA1*,[10] *TNF*-α,[34] *LEP*, *ADIPOQ*,[39] *MTHFR*,[36] and several CpG sites at *FIAM*,[40] and lower overall methylation at *ADRB3*[10] and *NPC1*[41] were associated with higher LDL-C concentrations. LDL-C was also found to be positively correlated to methylation at *MMP9*-CpG1,[42] *APOE*-CpG9,[37] and *GCK*-CpG3,[43] and negatively correlated with methylation at *MMP9*-CpG4[42] and *TCF7L2*-CpG27.[35] No significant associations were reported for LDL-C concentration in relation to methylation at *CD14*,[34] *Et-1*,[34] *HERV-W*,[34] *iNOS*,[34] *IGF2*,[44] *MCP-1*,[45] *ABCG1*,[38] *GALNT2*,[38] *HMGCR*,[38] or *DPP4*.[32] Finally, one study examined DNA methylation at *LPL* and *ABCA1* genes in placenta of pregnant women in relation to LDL-C and reported null associations.[30, 31]

Methylation of candidate genes and high-density lipoprotein-cholesterol

The association of DNA methylation in or near known candidate genes for blood lipids with HDL-C concentrations was investigated in 13 unique studies (**Table 3.1.5**). Ten studies assessed DNA methylation in blood and the other three studies examined methylation in several tissues, including placental tissue and visceral adipose tissue. Overall, these studies found that a lower degree of methylation of *FIAM*,[40] *ABCA1*,[46] *NPC1*,[41] *MTHFR*,[36] *LEP*,[39] *ADIPOQ*,[39] and higher levels of methylation of *LPL*[47] and *DPP4*[48] genes in the peripheral blood were associated with higher levels of plasma HDL-C. No associations were found between HDL-C and DNA methylation at *CD14*,[34] *Et-1*,[34] *HERV-W*,[34] *iNOS*,[34] *TNF*-α,[34] *TCF7L2*,[35] *IGF2*,[44] *MCP-1*,[45] *APOE*,[37] *ABCG1*,[38] *GALNT2*,[38] or *HMGCR*.[38] Two studies showed sex-specific associations between HDL-C and DNA methylation at *PLTP*,[10] *CETP*,[47] and *LIPC-CpGA2*[10]: these associations were all significant in men but not in women.

Table 3.1.3. Specific gene methylation and total cholesterol: gene and genome-wide approaches

Author	Study design	Population characteristics	Tissue type	Methods	Methylation sites	Adjustments	Main finding
Candidate gene approach							
Wu L et al. 2015[40]	Cross-sectional	China, n=98, 8-18 y	PBL	MassARRAY platform	*FIAM* gene promoter, 36 CpG site	Age, gender, BMI	Positive association
Houde A et al. 2015[39]	Cross-sectional	Canada, n=73, 34,7 ± 7.1 y	Blood, subcutaneous adipose tissue, and visceral adipose tissue	Pyrosequencing	*LEP* and *ADIPOQ*	Age, sex and waist girth.	Association with *ADIPOQ*-CpGE3 (direction not reported)
Guay et al., 2014[10]	Cross-sectional study	Canada, n=98, n=61 men 46.3 ±1.7 y, and n=37 women 36.9±2.0 y, untreated FH patients	Blood leucocytes	Bisulfite pyrosequencing	*Primary: ABCG1*, *LIPC*, *PLTP* and *SCARB1* *secondary*: *ABCA1*, *CETP*, and *LPL*	Age, waist circumference, blood pressure, fasting plasma lipids and glucose levels	Men, Negative association for *ABCG1*-CpGC3 Women: No significant associations
Guay SP et al. 2015[42]	Cross-sectional study	Canada, n=50, 53.8 ± 1.5 y, men	PBL	Bisulfite-pyrosequencing	*COL14A1*	Age, BMI and glucose levels	Positive association for *COL14A1*-CpG2
Guay et al., 2014[67]	Case-control for CAD	Canada, n= 88 (38 cases, 50 controls), median age 61 y, M, Caucasian, all recruited from patients who underwent heart surgery	Blood leucocytes	Bis-pyrosequencing	*ABCA1*	Age, CAD status and medication	Older men (≥61 years) Positive association Younger men (<61 years) no significant association
Jiang et al., 2013[12]	Case-control for CHD	China, n=72, cases: n=36 , age 62.5 ±5.5 y, controls: n= 36, 61.7 ±5.2 y	Blood	Bisulphite pyrosequencing	*PLA2G7* gene promotor	Age, history of smoking, diabetes, hypertension	Women: positive association Men: no significant association

Table 3.1.3. Specific gene methylation and total cholesterol: gene and genome-wide approaches (continued)

Author	Study design	Population characteristics	Tissue type	Methods	Methylation sites	Adjustments	Main finding
Afzali et al. 2013[41]	Case-control study for CVD	Iran, n=100, 59.9±12.1 y	Peripheral blood leukocytes	Nested-methylation specific polymerase chain reaction method	*NCP1*		Positive association
Canivell et al. 2014[35]	Case-control study for T2D	Spain, n=186, 67.9±10.5 y	Whole blood	Sequenom EpiTYPER system	*TCF7L2*	Age, gender, BMI, physical activity, smoking status and waist circumference	Positive association (CpG 27)
Deodati et al. 2013[44]	Cross-sectional study	Italy, n=85, 11.5 ± 2.2 y	Blood lymphocytes	Methyl-Profiler DNA Methylation qPCR Assay (SABiosciences)	*IGF2*		No significant association
Ghattas et al. 2014[36]	Case-control study for end-stage renal cases	Egypt, n=96, 30-70 y	Peripheral blood	EpiTech Bisulfite conversion kit	*MTHFR*		Positive association
Houde A et al. 2013[31]	Cross-sectional study	Canada, n=100, 20-39 y W	Placenta and cord blood	Bisulfite pyrosequencing	*ABCA1* gene promotor, 22 CpG sites	Age, BMI, TAG at first trimester and history of gestational diabetes	No significant association
Liu, Z. H. et al. 2012[45]	Case-control study for diabetes	China, n=47, 51.46±14.11 y	Peripheral blood	Bisulfite modification	*MCP-1* promoters		No significant association
Ma, Y. et al. 2015[37]	Cross-sectional study	United States, n=993, 47.88 y	CD4+ T cells	Bisulfite treatment	*APOE*, 13 CpG sites	Pedigree, sex, center, and the first principal component of cellular purity and population structure	Positive association for 1 CpG site (CpG 7) Negative association for 5 CpG sites (CpG 1,2,10,12, and 13)

Table 3.1.3. Specific gene methylation and total cholesterol: gene and genome-wide approaches (continued)

Author	Study design	Population characteristics	Tissue type	Methods	Methylation sites	Adjustments	Main finding
Peng, P. et al. 2014[38]	Case-control study for CHD	China, n=139, 59.35±9.12 y	Peripheral blood	Bisulfite treatment.	*ABCG1*, *GALNT2* and *HMGCR*	Age, gender, smoking, lipid level, history of hypertension, and history of diabetes	No significant associations
Tang, L. et al 2014[33]	Cases control study for diabetes	China, n=96, 59.1± 7.6 y	Peripheral blood	Sodium bisulfite	*BCL11A*, 5 CpG sites		No significant association
Turcot, V. et al 2011[48]	Cross-sectional study	Canada, n=92, 34 ± 0.7 y	Visceral adipose tissue	Sodium bisulfite	*DPP4*,102 CpG sites	Age, smoking, and BMI	No significant association
Turcot, V. et al 2013[32]	Cross-sectional study	Canada, n=92, 34 ± 0.7 y	Visceral adipose tissue	Sodium bisulfite	*DPP4*, CpG94 to CpG102	Age, sex, smoking and waist circumference	Negative association
Genome-wide approach							
Pfeifferm L et al. 2015[11]	Cross-sectional study	Germany, Discovery set: KORA F4, n=1776, 60.8 ± 8.9 years Replication set: KORA F3, n=499, 52.9 ± 9.6 years, InCHIANTI, n=472, 71.2 ± 16.0 years, MuTHER study, n=634, 59.4 ± 9.0 years, W	WB, adipose tissue and fibroblast	Infinium HumanMethylation27 BeadArray (Illumina)		Age, sex, BMI, smoking, alcohol, lipid-lowering drugs, physical activity, history of myocardial infarction, current hypertension, HbA1c levels, C-reactive protein and white blood cell count	No significant associations

FH, familial hypercholesterolemia; KORA, Cooperative health research in the Region of Augsburg; MuTHER, Multiple Tissue Human Expression Source; PBL,peripheral blood leukocytes; WB, whole blood; y, year

Table 3.1.4. Specific gene methylation and LDL-cholesterol: gene and genome-wide approaches

Author	Study design	Population/Age range/Follow-up	Tissue type	methods	Methylation sites	Adjustments	Main finding
Candidate gene approach							
Wu L et al. 2015[40]	Cross-sectional study	China, n=98, 8-18 y	PBL	MassARRAY platform	*FIAM* gene promoter, 36 CpG sites	Age, gender, BMI	Positive association
Houde A et al. 2015[39]	Cross-sectional study	Canada, n=73, 34,7 ± 7.1 y	Blood, subcutaneous adipose tissue, and visceral adipose tissue	Pyrosequencing	*LEP* and *ADIPOQ*	Age, sex and waist girth.	Positive associations for *LEP* and *ADIPOQ*
Guay SP et al. 2015[42]	Cross-sectional study	Canada, n=50, 53.8 ± 1.5 y, Men	PBL	Bisulfite-pyrosequencing	*MMP9*	Age and waist circumference	Positive association for CpG1 Negative association for CpG4.
Guay et al., 2014[10]	Cross-sectional study	Canada , n=61 FH men and n=30 severely obese men (BMI > 40 kg/m2) without FH, who underwent a biliopancreatic diversion with duodenal switch	Blood, Visceral adipose tissue	Pyrosequencing (sodium bisulfite conversion, PCR amplification, and sequencing by synthesis assay)	*ADRB3* gene promotor		Negative association (in subgroup of n=41 FH subjects without *ADRB3* rare alleles)
Guay et al., 2014[67]	Case-control for CAD	Canada, n= 88 men, cases: n=38, controls: 50, median age 61 y, Caucasian , all recruited from patients who underwent heart surgery.	Blood leukocytes	Bis-pyrosequencing	*ABCA1*	Age, CAD status and medication	In older men (≥61 years): positive association In younger men (<61 years): no significant association

Table 3.1.4. Specific gene methylation and LDL-cholesterol: gene and genome-wide approaches (continued)

Author	Study design	Population/Age range/Follow-up	Tissue type	methods	Methylation sites	Adjustments	Main finding
Afzali et al. 2013[41]	Case-control study	Iran, 100, 59.9±12.1 y	Peripheral blood leukocytes	Nested-methylation specific polymerase chain reaction method.	*NCP1*		Positive association
Bollati et al. 2014[34]	Cross sectional study	Italy, n=165, 50.3±11.5 y	Whole blood	Pyromark MD System (Pyrosequencing, Inc. Westborough, MA, USA).	*CD14*, *Et-1*, *HERV-W*, *iNOS* and *TNF*-α	Age, sex, BMI, smoking, and % of neutrophils	Positive association for TNFα. No significant associations for other genes.
Canivell et al. 2014[35]	Case-control study for T2D	Spain, n=186, 67.9±10.5 y	Whole blood	Sequenom EpiTYPER system	*TCF7L2*	Age, gender, BMI, physical activity, smoking status and waist circumference	Positive association for CpG 27
Deodati et al. 2013[44]	Cross-sectional study	Italy, n=85, 11.5 ± 2.2 y	Blood lymphocytes	Methyl-Profiler DNA Methylation qPCR Assay (SABiosciences)	*IGF2*		No significant association
Ghattas et al. 2014[36]	Case-control study (end-stage renal cases)	Egypt, n=96, 30-70 y	Peripheral blood	EpiTech Bisulfite conversion kit	*MTHFR*		Positive association
Houde A.A. et al. 2014[30]	Cross-sectional study	Canada, n=126, 20-39 y , Women	Placenta	Bisulfite pyrosequencing and quantitative real-time PCR.	*LPL* gene, 3 CpG sites	Age, BMI, and history of gestational diabetes	No significant association
Houde A et al. 2013[31]	Cross-sectional study	Canada, n=100, 20-39 y, Women	Placenta and cord blood	Bisulfite pyrosequencing	*ABCA1* gene promotor, 22 CpG sites	Age, BMI, TAG at first trimester and history of gestational diabetes	No significant association

Table 3.1.4. Specific gene methylation and LDL-cholesterol: gene and genome-wide approaches (continued)

Author	Study design	Population/Age range/Follow-up	Tissue type	methods	Methylation sites	Adjustments	Main finding
Liu, Z. H. et al. 2012[45]	Case-control study for diabetes	China, n=47, 51.46±14.11 y	Peripheral blood	Bisulfite modification	*MCP-1* promoters		No significant association
Ma, Y. et al. 2015[37]	Cross-sectional study	USA, n=993, 47.88 y	CD4+ T cells	Bisulfite treatment	*APOE*, 13 CpG sites	Pedigree, sex, center, and the first principal component of cellular purity and population structure	Positive association for CpG 9
Peng, P. et al. 2014[38]	Case-control study for CHD	China, n=139, 59.35±9.12 y	Peripheral blood	Bisulfite treatment	*ABCG1*, *GALNT2* and *HMGCR*	Age, gender, smoking, lipid level, history of hypertension, and history of diabetes	No significant associations
Turcot, V. et al 2013[32]	Cross-sectional study	Canada, n=92, 34 ± 0.7 y	Visceral adipose tissue	Sodium bisulfite	*DPP4*, CpG94 to CpG102	Age, sex, smoking, and waist circumference	No significant association
Xu, L. et al 2014[43]	Cases control study for CHD	China, n=72, age not reported	Peripheral blood	Bisulfite pyrosequencing	*GCK*, 4 CpG sites	Age, history of smoking, diabetes, and hypertension	Positive association for 1 CpG site
Genome-wide approach							

Table 3.1.4. Specific gene methylation and LDL-cholesterol: gene and genome-wide approaches (continued)

Author	Study design	Population/Age range/Follow-up	Tissue type	methods	Methylation sites	Adjustments	Main finding
Pfeifferm L et al. 2015[11]	Cross-sectional study	Germany, Discovery set: KORA F4, n=1776, 60.8 ± 8.9 years Replication set: KORA F3, n=499, 52.9 ± 9.6 years, InCHIANTI, n=472, 71.2 ± 16.0 years, and MuTHER study, n=634, 59.4 ± 9.0 years, women only	WB, adipose tissue and fibroblast	Infinium HumanMethylation450 BeadChip (Illumina)		Age, sex, BMI, smoking, alcohol, lipid-lowering drugs, physical activity, history of myocardial infarction, current hypertension, HbA1c levels, C-reactive protein and white blood cell count	Positive association for *TNIP1*

FH, familial hypercholesterolemia; KORA, Cooperative health research in the Region of Augsburg; MuTHER, Multiple Tissue Human Expression Source; PBL,peripheral blood leukocytes; WB, whole blood; y, year

Table 3.1.5. Specific gene methylation and HDL-cholesterol: gene and genome-wide approaches

Author	Study design	Population/Age range/Follow-up	Tissue type	Methods	Methylation sites	Adjustments	Main finding
Candidate gene approach							
Wu L et al. 2015[40]	Cross-sectional study	China, n=98, 8-18 y	PBL	MassARRAY platform	*FIAM* gene promoter, 36 CpG sites	Age, gender, BMI	Negative association
Houde A et al. 2015[39]	Cross-sectional study	Canada, n=73, 34,7 ± 7.1 y	Blood, subcutaneous adipose tissue, and visceral adipose tissue	Pyrosequencing	*LEP* and *ADIPOQ*	Age, sex and waist girth.	Association with LEP-CpG7 (direction not reported)
Guay et al., 2014[10]	Cross-sectional study	Canada , n=98 untreated FH patients, men: n=61, 46.3 ±1.7 y, women: n= 37, 36.9±2.0y	Blood leucocytes	Bisulfite pyrosequencing	*Primary: ABCG1*, *LIPC*, *PLTP* and *SCARB1* *secondary*: *ABCA1*, *CETP*, and *LPL*	Age, waist circumference, blood pressure, fasting plasma lipids and glucose levels	Men: Negative association for *LIPC*-CpGA2 *PLTP*-CpGC Women: No significant associations
Guay et al., 2012[46]	Cross-sectional study	Canada , n=97, men: n=59, 45.6 ± 1.7 y, women: n=38, 36.9 ± 1.9 y Caucasian subjects with familial hypercholesterolemia (FH), all with same LDLR gene mutation	Blood leucocytes	Pyrosequencing (sodium bisulfite conversion, PCR amplification, and sequencing by synthesis assay)	*ABCA1* gene promotor	Age, gender, waist circumference, fasting triglycerides	Negative association
Guay et al., 2013[47]	Cross-sectional study	Canada, n=98, men: 61, 46.3 ±1.7 y, women: n=37, 36.9 ± 2.0 y Untreated FH patients	Blood leucocytes	Bisulfite pyrosequencing	*LDLR*, *CETP*, *LCAT* and *LPL* gene promotors	Age, gender, waist circumference, fasting triglycerides	Positive association for LPL Men: Negative association for CETP No significant associations for LCAT or LDLR
Afzali et al. 2013[41]	Case-control study	Iran, n=100, 59.9±12.1 y	Peripheral blood leukocytes	Nested-methylation specific polymerase chain reaction method	*NCP1*		Negative association

Table 3.1.5. Specific gene methylation and HDL-cholesterol: gene and genome-wide approaches (continued)

Author	Study design	Population/Age range/Follow-up	Tissue type	Methods	Methylation sites	Adjustments	Main finding
Bollati et al. 2014[34]	Cross sectional study	Italy, n=165, 50.3±11.5 y	Whole blood	Pyromark MD System (Pyrosequencing, Inc. Westborough, MA, USA)	*CD14*, *Et-1*, *HERV-W*, *iNOS* and *TNF*-α	Age, sex, BMI, smoking, and % of neutrophils	No significant associations
Canivell et al. 2014[35]	Case-control study for T2D	Spain, n=186, 67.9±10.5 y	Whole blood	Sequenom EpiTYPER system	*TCF7L2*	Age, gender, BMI, physical activity, smoking status and waist circumference.	No significant association
Deodati et al. 2013[44]	Cross-sectional study	Italy, n=85, 11.5 ± 2.2 y	Blood lymphocytes	Methyl-Profiler DNA Methylation qPCR Assay (SABiosciences)	*IGF2*		No significant association
Ghattas et al. 2014[36]	Case-control study (end-stage renal cases)	Egypt, n=96, 30-70 y	Peripheral blood	EpiTech Bisulfite conversion kit	*MTHFR*		Negative association
Houde A.A. et al. 2014[30]	Cross-sectional study	Canada, n=126, 20-39 y, Women	Placenta	Bisulfite pyrosequencing and quantitative real-time PCR	*LPL* gene, 3 CpG sites	Age, BMI, TAG at first trimester, and history of gestational diabetes	Negative association
Houde A et al. 2013[31]	Cross-sectional study	Canada, n=100, 20-39 y, Women	Placenta and cord blood	Bisulfite pyrosequencing	*ABCA1* gene promotor, 22 CpG sites	Age, BMI, TAG at first trimester and history of gestational diabetes	Negative association
Liu, Z. H. et al. 2012[45]	Case-control study for diabetes	China, n=47, 51.46±14.11 y	Peripheral blood	Bisulfite modification	*MCP-1* promoters		No significant association
Ma, Y. et al. 2015[37]	Cross-sectional study	United States, n=993, 47.88 y	CD4+ T cells	Bisulfite treatment	*APOE*, 13 CpG sites	Pedigree, sex, center, and the first principal component of cellular purity and population structure	No significant association

Table 3.1.5. Specific gene methylation and HDL-cholesterol: gene and genome-wide approaches (continued)

Author	Study design	Population/Age range/Follow-up	Tissue type	Methods	Methylation sites	Adjustments	Main finding
Peng, P. et al. 2014[38]	Case-control study for CHD	China, n=139, 59.35±9.12 y	Peripheral blood	Bisulfite treatment	*ABCG1*, *GALNT2* and *HMGCR*	Age, gender, smoking, lipid level, history of hypertension, and history of diabetes	No significant associations
Turcot, V. et al 2011[48]	Cross-sectional study	Canada, n=92, 34 ± 0.7 y	Visceral adipose tissue	Sodium bisulfite	*DPP4*, 102 CpG sites	Age, smoking, and BMI	Positive association
Turcot, V. et al 2013[32]	Cross-sectional study	Canada, n=92, 34 ± 0.7 y	Visceral adipose tissue	Sodium bisulfite	*DPP4*, CpG94 to CpG102	Age, sex, smoking and waist circumference	No significant association
Genome-wide approach							
Pfeifferm L et al. 2015[11]	Cross-sectional study	Germany, Discovery set: KORA F4, n=1776, 60.8 ± 8.9 years Replication set: KORA F3, n=499, 52.9 ± 9.6 years, InCHIANTI, n=472, 71.2 ± 16.0 years, and MuTHER study, n=634, 59.4 ± 9.0 years, women only	WB, adipose tissue and fibroblast	Infinium HumanMethylation27 BeadArray (Illumina)		Age, sex, BMI, smoking, alcohol, lipid-lowering drugs, physical activity, history of myocardial infarction, current hypertension, HbA1c levels, C-reactive protein and white blood cell count	Discovery set: Positive association for 5 CpG sites located in *ABCG1*, *SREBF1*, and *CPT1A* Replication: All results from discovery cohort were replicated
Guay et al., 2012[50]	Cross-sectional study	Canada, n=21 men with FH, 10 with low HDL-c, age 44.2±3.7 y, and 11 with high HDL-c, 60.0 ±3.0 y Replication, n=276 FH subjects	Blood leucocytes	Infinium HumanMEthylation27 BeadChip and validation for seven genes with bisulfite DNA pyrosequencing		Age, gender, waist circumference, fasting triglycerides	Discovery cohort: 409 differentially methylated CpG sites identified Replication cohort: Positive association for *TNNT1* replicated

FH, familial hypercholesterolemia; KORA, Cooperative health research in the Region of Augsburg; MuTHER, Multiple Tissue Human Expression Source; PBL,peripheral blood leukocytes; WB, whole blood; y, year

Epigenome-wide analyses

Epigenome-wide association studies (EWAS) have recently become available, in which methylation levels of all measurable CpG sites are tested in relation to a phenotype or biological trait. This approach provides the opportunity to identify novel CpG sites in relation to disease phenotypes. In contrast to the candidate gene approach, this approach is hypothesis-free. This relatively new method provides many new opportunities, but also new challenges arise with the introduction of this method. For instance, due to the possibility to measure a large amount of CpG sites, this introduces a multiple testing issue, and an appropriate adjustment is necessary to avoid false positive results. Since this is a hypothesis free approach, results obtained in a first study (a discovery cohort) can be used to generate hypotheses, that subsequently need to be tested in a replication cohort.

Epigenome-wide analyses and blood lipids

Three studies examined blood lipids-associated differentially methylated sites in blood. One study investigated solely TAG,[49] one study solely HDL-C,[50] and the other study investigated TAG, TC, LDL-C and HDL-C simultaneously (**Tables 3.1.2-3.1.5**).[11] This latter study, using samples from blood, adipose tissue and fibroblasts, found no CpG site to be associated with TC concentrations (**Table3.1.3**).[11] The same study identified 11 CpG sites, located in *ABCG1* (3 CpGs), *MIR33B* (1 CpGs), *SREBF1* (1CpG), *CPT1A* (1 CpG), *TXNIP* (1CpG), *APOA5* (1 CpG), *TNIP1*(1 CpG), and 2 other CpGs with no gene annotation (cg07504977, cg07815238), that were associated with either HDL-C, LDL-C, or TAG (**Tables 3.1.2., 3.1.3. and 3.1.5.**). Associations between *ABCG1*, *SREBF1*, and *CPT1A* and HDL-C were replicated in other blood samples or adipose tissues. The association between *TNIP1* and LDL-C was also replicated. Furthermore, associations between *MIR33B*, *ABCG1*, and one CpG site not annotated to a gene and TAG concentrations were replicated in other blood samples and in adipose and skin tissue. Irvin et al.[49] investigating DNA methylation at CD4+ T cells reported four CpG sites (cg00574958, cg17058475, cg01082498, and cg09737197) in intron 1 of carnitine palmitoyltransferase 1A (*CPT1A*) to positively correlate with TAG levels (P=1.6×10^{-26} to 1.5×10^{-9}): DNA methylation at *CPT1A* cg00574958 explained 11.6% and 5.5% of the variation in TAG in the discovery and replication cohorts, respectively. Another study performed EWAS for HDL-C and identified 409 differentially methylated loci in the discovery cohort, without correction for multiple testing.[50] Pathway analyses showed that 37 of these differentially methylated loci were involved in pathways related to lipid metabolism. One of the top hits, a CpG site located in the promoter of the *TNNT1* gene, was replicated: *TNNT1* methylation levels were positively correlated with HDL-C levels.

GENERAL DISCUSSION

The present work is the first to systematically review the currently available evidence on the association of DNA methylation and histone modifications with plasma lipid levels. Overall, the scarce evidence suggests no consistent association between global DNA methylation with blood concentrations of triacylglycerol or cholesterol. Several studies did report that DNA methylation at specific individual genes was associated with concentrations of one or more blood lipids, suggesting a role of epigenetics in blood lipids levels. However, most studies were of relatively small size and replication is lacking

for most results. Care must thus be taken in the interpretation of the specific findings and further studies are required. Furthermore, there is lack of evidence on the role of histone modifications in dyslipidaemia.

Summary and interpretation of main findings

Only four studies examined the association between global DNA methylation and specific blood lipids and reported contradictory results.[24, 26, 27, 29] These findings are in line with other studies showing no consistent association between global DNA methylation and other health outcomes, such as cardiovascular disease and type 2 diabetes.[8, 9] Global methylation may not be specific enough as a measure of epigenetics to predict health, as methylation may have different effects at different positions in the DNA. For example, hypermethylation of promoter CpG islands has been linked to reduced expression of the associated gene,[51] whereas hypermethylation of the body of the gene is associated with increased gene expression.[52] Therefore, gene-specific methylation may provide better insight on the association of epigenetic modifications and health.

In the current review, several studies were identified that examined the association between methylation of individual genes, either in the full genome or focused on specific genes, with blood lipids. The majority of these studies used a candidate-gene approach, i.e., they studied methylation at genes for which previous studies suggested that their genetic code is associated with blood lipid levels. These candidate-gene studies reported that epigenetic regulation of *FIAM*, *PLA2G7*, *NPC1*, *MTHFR*, *DPP4*, *IGF2*, *BCL11A*, *LEP*, *ADIPOQ*, *APOA5*, and *ABCG1* were associated with lipid levels. Future studies are needed to replicate these findings.

In addition to the candidate-gene approach, gene-specific DNA methylation can also be studied for the whole genome. Only three studies were identified that examined this genome-wide DNA methylation in relation to blood cholesterol or TAG levels.[11, 49, 50] Their results confirmed some associations of known lipid-associated genes, such as *ABCG1*, but also identified new potential epigenetic predictors of blood lipids levels. *ABCG1*, *CPT1A*, *TNNT1*, *MIR33B*, *SREBF1*, and *TNIP* was associated with concentrations of TC, TAG, LDL-C and HDL-C. Blood lipid levels in relation to DNA methylation at *ABCG1*, *CPT1A*, *TNNT1*, *MIR33B*, *SREBF1*, and *TNIP* were replicated.

No studies were identified that studied histone modification in relation to blood cholesterol or TAG levels. We previously also identified a scarcity of studies examining this epigenetic mark in relation to cardiovascular disease and type 2 diabetes.[8, 9] Modification of histones associated with specific genes, and in particular histone acetylation, may play a role in controlling eukaryotic gene expression.[53] Histone modifications can also affect processes such as transcription, mitosis and chromosome stability.[53] More studies are thus needed to explore the potential relation of histone modifications with health.

Underlying mechanisms

DNA methylation, the addition of a methyl group to the 5' position of cytosine in a dinucleotide CpG site, is an important mechanism in gene expression regulation.[54] Until recently, it was believed that loss of DNA methylation promotes gene expression.[55] However, recent studies have shown that underlying mechanisms are more complex, and that the association of DNA methylation with gene expression depends on where within the gene sequence the methylation occurs.[51]

Studies included in this systematic review report that differentially methylated CpG sites annotated to *FIAM*, *PLA2G7*, *NPC1*, *MTHFR*, *DPP4*, *IGF2*, *BCL11A*, *LEP*, *ADIPOQ*, *APOA5*, *ABCG1*, *CPT1A*, *TNNT1*, *MIR33B*, *SREBF1*, and *TNIP* were associated with concentrations of TC, TAG, LDL-C and HDL-C. Of all these CpG sites that were reported to be related to blood lipid levels, CpG sites annotated to *ABCG1*, *CPT1A*, *TNNT1*, *MIR33B*, *SREBF1*, and *TNIP* were replicated. An overview of these replicated genes with a description of the protein these genes encode is provided in **Table 3.1.6**. Methylation at *ABCG1* was reported to be associated with lipid levels in a candidate-gene study.[10] Additionally, methylation of this gene in relation to TAG and HDL-C levels was identified and replicated in an EWAS.[11] *ABCG1* encodes for ATP-binding cassette sub-family G member 1 protein, and is known to be involved in lipid homeostasis as cholesterol transporter.[56] The other replicated CpG sites, including *CPT1A*, *TNNT1*, *MIR33B*, *SREBF1*, and *TNIP* were not previously reported by candidate gene studies, but identified by EWAS. The majority of these CpG sites are located in genes that are involved in cholesterol and lipid metabolism. For example, *CPT1A* gene, encodes for carnitine palmitoytransferase-1, which has an important role of transporting long-chain fatty acids in mitochondria.[57] *MIR33B*, a non-coding microRNA, is involved in regulation of cholesterol biosynthesis by *SREBF1*, which encodes for sterol-regulatory-element-binding-proteins (SREBP).[58] *TNIP* encodes for TNFa-induced protein 3 (TNFAIP3)-interacting protein 1, which regulates transcriptional activity of PPAR and RAR.[59]

Table 3.1.6. Description of genes at which methylation levels are associated with blood lipid levels.

Gene	Common name	Encodes for	Function
ABCG1	ATP-binding cassette transporter 8	ATP-binding cassette sub-family G member 1 protein	Macrophage cholesterol and phospholipid transporter
CPT1A	Carnitine palmitoytransferase-1A	Carnitine palmitoytransferase-1	Transporting long-chain fatty acids in mitochondria.
TNNT1	Troponin T Type 1	Troponin subunit protein	Regulates striated muscle contraction in response to intracellular calcium concentrations
MIR33B	MicroRNA 33b	Non-coding RNA	Regulation of cholesterol biosynthesis by SREBP
SREBF1	Sterol regulatory element-binding transcription factor 1	Sterol-regulatory-element-binding-proteins (SREBP)	Activate the uptake and synthesis of cholesterol and active the synthesis of fatty acids
TNIP		TNFa-induced protein 3 (TNFAIP3)-interacting protein 1	Regulation of multiple receptor-mediated activity of PPAR and RAR

Quality of the included studies and directions for future research

The current evidence suggests that epigenetic changes at certain genes may play a role in the development of dyslipidaemia, but quality of the included studies should be taken into account in the interpretation of the their findings. There are several components that determine the quality of the studies, such as design, sample size, use of tissue, confounder adjustment, and replication.

We identified 23 unique studies that examined the association between DNA methylation and blood cholesterol and TAG levels, but all had a cross-sectional design. Therefore it is not possible to

draw inferences on the temporal relation between DNA methylation and blood lipids. Furthermore, many of the included studies were of limited sample size and may have been underpowered to detect associations. Considering that DNA methylation is tissue-specific, tissue used for analyses in the included studies is another quality aspect that should be to be taken into account when interpreting the results. Most studies used blood tissue, which is one of the most accessible tissues. For examining associations with blood lipids, blood might be the appropriate tissue to measure DNA methylation. However, not all CpG sites related to lipid concentrations may be methylated in blood, but rather in adipose or liver tissue. Therefore, some CpG sites may have been overlooked in studies that only investigated blood. An another important quality aspect in epigenetic studies is adjustment for confounders. In contrast to genetics, epigenetic studies are prone to confounding. For instance, since the majority of the studies used whole blood to measure DNA methylation,[10-12, 31, 33-36, 38, 40-47, 50, 60, 61] results may have been confounded by cell type proportion.[62] Although some studies adjusted for measured or estimated cell type proportions,[11, 34] several studies did not which may have resulted in spurious findings.[10, 12, 31, 33, 35, 36, 38, 40-47, 50, 60, 61] Besides cell type proportion, there are many other potential confounding factors in the association between DNA methylation and lipid levels, such as age, sex, BMI, medication, alcohol consumption, and smoking. Of all studies included, only one study adjusted for all these covariates.[63] Finally, although all EWASs included in our review performed replication of their findings in independent cohorts, replication is still lacking for results from the candidate-gene studies.

Therefore, more research is needed in order to draw stronger conclusions on the likely complex association between epigenetics and blood lipids. More specifically, more studies should examine gene-specific DNA methylation in genome-wide analyses in large samples and findings from these studies should be replicated in other populations. Furthermore, longitudinal studies are needed to examine temporal relation of epigenetics with blood lipids, i.e., whether differences in epigenetic marks may predict changes in blood lipid levels and subsequently cardiometabolic disease risk. If these studies confirm that epigenetic marks may predict blood lipids levels, methylation profiling may be used to identify individuals at high risk of dyslipidaemia and cardiometabolic disease before actual changes are detectable in lipid levels.[64] Additionally, these results forthcoming these future studies may be used to identify pathways underlying the development of dyslipidaemia and cardiometabolic disease and provide insight in better preventive strategies. As methylation of DNA, in contrast to the DNA sequence, can be modified by several other environmental factors, the association between DNA methylation and cholesterol and TAG levels might be due to potential environmental influences, such as nutrition, smoking, and exercise.[65, 66] Since blood lipids are also influenced by many of these lifestyle factors, methylation may be a possible mediator between these external factors and cholesterol and TAG levels. New findings from studies on DNA methylation, lifestyle and lipids may thus provide further insight in underlying pathways and how to treat and prevent dyslipidaemia.

Conclusions

The present systematic review suggests that DNA methylation at specific genes is associated with blood cholesterol and TAG levels. However, most studies were small and of relatively small size and replication is lacking. Future studies using larger sample sizes, longitudinal data, relevant tissues, and

appropriate confounder adjustment are needed to establish whether differences in epigenetic marks are associated with changes in blood lipid levels and subsequently cardiometabolic disease risk.

REFERENCES

1. Otocka-Kmiecik A, Mikhailidis DP, Nicholls SJ, Davidson M, Rysz J, Banach M. Dysfunctional HDL: a novel important diagnostic and therapeutic target in cardiovascular disease? Progress in lipid research. 2012;51(4):314-24.
2. Fiorucci S, Cipriani S, Baldelli F, Mencarelli A. Bile acid-activated receptors in the treatment of dyslipidemia and related disorders. Progress in lipid research. 2010;49(2):171-85.
3. Yusuf S, Hawken S, Ôunpuu S, Dans T, Avezum A, Lanas F, et al. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. The Lancet. 2004;364(9438):937-52.
4. Mooradian AD. Dyslipidemia in type 2 diabetes mellitus. Nature clinical practice Endocrinology & metabolism. 2009;5(3):150-9.
5. Franklin BA, Durstine JL, Roberts CK, Barnard RJ. Impact of diet and exercise on lipid management in the modern era. Best Pract Res Clin Endocrinol Metab. 2014;28(3):405-21.
6. Sahebkar A, Chew GT, Watts GF. Recent advances in pharmacotherapy for hypertriglyceridemia. Progress in lipid research. 2014;56:47-66.
7. Global Lipids Genetics C. Discovery and refinement of loci associated with lipid levels. Nature genetics. 2013;45(11):1274-83.
8. Muka T, Koromani F, Portilla E, O'Connor A, Bramer WM, Troup J, et al. The role of epigenetic modifications in cardiovascular disease: A systematic review. International Journal of Cardiology. 2016;212: 174-83.
9. Muka T, Nano J, Voortman T, Braun KVE, Ligthart S, Stranges S, et al. The role of Global and Regional DNA Methylation and Histone Modifications in glycemic traits and Type 2 Diabetes: a Systematic Review. Nutrition, Metabolism and Cardiovascular Diseases. 2016.
10. Guay SP, Brisson D, Lamarche B, Gaudet D, Bouchar L. Epipolymorphisms within lipoprotein genes contribute independently to plasma lipid levels in familial hypercholesterolemia. Epigenetics. 2014; 9(5):718-29.
11. Pfeifferm L, Wahl S, Pilling LC, Reischl E, Sandling JK, Kunze S, et al. DNA Methylation of Lipid-Related Genes Affects Blood Lipid Levels. Circ Cardiovasc Genet. 2015.
12. Jiang D, Zheng D, Wang L, Huang Y, Liu H, Xu L, et al. Elevated PLA2G7 Gene Promoter Methylation as a Gender-Specific Marker of Aging Increases the Risk of Coronary Heart Disease in Females. PLoS ONE. 2013;8(3).
13. Feinberg AP. Epigenetics at the epicenter of modern medicine. Jama. 2008;299(11):1345-50.
14. Martin-Nunez GM, Rubio-Martin E, Cabrera-Mulero R, Rojo-Martinez G, Olveira G, Valdes S, et al. Type 2 diabetes mellitus in relation to global LINE-1 DNA methylation in peripheral blood: a cohort study. Epigenetics. 2014;9(10):1322-8.
15. Kim M, Long TI, Arakawa K, Wang R, Yu MC, Laird PW. DNA methylation as a biomarker for cardiovascular disease risk. PLoS One. 2010;5(3):e9692.
16. Feinberg AP. Epigenetics at the epicenter of modern medicine. Jama. 2008;299(11):1345-50.
17. Muka T, Koromani F, Portilla E, O'Connor A, Bramer WM, Troup J, et al. The role of epigenetic modifications in cardiovascular disease: A systematic review. Int J Cardiol. 2016;212:174-83.
18. Bollati V, Schwartz J, Wright R, Litonjua A, Tarantini L, Suh H, et al. Decline in genomic DNA methylation through aging in a cohort of elderly subjects. Mechanisms of ageing and development. 2009;130(4): 234-9.

19. Ehrlich M, Gama-Sosa MA, Huang L-H, Midgett RM, Kuo KC, McCune RA, et al. Amount and distribution of 5-methylcytosine in human DNA from different types of tissues or cells. Nucleic acids research. 1982;10(8):2709-21.
20. Wilson AS, Power BE, Molloy PL. DNA hypomethylation and human diseases. Biochimica et Biophysica Acta (BBA)-Reviews on Cancer. 2007;1775(1):138-62.
21. Weisenberger DJ, Campan M, Long TI, Kim M, Woods C, Fiala E, et al. Analysis of repetitive element DNA methylation by MethyLight. Nucleic acids research. 2005;33(21):6823-36.
22. Yang AS, Estécio MRH, Doshi K, Kondo Y, Tajara EH, Issa JPJ. A simple method for estimating global DNA methylation using bisulfite PCR of repetitive DNA elements. Nucleic acids research. 2004;32(3): e38-e.
23. Mohsen K, Johansson S, Ekström TJ. Using LUMA: a Luminometric-based assay for global DNA-methylation. Epigenetics. 2006;1(1):46-9.
24. Pearce MS, McConnell JC, Potter C, Barrett LM, Parker L, Mathers JC, et al. Global LINE-1 DNA methylation is associated with blood glycaemic and lipid profiles. Int J Epidemiol. 2012;41(1):210-7.
25. Nicoletti CF, Nonino CB, de Oliveira BAP, Pinhel MADS, Mansego ML, Milagro FI, et al. DNA Methylation and Hydroxymethylation Levels in Relation to Two Weight Loss Strategies: Energy-Restricted Diet or Bariatric Surgery. Obes Surg. 2015.
26. Turcot V, Tchernof A, Deshaies Y, Perusse L, Belisle A, Marceau S, et al. LINE-1 methylation in visceral adipose tissue of severely obese individuals is associated with metabolic syndrome status and related phenotypes. Clin Epigenetics. 2012;4(1):10.
27. Kato S, Lindholm B, Stenvinkel P, Ekstrom TJ, Luttropp K, Yuzawa Y, et al. DNA hypermethylation and inflammatory markers in incident Japanese dialysis patients. Nephron Extra. 2012;2(1):159-68.
28. Nicoletti CF, Nonino CB, de Oliveira BAP, de Souza Pinhel MA, Mansego ML, Milagro FI, et al. DNA methylation and hydroxymethylation levels in relation to two weight loss strategies: energy-restricted diet or bariatric surgery. Obesity surgery. 2015:1-9.
29. Cash HL, McGarvey ST, Houseman EA, Marsit CJ, Hawley NL, Lambert-Messerlian GM, et al. Cardiovascular disease risk factors and DNA methylation at the LINE-1 repeat region in peripheral blood from Samoan Islanders. Epigenetics. 2011;6(10):1257-64.
30. Houde AA. Placental lipoprotein lipase DNA methylation levels are associated with gestational diabetes mellitus and maternal and cord blood lipid profiles. J Dev Orig Health Dis. 2014;5(2):132-41.
31. Houde AA, Guay SP, Desgagne V, Hivert MF, Baillargeon JP, St-Pierre J, et al. Adaptations of placental and cord blood ABCA1 DNA methylation profile to maternal metabolic status. Epigenetics. 2013;8(12): 1289-302.
32. Turcot V, Tchernof A, Deshaies Y, Perusse L, Belisle A, Marceau P, et al. Comparison of the dipeptidyl peptidase-4 gene methylation levels between severely obese subjects with and without the metabolic syndrome. Diabetol Metab Syndr. 2013;5(1).
33. Tang L, Wang L, Ye H, Xu X, Hong Q, Wang H, et al. BCL11A gene DNA methylation contributes to the risk of type 2 diabetes in males. Exp Ther Med. 2014;8(2):459-63.
34. Bollati V, Favero C, Albetti B, Tarantini L, Moroni A, Byun HM, et al. Nutrients intake is associated with DNA methylation of candidate inflammatory genes in a population of obese subjects. Nutrients. 2014; 6(10):4625-39.
35. Canivell S, Ruano EG, Siso-Almirall A, Kostov B, Gonzalez-De Paz L, Fernandez-Rebollo E, et al. Differential Methylation of TCF7L2 promoter in peripheral blood DNA in newly diagnosed, drug-naive patients with type 2 diabetes. PLoS ONE. 2014;9(6).

36. Ghattas M, El-shaarawy F, Mesbah N, Abo-Elmatty D. DNA methylation status of the methylenetetrahydrofolate reductase gene promoter in peripheral blood of end-stage renal disease patients. Mol Biol Rep. 2014;41(2):683-8.
37. Ma Y, Smith CE, Lai CQ, Irvin MR, Parnell LD, Lee YC, et al. Genetic variants modify the effect of age on APOE methylation in the Genetics of Lipid Lowering Drugs and Diet Network study. Aging Cell. 2015;14(1):49-59.
38. Peng P, Wang L, Yang X, Huang X, Ba Y, Chen X, et al. A preliminary study of the relationship between promoter methylation of the ABCG1, GALNT2 and HMGCR genes and coronary heart disease. PLoS ONE. 2014;9(8).
39. Houde AA, Légaré C, Biron S, Lescelleur O, Biertho L, Marceau S, et al. Leptin and adiponectin DNA methylation levels in adipose tissues and blood cells are associated with BMI, waist girth and LDL-cholesterol levels in severely obese men and women. BMC Med Genet. 2015;16(1).
40. Wu L, Zhao X, Shen Y, Zhang MX, Yan Y, Hou D, et al. Promoter methylation of fas apoptotic inhibitory molecule 2 gene is associated with obesity and dyslipidaemia in Chinese children. Diab Vasc Dis Res. 2015.
41. Afzali M, Nakhaee A, Tabatabaei SP, Tirgar-Fakheri K, Hashemi M. Aberrant promoter methylation profile of niemann-pick type C1 gene in cardiovascular disease. Iran Biomed J. 2013;17(2):77-83.
42. Guay SP, Brisson D, Mathieu P, Bosse Y, Gaudet D, Bouchard L. A study in familial hypercholesterolemia suggests reduced methylomic plasticity in men with coronary artery disease. Epigenomics. 2015; 7(1):17-34.
43. Xu L, Zheng D, Wang L, Jiang D, Liu H, Xu L, et al. GCK gene-body hypomethylation is associated with the risk of coronary heart disease. BioMed Res Int. 2014;2014.
44. Deodati A, Inzaghi E, Liguori A, Puglianiello A, Germani D, Brufani C, et al. IGF2 methylation is associated with lipid profile in obese children. Horm Res Paediatr. 2013;79(6):361-7.
45. Liu ZH, Chen LL, Deng XL, Song HJ, Liao YF, Zeng TS, et al. Methylation status of CpG sites in the MCP-1 promoter is correlated to serum MCP-1 in type 2 diabetes. J Endocrinol Invest. 2012;35(6): 585-9.
46. Guay SP, Brisson D, Munger J, Lamarche B, Gaudet D, Bouchard L. ABCA1 gene promoter DNA methylation is associated with HDL particle profile and coronary artery disease in familial hypercholesterolemia. Epigenetics. 2012;7(5):464-72.
47. Guay SP, Brisson D, Lamarche B, Marceau P, Vohl MC, Gaudet D, et al. DNA methylation variations at CETP and LPL gene promoter loci: New molecular biomarkers associated with blood lipid profile variability. Atherosclerosis. 2013;228(2):413-20.
48. Turcot V, Bouchard L, Faucher G, Tchernof A, Deshaies Y, Perusse L, et al. DPP4 gene DNA methylation in the omentum is associated with its gene expression and plasma lipid profile in severe obesity. Obesity. 2011;19(2):388-95.
49. Irvin MR, Zhi D, Joehanes R, Mendelson M, Aslibekyan S, Claas SA, et al. Epigenome-wide association study of fasting blood lipids in the genetics of lipid-lowering drugs and diet network study. Circulation. 2014;130(7):565-72.
50. Guay SP, Voisin G, Brisson D, Munger J, Lamarche B, Gaudet D, et al. Epigenome-wide analysis in familial hypercholesterolemia identified new loci associated with high-density lipoprotein cholesterol concentration. Epigenomics. 2012;4(6):623-39.
51. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nature reviews Genetics. 2012;13(7):484-92.
52. Aran D, Toperoff G, Rosenberg M, Hellman A. Replication timing-related and gene body-specific methylation of active human genes. Hum Mol Genet. 2011;20(4):670-80.

53. Strahl BD, Allis CD. The language of covalent histone modifications. Nature. 2000;403(6765):41-5.
54. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet. 2003;33 Suppl:245-54.
55. Post WS, Goldschmidt-Clermont PJ, Wilhide CC, Heldman AW, Sussman MS, Ouyang P, et al. Methylation of the estrogen receptor gene is associated with aging and atherosclerosis in the cardiovascular system. Cardiovasc Res. 1999;43(4):985-91.
56. Gelissen IC, Harris M, Rye KA, Quinn C, Brown AJ, Kockx M, et al. ABCA1 and ABCG1 synergize to mediate cholesterol export to apoA-I. Arterioscler Thromb Vasc Biol. 2006;26(3):534-40.
57. McGarry JD, Brown NF. The mitochondrial carnitine palmitoyltransferase system. From concept to molecular analysis. Eur J Biochem. 1997;244(1):1-14.
58. Gerin I, Clerbaux LA, Haumont O, Lanthier N, Das AK, Burant CF, et al. Expression of miR-33 from an SREBP2 intron inhibits cholesterol export and fatty acid oxidation. J Biol Chem. 2010;285(44): 33652-61.
59. Gurevich I, Aneskievich BJ. Liganded RARalpha and RARgamma interact with but are repressed by TNIP1. Biochem Biophys Res Commun. 2009;389(3):409-14.
60. Houde A-A, Légaré C, Biron S, Lescelleur O, Biertho L, Marceau S, et al. Leptin and adiponectin DNA methylation levels in adipose tissues and blood cells are associated with BMI, waist girth and LDL-cholesterol levels in severely obese men and women. BMC medical genetics. 2015;16(1):1.
61. Guay SP, Brisson D, Lamarche B, Biron S, Lescelleur O, Biertho L, et al. ADRB3gene promoter DNA methylation in blood and visceral adipose tissue is associated with metabolic disturbances in men. Epigenomics. 2014;6(1):33-43.
62. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC bioinformatics. 2012;13(1): 86.
63. Pfeifferm L, Wahl S, Pilling LC, Reischl E, Sandling JK, Kunze S, et al. DNA Methylation of Lipid-Related Genes Affects Blood Lipid Levels. Circulation: Cardiovascular Genetics. 2015:CIRCGENETICS. 114.000804.
64. Heyn H, Moran S, Hernando-Herraez I, Sayols S, Gomez A, Sandoval J, et al. DNA methylation contributes to natural human variation. Genome Res. 2013;23(9):1363-72.
65. Anderson OS, Sant KE, Dolinoy DC. Nutrition and epigenetics: an interplay of dietary methyl donors, one-carbon metabolism and DNA methylation. J Nutr Biochem. 2012;23(8):853-9.
66. Milagro FI, Mansego ML, De Miguel C, Martinez JA. Dietary factors, epigenetic modifications and obesity outcomes: progresses and perspectives. Mol Aspects Med. 2013;34(4):782-812.
67. Guay SP, Legare C, Houde AA, Mathieu P, Bosse Y, Bouchard L. Acetylsalicylic acid, aging and coronary artery disease are associated with ABCA1 DNA methylation in men. Clin Epigenetics. 2014;6(1):14.
68. Irvin MR, Zhi D, Joehanes R, Mendelson M, Aslibekyan S, Claas SA, et al. Epigenome-wide association study of fasting blood lipids in the genetics of lipid-lowering drugs and diet network study. Circulation. 2014;130(7):565-72.

SUPPLEMENTARY MATERIAL

Supplement 3.1.1. Search strategy

Embase.com

(epigenetics/exp OR 'DNA methylation'/exp OR 'histone modification'/exp OR 's adenosylmethionine'/exp OR 'CpG island'/exp OR (((histone* OR dna OR 'long interspersed') NEAR/3 (acetylat* OR demethylat* OR methylat* OR phosphorylat* OR ubiquitinat* OR modif*)) OR 's adenosylmethionine' OR cpg OR epigenetic* OR epigenomic*):ab,ti) AND ('lipid blood level'/exp OR (((lipid* OR cholester* OR triacylglycerol* OR triglyceride* OR HDL* OR LDL* OR VLDL* OR VHDL*) NEAR/6 (level* OR blood OR serum OR plasma* OR concentration*)) OR hypercholesterol*):ab,ti) NOT ([animals]/lim NOT [humans]/lim) NOT ([Conference Abstract]/lim OR [Letter]/lim OR [Note]/lim OR [Editorial]/lim)

Medline ovid

(Epigenomics/ OR DNA methylation/ OR S-Adenosylmethionine/ OR CpG Islands/ OR (((histone* OR dna OR long interspersed) ADJ3 (acetylat* OR demethylat* OR methylat* OR phosphorylat* OR ubiquitinat* OR modif*)) OR s adenosylmethionine OR cpg OR epigenetic*).ab,ti.) AND (exp lipids/ bl OR (((lipid* OR cholester* OR triacylglycerol* OR triglyceride* OR HDL* OR LDL* OR VLDL* OR VHDL*) ADJ6 (level* OR blood OR serum OR plasma* OR concentration*)) OR hypercholesterol*).ab,ti.) NOT (exp animals/ NOT humans/) NOT (letter OR news OR comment OR editorial OR congresses OR abstracts).pt.

Cochrane

((((histone* OR dna OR 'long interspersed') NEAR/3 (acetylat* OR demethylat* OR methylat* OR phosphorylat* OR ubiquitinat* OR modif*)) OR 's adenosylmethionine' OR cpg OR epigenetic* OR epigenomic*):ab,ti) AND ((((lipid* OR cholester* OR triacylglycerol* OR triglyceride* OR HDL* OR LDL* OR VLDL* OR VHDL*) NEAR/6 (level* OR blood OR serum OR plasma* OR concentration*)) OR hypercholesterol*):ab,ti)

Web-of-science

TS=(((((histone* OR dna OR "long interspersed") NEAR/2 (acetylat* OR demethylat* OR methylat* OR phosphorylat* OR ubiquitinat* OR modif*)) OR "s adenosylmethionine" OR cpg OR epigenetic* OR epigenomic*)) AND ((((lipid* OR cholester* OR triacylglycerol* OR triglyceride* OR HDL* OR LDL* OR VLDL* OR VHDL*) NEAR/5 (level* OR blood OR serum OR plasma* OR concentration*)) OR hypercholesterol*)) NOT ((animal* OR rat OR rats OR mouse OR mice OR murine) NOT human*)) AND dt=(article)

Pubmed publisher

(Epigenomics[mh] OR DNA methylation[mh] OR S-Adenosylmethionine[mh] OR CpG Islands[mh] OR (((histone*[tiab] OR dna OR long interspersed) AND (acetylat*[tiab] OR demethylat*[tiab] OR methylat*[tiab] OR phosphorylat*[tiab] OR ubiquitinat*[tiab] OR modif*[tiab])) OR "s adenosylme-

thionine" OR cpg OR epigenetic*[tiab])) AND (lipids/bl[mh] OR (((lipid*[tiab] OR cholester*[tiab] OR triacylglycerol*[tiab] OR triglyceride*[tiab] OR HDL*[tiab] OR LDL*[tiab] OR VLDL*[tiab] OR VHDL*[tiab]) AND (level*[tiab] OR blood OR serum OR plasma*[tiab] OR concentration*[tiab])) OR hypercholesterol*[tiab])) NOT (animals[mh] NOT humans[mh]) NOT (letter[pt] OR news[pt] OR comment[pt] OR editorial[pt] OR congresses[pt] OR abstracts[pt]) AND (publisher[sb] OR inprocess [sb])

Google scholar

"histone|dna methylation|modification"|epigenetics "lipid|cholesterol|triacylglycerol|triglyceride level|levels|blood|serum|plasma|concentration"|"blood|plasma lipids|cholesterol"|hypercholesterolemia|hypercholesterolaemia

Supplement 3.1.2. Selection criteria

Inclusion criteria:

1. Include randomized controlled trials studies, clinical trials, intervention studies, cohort, case-control or cross-sectional studies.
2. Include studies that evaluate DNA methylation (either global, site-specific or genome-wide methylation of DNA) and histone modifications
3. Include studies that evaluate the association of DNA-methylation or histone modifications with at least one of the following outcomes:
   - Hypercholesterolemia
   - Total Cholesterol
   - HDL-cholesterol
   - Triglycerides
   - LDL
4. Include studies conducted in humans.
5. Include studies conducted in all ages
6. No language or date restriction.

Exclusion criteria

1. Exclude abstracts, cost effectiveness studies, letters to the editor, conference proceedings, systematic reviews or meta-analyses.
2. Exclude studies conducted in animals.

3.2

Epigenome-wide association study (EWAS) on lipids: the Rotterdam Study.

Based on:

Braun KVE, Dhana K, de Vries PS, Voortman T, van Meurs JBJ, Uitterlinden AG, Hofman A, Hu FB, BIOS consortium, Franco OH, Dehghan A. Epigenome-wide association study (EWAS) on lipids: the Rotterdam Study. *Clinical Epigenetics. 2017;9:15.*

ABSTRACT

Background: DNA methylation is a key epigenetic mechanism that is suggested to be associated with blood lipid levels. We aimed to identify CpG sites at which DNA methylation levels are associated with blood levels of triglycerides, high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), and total cholesterol in 725 participants of the Rotterdam Study, a population-based cohort study. Subsequently, we sought replication in a non-overlapping set of 760 participants.

Results: Genome-wide methylation levels were measured in whole blood using the Illumina Methylation 450 array. Associations between lipid levels and DNA methylation beta-values were examined using linear mixed-effect models. All models were adjusted for sex, age, smoking, white blood cell proportions, array number, and position on array. A Bonferroni corrected p-value lower than 1.08×10^{-7} was considered statistically significant. Five CpG sites annotated to genes including *DHCR24*, *CPT1A*, *ABCG1*, and *SREBF1*, were identified and replicated. Four CpG sites were associated with triglycerides, including CpG sites annotated to *CPT1A* (cg00574958 and cg17058475), *ABCG1* (cg06500161), and *SREBF1* (cg11024682). Two CpG sites were associated with HDL-C, including *ABCG1* (cg06500161), and *DHCR24* (cg17901584). No significant associations were observed with LDL-C or total cholesterol.

Conclusions: We report an association of HDL-C levels with methylation of a CpG site near *DHCR24*, a protein coding gene involved in cholesterol biosynthesis, which has previously been reported to be associated with other metabolic traits. Furthermore, we confirmed previously reported associations of methylation of CpG sites within *CPT1A*, *ABCG1*, and *SREBF1* and lipids. These results provide insight in the mechanisms that are involved in lipid metabolism.

BACKGROUND

Genetics is an important determinant of lipid levels which may affect lipid levels by changing expression levels of the genes.[1] Gene expression levels, however, are also regulated by DNA methylation, which is one of the most studied mechanism in the field of epigenetics, and may therefore have an effect on lipid levels.[2] In contrast to the DNA sequence, DNA methylation is dynamic over time and responsive to the environment, therefore, DNA methylation could also change in response to blood lipid levels.[3]

A few studies using candidate gene approaches have reported that DNA methylation at several loci, such as *APOE* and *ABCA1*, are associated with lipid levels.[4, 5] In addition, epigenome-wide association studies (EWAS) have recently become available, providing the possibility to identify associations between blood lipid levels and DNA methylation at novel loci.[3] To date, EWAS have identified associations between differentially methylated genes at a few novel loci, such as *TNNT1*, *CPT1A*, and *ABCG1*, and blood lipid levels.[6-8] However, so far most studies investigating the association between DNA methylation and lipids have been performed in patient populations, while only one study has been performed within a population-based study. As DNA methylation may vary across different states of health, further population-based studies are needed to explore these associations in the general population.

In this study we aimed to investigate the association between blood DNA methylation levels and blood levels of triglycerides, high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), and total cholesterol in of 725 participants of the Rotterdam Study, a population-based cohort study. Subsequently, we sought replication in a non-overlapping set of 760 participants.

RESULTS

Participant characteristics

Participant characteristics of the discovery cohort (n=725) and the replication cohort (n=760) are presented in **Table 3.2.1**. Levels of triglycerides and total cholesterol were similar in both cohorts. Mean levels of HDL were slightly lower in the discovery cohort compared to the replication cohort (1.4 vs. 1.5 mmol/L, p<0.001). Mean levels of LDL-C were higher in the discovery cohort compared to the replication cohort (3.9 vs. 3.7 mmol/L, p=0.02). The mean age was significantly higher (p<0.001) in the replication cohort (59.9 ± 8.2 years) compared to the discovery cohort (67.7 ± 5.9 years). In the discovery cohort, 26% of the population were current smokers, whereas in the replication cohort, this was 10% (p<0.001).

Table 3.2.1. Participant characteristics

	Discovery[1]	Replication[1]	P value[2]
N	725	760	
Gender (male)	336 (46%)	324 (42%)	0.14
Age (years)	59.9 ± 8.2	67.7 (5.9)	<0.001
BMI (kg/m^2)	27.6 ± 4.6	27.8 (4.2)	0.48
Obesity (BMI>30)	173 (24%)	194 (26%)	0.54
Waist circumference	93.7 ± 12.8	94.4 ± 12.0	0.318
Current smoking (yes)	197 (27%)	79 (10%)	<0.001
Triglycerides (mmol/L)	1.3 [0.9-1.8]	1.3 [1.0-1.7]	0.66
HDL-cholesterol (mmol/L)	1.4 (0.41)	1.5 (0.44)	<0.001
LDL-cholesterol (mmol/L)	3.9 (1.00)	3.7 (0.95)	0.001
Total cholesterol (mmol/L)	5.6 (1.07)	5.5 (1.03)	0.26
Lipid lowering medication (yes)	191 (26%)	238 (31%)	<0.001
CHD	42 (6%)	61 (8%)	0.003
DM	72 (10%)	94 (12%)	0.13

[1] *Values are presented as mean ± SD, median [IQR], or N (%).*
[2] *Characteristics of the discovery cohort and replication cohort were compared with ANOVA*

Discovery panel

The associations between DNA methylation probes and blood lipid levels are presented in Manhattan plots (**Figure 3.2.1** and **Figure 3.2.2**). **Table 3.2.2** shows the Bonferroni significant CpGs. We identified five CpG sites associated with triglyceride levels. These CpG sites were annotated to *CPT1A* (cg00574958 and cg17058475), *ABCG1* (cg06500161), *SREBF1* (cg11024682), and *DHCR24* (cg17901584). CpG sites annotated to *CPT1A* and *DHCR24* were negatively associated with triglycerides, whereas CpG sites annotated to *ABCG1* and *SBREF1* were positively associated with triglycerides. We identified three CpG sites associated with HDL-C. These CpG sites were annotated to *ABCG1* (cg06500161), *DHCR24* (cg17901584), and one CpG site that was not annotated to a gene (cg14816825). The CpG site annotated to *DHCR24* was positively associated with HDL-C, the other two CpG sites were negatively associated with HDL-C. We did not find significantly associated CpG sites for LDL-C and total cholesterol levels in the discovery cohort.

Replication panel

Of the five CpG sites significantly associated with triglycerides in the discovery cohort, four replicated in the replication cohort, including *CPT1A* (cg00574958 and cg17058475), *ABCG1* (cg06500161), and *SREBF1* (cg11024682). Of the three CpG sites significantly associated with HDL-C in the discovery cohort, two replicated in the replication cohort, including *ABCG1* (cg06500161), and *DHCR24* (cg17901584). Results from the discovery and replication cohort were combined using fixed effects meta-analyses.

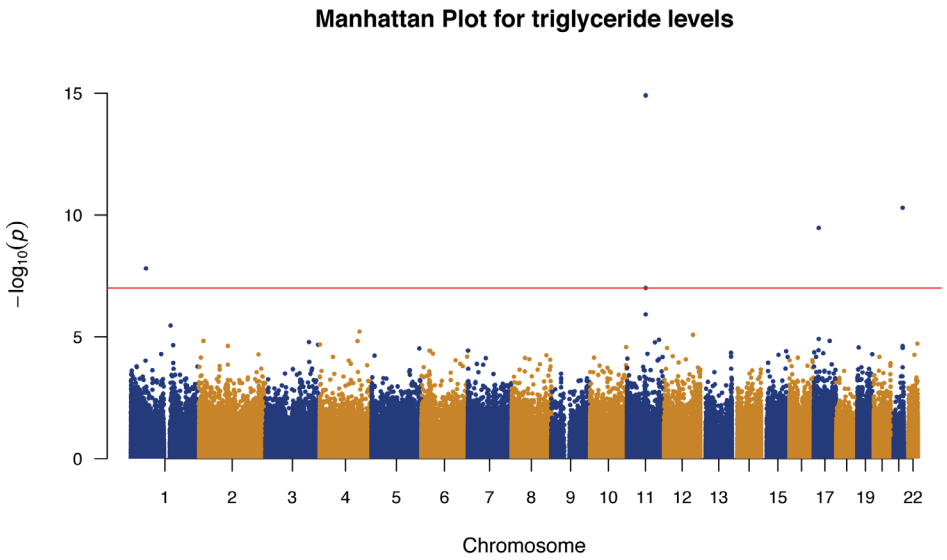


Figure 3.2.1. Manhattan plot epigenome-wide associations between genome wide DNA methylation and triglycerides.

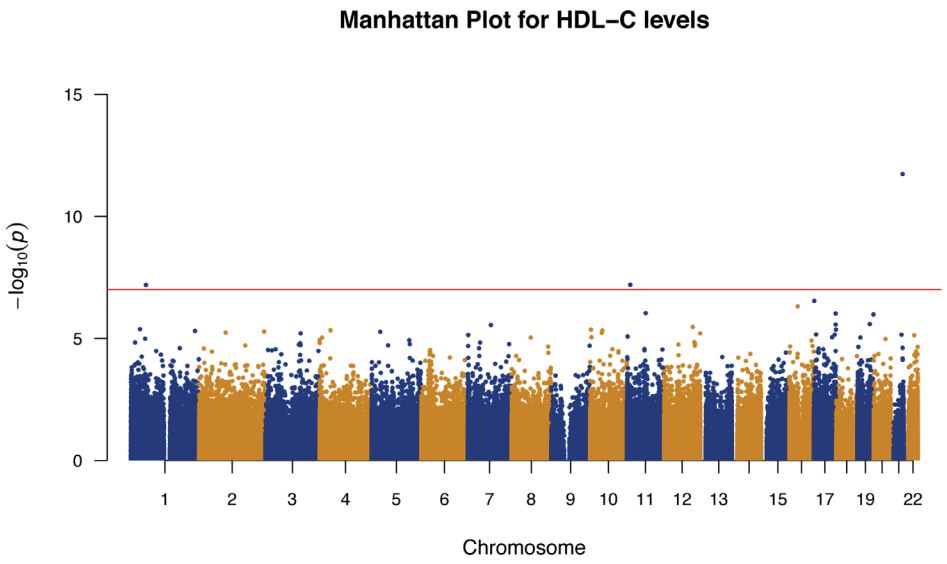


Figure 3.2.2. Manhattan plot epigenome-wide associations between genome wide DNA methylation and HDL-C.

Table 3.2.2. Epigenome-wide associations between genome wide DNA methylation and lipid levels.[1]

				Discovery		Replication		
	CpG site	Chromosome	Position	Regression coefficient	P[2]	Regression coefficient	P[2]	Gene
Triglycerides	cg00574958	11	68607622	-0.0206	**1.23×10^{-15}**	-0.0114	**7.24×10^{-13}**	*CPT1A*
	cg06500161	21	43656587	0.0149	**5.02×10^{-11}**	0.0201	**1.46×10^{-15}**	*ABCG1*
	cg11024682	17	17730094	0.0135	**3.40×10^{-10}**	0.0172	**3.27×10^{-11}**	*SREBF1*
	cg17901584	1	55353706	-0.0162	**1.56×10^{-08}**	-0.0068	9.54×10^{-02}	*DHCR24*
	cg17058475	11	68607737	-0.0183	**9.84×10^{-08}**	-0.0118	**1.71×10^{-06}**	*CPT1A*
HDL-C	cg06500161	21	43656587	-0.0202	**1.84×10^{-12}**	-0.0189	**3.29×10^{-12}**	*ABCG1*
	cg14816825	11	12128203	-0.0143	**6.31×10^{-08}**	-0.0051	1.15×10^{-01}	*NA*[3]
	cg17901584	1	55353706	0.0196	**6.44×10^{-08}**	0.0142	**1.03×10^{-03}**	*DHCR24*

Values are regression coefficients based on linear mixed models and reflect differences in methylation beta values per increase in HDL-C or log transformed triglycerides unit.
[1]Models are adjusted for age, gender, current smoking, leukocyte proportions, array number, and position on array.
[2]Level of significance: $p<1.08\times10^{-07}$ (discovery cohort), $p<7.14\times10^{-03}$ (replication).
[3]Not annotated.

Meta analyses

In order to test potential confounding, additional models with adjustment for lipid-lowering medication, waist circumference, and other lipids were performed in the combined analyses. Some of the effect estimates decreased in strength, but overall the results remained similar to those of model 1 (**Table 3.2.3**). Furthermore, results from meta-analyses revealed seven new CpG sites associated with triglycerides, including *TXNIP*, *TMEM49*, *SLC7A11*, and *KCNA3*. An additional 55 CpG sites were associated with HDL-C, Methylation of four CpG sites were associated with total cholesterol, including, CpGs annotated to *IFFO1*, *ABCG1*, and *DHCR24* **(Supplementary Table 3.2.1).**

Additional analyses

Since lipid levels might be affected by dietary fat intake, we explored whether intake of total fat, poly-unsaturated fatty acids (PUFA), mono-unsaturated fatty acids (MUFA) and saturated fatty acids (SFA) was associated with DNA methylation of significantly replicated CpG sites, using linear regression models. From these models we observed no significant association between fat intake and methylation of *DHCR24*, *SREBF1*, *ABCG1*, or *CPT1A* **(Supplementary Table 3.2.2).** To test if there was an interaction between lipid-lowering medication or fat intake and the CpG site located in the *DHCR24* gene on blood lipid levels, interaction terms were added to the regression model. However, none of these interaction terms were significant. Sensitivity analyses in which we replaced beta-values with M-values showed similar results (**Supplementary Table 3.2.3**).

Table 3.2.3. Associations between genome wide DNA methylation and lipid levels (meta-analyses).[1]

		Model 2[2]		Model 3[3]		Model 4[4]		
	ProbeID	Regression coefficient	P[5]	Regression coefficient	P[5]	Regression coefficient	P[5]	Gene
TG	cg00574958	-0.0142	2.6×10^{-26}	-0.0114	1.9×10^{-16}	-0.0161	1.5×10^{-22}	*CPT1A*
	cg06500161	0.0167	1.4×10^{-24}	0.0153	3.9×10^{-18}	0.0151	7.0×10^{-14}	*ABCG1*
	cg11024682	0.0147	3.0×10^{-19}	0.0124	8.4×10^{-13}	0.0138	1.0×10^{-11}	*SREBF1*
	cg17058475	-0.0146	3.1×10^{-13}	-0.0120	8.8×10^{-09}	-0.0196	9.2×10^{-16}	*CPT1A*
	cg17901584	-0.013	1.9×10^{-08}	-0.0104	2.5×10^{-05}	-0.012	2.5×10^{-05}	*DHCR24*
HDL-C	cg06500161	-0.0187	9.5×10^{-23}	-0.0166	9.5×10^{-16}	-0.0116	2.5×10^{-07}	*ABCG1*
	cg14816825	-0.0104	3.2×10^{-07}	-0.0145	3.8×10^{-07}	-0.0127	1.7×10^{-07}	*NA*[6]
	cg17901584	0.0164	2.4×10^{-09}	0.0128	1.3×10^{-05}	0.0116	3.6×10^{-04}	*DHCR24*

Values are regression coefficients based on linear mixed models and reflect differences in methylation beta values per increase in HDL-C or log transformed triglycerides unit.
[1]All models are adjusted for age, gender, current smoking, leukocyte proportions, array number, and position on array.
[2]Model 2: model1 + lipid lowering medication use.
[3]Model 3: model1 + waist circumference.
[4]Model 4: model1 + other lipids.
[5]Level of significance: $p<7.14\times10^{-03}$.
[6] Not annotated

Methylation risk scores

The methylation risk score was calculated using eight CpG sites for triglycerides and seven CpG sites for HDL-C, based on current and previously reported findings.(7, 8) Correlation coefficients of these CpG sites are presented in **Supplementary Table 3.2.4** and **Supplementary Table 3.2.5.** For triglycerides, 9% of the variance was explained by the methylation risk score. For HDL-C, 5 % of the variance was explained by the methylation risk score (**Supplementary Table 3.2.6**). To test whether the association between methylation risk scores and lipids differed by lipid lowering medication use, interaction terms were tested. For both triglycerides as HDL-C, none of the interaction terms were significant. The difference in levels of triglycerides and HDL-C per quartile of methylation risk score are presented in **Figure 3.2.3** and **Figure 3.2.4.** HDL-C levels decrease as quartiles of methylation risk score increase. Triglyceride levels increased from the first quartile to the second quartile, but remained similar for the third quartile and fourth quartile.

Expression levels

The associations between DNA methylation of ABCG1 and CPT1A and there expression levels are presented in **Supplementary Table 3.2.7**. Of the four genes of which the significantly associated CpG sites were annotated, only data on expression levels for *CPT1A* and *ABCG1* passed quality control. For *CPT1A*, expression (ILMN_1696316) was negatively associated with methylation of *CPT1A* at both identified CpG sites (cg00574958: $p=7.6\times10^{-4}$, cg17058475: $p= 5.1\times10^{-3}$). Expression of *ABCG1* at two of the five transcripts (ILMN_1794782 and ILMN_2329927) were negatively associated with methylation of the *ABCG1* CpG site ($p=6.6\times10^{-14}$ and 3.7×10^{-11}).

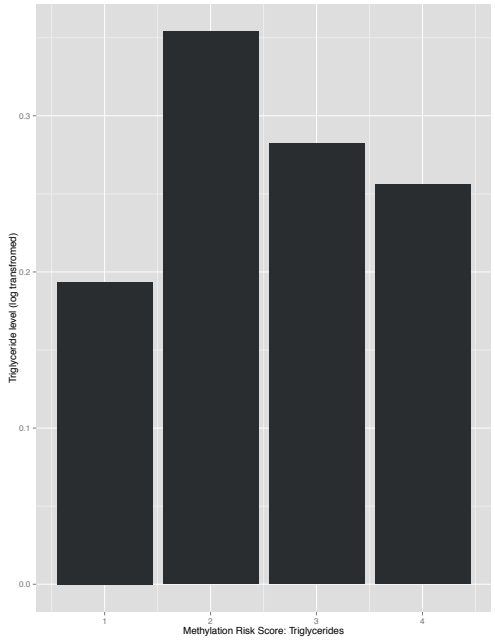


Figure 3.2.3. Methylation risk score in quartiles and levels of triglycerides.

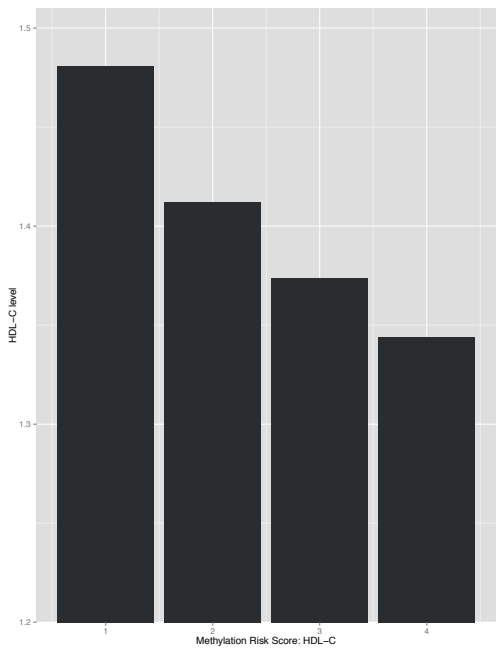


Figure 3.2.4. Methylation risk score in quartiles and levels of HDL-C.

DISCUSSION

The current study reports the results of an epigenome-wide association study of fasting triglycerides, HDL-C, LDL-C and total cholesterol blood levels. We identified and replicated a CpG site near *DHCR24* to be related with HDL-cholesterol. Moreover, we confirmed former reports that methylation of CpG sites within *ABCG1* are associated with HDL-C, and methylation of CpG sites within *CPT1A*, *ABCG1*, and *SREBF1* were associated with triglycerides. The majority of these associations were independent of several potentially confounding factors, such as use of lipid lowering medication, waist circumference, or other lipid traits. The differentially methylated CpG sites combined in a methylation risk score explained up to 9% of the variance in triglycerides and 5% of the variance in HDL-C.

We observed that methylation of a CpG site (cg17901584) annotated to the *DHCR24* gene was associated with HDL-cholesterol. This CpG site is located within 1 Kb upstream of the *DHCR24* gene (**Figure 3.2.5**). The *DHCR24* gene encodes for the cholesterol biosynthesis enzyme 3-hydroxysterol-24 reductase, which catalyzes the conversion of desmosterol to cholesterol.[9, 10] Mutations in the *DHCR24* gene may cause desmosterolosis, an autosomal recessive disease characterized by high levels of desmosterol.[11] Since *DHCR24* is involved in the cholesterol metabolism, an association between methylation in this gene and HDL-C is plausible. However, to date, it is not clear how methylation of *DHCR24* is involved in sterol regulation. It has been suggested that regulation of *DHCR24* expression is mediated by sterol-regulatory-element-binding-proteins (SREBP) in response to cholesterol availability.[12] We observed in meta-analysis of the discovery and replication cohort that methylation of a CpG site in the *SREBF1* gene, which encodes for *SREBP*, was associated with HDL-C levels. These findings suggest that mechanisms in which these two genes are involved might interact. Unfortunately, as the probes for these genes did not pass quality control, we did not have data on expression levels of these genes to further investigate this hypothesis. Moreover, the *DHCR24* gene is located near *PCSK9*, which is associated with cholesterol levels.[1] Therefore, the association we observed between *DHCR24* methylation and HDL-C might be due to SNP variation in *PCSK9*. However, when we adjusted our models on HDL-C for the top SNP from GWAS, the association remained similar (data not shown). To our knowledge, an association between methylation of cg17901584 (*DHCR24*) and HDL-C has not been previously reported. However, an EWAS performed by Dekkers et al. showed that another CpG site located in the *DHCR24* gene, cg2716885, was associated with LDL-C levels.[13] Although there are no previous studies that report an association between methylation of cg17901584 (*DHCR24*) and HDL-C, previous studies reported associations between methylation of this CpG site with waist circumference and phosphatidylcholine (PC ae C36:5).[14-16] In additional models we adjusted for waist circumference, where we indeed observed a decrease in strength compared to the main model. However, the association stayed significant. Two scenarios may explain these results. A higher waist circumference may affect lipid levels and consequently modify DNA methylation. Alternatively, waist circumference could be a confounding factor since metabolic traits are highly correlated. Due to the cross-sectional design of our study, it is difficult to make strong conclusions whether waist circumference is a confounder or a precursor in this association.

In addition to our novel finding, we also replicated findings previously reported by the GOLDN study.[7] In line with findings from GOLDN, we observed a significant negative association between

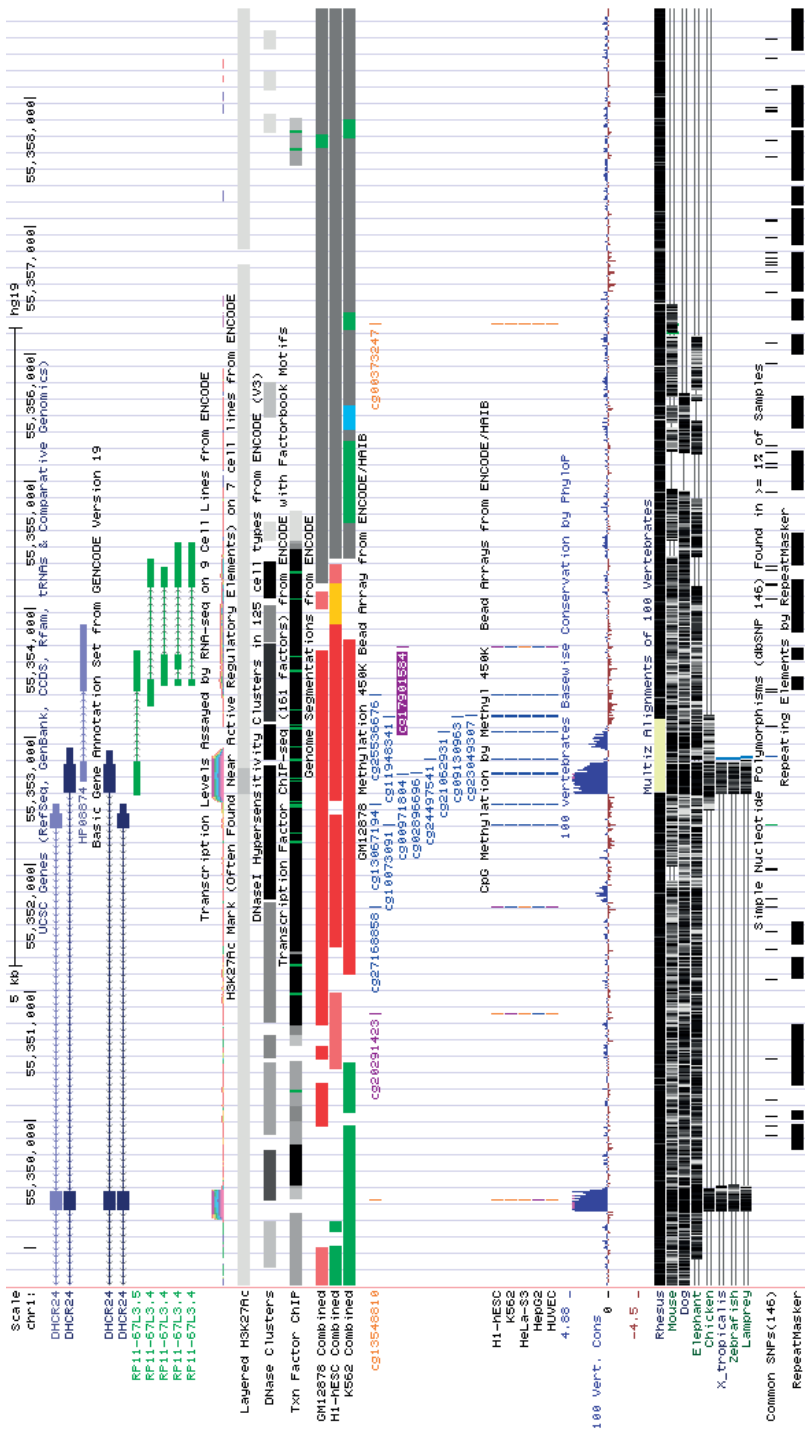


Figure 3.2.5. Annotation of cg17901584 to *DHCR24*

two CpG sites (cg00574958 and cg17058475) located in the *CPT1A* gene and triglycerides. Carnitine palmitoytransferase-1, which is encoded by the *CPT1A* gene, has an important role of transporting long-chain fatty acids in mitochondria. In order to obtain more knowledge on the epigenetic mechanisms of *CPT1A*, we explored expression levels of this gene in relation to methylation. Similar to findings from the GOLDN study, we observed a negative association between DNA methylation and expression levels of one probe, suggesting that increased methylation at *CPT1A* decreases expression of this gene.[7]

Furthermore, some of the findings previously reported by KORA were also replicated.[8] In line with findings from KORA, we also found a significant negative association between a CpG site (cg06500161), located in the *ABCG1* gene and HDL-C. We also found a significant positive association between CpG sites located in *ABCG1* (cg06500161) and *SBREF1* (cg11024682) and triglycerides. Although these findings regarding triglycerides correspond to findings from KORA, the CpG sites in *ABCG1* and *SBREF1* differed from the CpG sites found in our study. To further explore the epigenetic mechanisms of *ABCG1*, we investigated whether DNA methylation of the CpG sites at *ABCG1* was associated with gene expression. Similar to KORA, we observed a negative correlation between DNA methylation at CpG sites annotated to *ABCG1* with two mRNA expression probes.[8] In our study ILMN_1794782 showed the strongest negative correlation (r=-0.30) with methylation of *ABCG1* (cg06500161), which was also reported by the authors from KORA as one of the most strongly related transcripts to *ABCG1* methylation. These results are in agreement with a mediatory role of gene *ABCG1* expression in the association between DNA methylation and lipid levels. In addition to lipid levels, other studies have also demonstrated an association between methylation levels at CpG sites located in *CPT1A*, *ABCG1*, and *SREBF1* with other cardiometabolic traits, such as myocardial infarction, BMI, waist circumference, insulin and several metabolomic traits.[8, 14, 15, 17] These previously reported findings and our current results suggest that methylation of these genes are involved in metabolic mechanisms, and may be potential therapeutic targets for cardiovascular and metabolic related diseases.

Results from the discovery and replication cohort were combined in meta-analysis, which revealed a large amount of additional differentially methylated CpG sites that are associated with HDL-C levels. These additional findings include several CpG sites annotated to genes that were previously reported in GWAS on lipid levels, such as *LDLR* and *CMIP*.[1] *LDLR*, Low-Density Lipoprotein Receptor, is a protein coding gene, and mutations in this gene can cause familial hypercholesterolemia.[18] *CMIP* encodes a c-Maf inducing protein, which is involved in T-cell signaling pathway, and this gene is known to be associated with speech impairment.[19] However, less is known on the role of *CMIP* in cholesterol metabolism. Considering previously reported and current findings, there might be a genetic effect as well as an epigenetic effect of these genes on HDL-C levels. Nevertheless, these findings should be interpreted with caution, as results of our meta-analyses are not replicated by other studies yet.

This study has several strengths and limitations that should be considered with the interpretation of the currently reported results. The strength of this study is the availability of genome-wide DNA methylation data in a large sample of adults from the general population and an internal replication set. Although we had a large sample size, more loci can be identified with even larger sample sizes. To illustrate, when the results from the discovery and replication cohort were combined in meta-analysis, an additional set of 55 new significant associations for HDL-C were observed (**Supplemen-**

tary Table 3.2.1), which emphasizes the benefits of using a larger sample size in EWAS. A limitation of this study is the use of whole blood samples to determine DNA methylation levels, whereas DNA methylation is cell type specific. When methylation is studied as a consequence of lipid levels, then leukocytes are relevant tissue. However, when the aim is to study methylation sites that are causal of lipid levels, leukocytes might not be relevant tissue. Therefore, certain important CpG sites could have been overlooked in our study. Nevertheless, previous studies have demonstrated that results can be replicated across tissues, suggesting that the use of blood tissue is not necessarily a major issue.[14, 20] Moreover, associations may have been overlooked be due to the type of measurements used in our study. Levels of LDL-C in our study were calculated with use of the Friedewald formula,[21] which may not have been specific enough to identify an association between genome-wide DNA methylation and LDL-C. Results from another study performed within the GOLDN study showed that methylation of two CpG sites in the *CPT1A* gene are associated with LDL-C and VLDL-C.[22] These results overlap with our findings of triglycerides, but we did not identify CpG sites associated with LDL. This discrepancy in results may be due to a more detailed quantification of lipids by NMR spectroscopy in the GOLDN study. In our current study we used beta-values to analyze methylation levels. Since it has been suggested that there could be heteroscedasticity for CpG sites with very low or high methylation, the use of M-values is recommended. These M-values represent the Logit transformation of the beta-values and are considered to have a better detection rate and true positive rate.[23] In order to explore whether this could have affected our results, we performed additional analyses in which we replaced beta-values with M-values. However, these sensitivity analyses showed that results were similar for beta-values and M-values (**Supplementary Table 3.2.3**), which suggests that in our study heteroscedasticity may not be an issue. Due to the cross-sectional design we cannot determine the temporal direction of the association between DNA methylation and blood lipids. Furthermore, the observed associations could be explained by an unidentified common factor, as residual confounding is always an issue in observational studies. Another possibility is that the results may be confounded by differences in cell type proportion. In order to avoid this source of confounding, we adjusted all analyses for measured or estimated cell type proportions.[24] Finally, our replication cohort had a 7.8 years higher mean age compared to the discovery cohort, which could have resulted in differences in association between DNA methylation and lipid levels. However, when we tested our models for interaction with age we did not find any evidence that the strength of associations was affected by age.

CONCLUSION

In conclusion, we report an association of HDL-C levels with methylation of a CpG site near *DHCR24*, a protein coding gene involved in cholesterol biosynthesis. This CpG site has previously been reported to be associated with other metabolic traits, such as waist circumference. In addition, we replicated associations previously reported by other studies. These results provide insight in the mechanisms that are involved in cholesterol and lipid metabolism, and identify potential new therapeutic targets. Future studies should include a larger sample size and further investigate the independency and causality of the observed associations.

METHODS

Design and subjects

This study was embedded within the Rotterdam Study, a population-based cohort study in Rotterdam, the Netherlands. The design of The Rotterdam Study has been previously described in detail elsewhere.[25] Briefly, residents of Ommoord, a district in Rotterdam, aged 45 years and older were invited to participate. The Rotterdam Study includes three sub-cohorts. We used data from the baseline and second visit of the third cohort (RSIII-1 and RSIII-2), the third visit from the second cohort (RSII-3).

Discovery panel

We used the data from RSIII-1 as the discovery panel: between February 2006 and December 2008, 3,932 participants were examined. EWAS measurements were performed on a random subset of 731 subjects, of whom 725 had fasting blood samples available and were included in the current analyses.

Replication panel

We sought replication in a set of 767 participants from RSII-3 and RSIII-2. Between February 2011 and February 2012, 1,887 participants from RSII-3 were examined. Between March 2012 and December 2013, approximately 3,000 participants from RSIII-2 were examined. From the participants included in the replication study, 760 had fasting blood samples available and were included for analyses. None of the participants included in the replication study were included in the discovery cohort.

DNA methylation

DNA was extracted from whole blood (stored in EDTA tubes) by standardized salting out methods. Genome-wide methylation levels were measured using the Illumina Infinium Human Methylation450 Beadchip (Illumina Inc., San Diego, CA).[26] Briefly, samples (500 ng of DNA per sample) were bisulfite-treated with use of the Zymo EZ-96 DNA methylation kit (Zymo Research, Irvine, CA, USA). Thereafter, the samples were hybridized to the arrays according to the protocol of the manufacturer. During quality control in RSIII-1, samples showing incomplete bisulfite treatment were excluded (n=5) as were samples with a low detection rate (<99%) (n=7), and gender swaps (n=4). Probes with a detection p-value >0.01 in >1% of the samples, were filtered out.[27-29] In RSII-3 and RSIII-2, outlying samples were checked using the first two principal components obtained using principal component analysis (PCA). None of the samples failed the quality control checks, indicating high quality data. Per individual probe, participants with methylation levels higher than three times the inter-quartiles range (IQR) were excluded. The methylation proportion of a CpG site was reported as a beta-value ranging from 0 to 1.[30] We used the genome coordinates provided by Illumina (GRCh37/hg19) to identify independent loci.

mRNA expression data

Total RNA was isolated (PAXGene Blood RNA kits - Qiagen) from whole blood (PAXGene Tubes – Becton Dickinson). All RNA samples were analyzed using the Labchip GX (Calliper) according to manufacturer's instructions, to ensure a constant high quality of RNA preparations. Samples with an RNA quality score >7 were amplified and labeled (Ambion TotalPrep RNA) and hybridized to the

Illumina HumanHT12v4 Expression Beadchips, as described by the manufacturers protocol. RNA samples were processed at the Genetic Laboratory of Internal Medicine, Erasmus University Medical Centre Rotterdam. The dataset including 881 expression samples from RSIII-1 is available at GEO (Gene Expression Omnibus) public repository under the accession GSE338828. Gene expression data was quantile-normalized to the median distribution and log2-transformed. Probe and sample means were centered to zero. Genes were considered significantly expressed when detection p-values calculated by Genome Studio were less than 0.05 in more than 10% of all discovery samples, which added to a total number of 21,238 probes. The eQTL-mapping pipeline was used to perform quality control.[31]

Blood lipids

All participants had blood samples taken during the visits to the research center. From the blood samples, concentrations of triglycerides, HDL-C, and total cholesterol were measured using an automated enzymatic method. LDL-C was calculated using the Friedewald formula (total cholesterol – HDL-C – triglycerides/5).[21] Participants with non-fasting blood samples were excluded from the current analyses (n=6).

Covariates

Height and weight were measured during the centre visit and BMI was calculated (kg/m^2). During home visit interviews, data on tobacco smoking, dietary intake, and medication use were collected. Information on smoking history was acquired from questionnaires, and categorized as never, former or current smoking. Nutritional data was collected using semi-quantitative FFQs, and information on intake of different types of fatty acids were obtained. Fat intake was reported as total fat, PUFA, MUFA, and SFA. Information regarding the use of lipid lowering medication was derived from both structured home interviews and linkage to pharmacy records.

Statistical analysis

Triglyceride level was log-transformed using a natural-log to obtain a normal distribution. The associations between lipid levels and DNA methylation beta-values were examined using linear mixed-effect models.

Discovery

All models were adjusted for sex, age, smoking (current, former, or never), white blood cell proportions, and technical covariates (model 1). Gender, age, and smoking were added to the model as fixed effects. To correct for cell mixture distribution, leukocyte proportions (CD8+ T-cells, CD4+ T-cells, NK-cells, B-cells, monocytes, and granulocytes) were estimated using the Houseman method and were added to the model as fixed effects.[24] Technical covariates included array number and position on the array, and these were added to the models as random effects. To account for multiple testing, we used a Bonferroni corrected p-value of 1.08×10^{-7} (0.05/463,456 probes).

Replication

Identified probes were replicated using the same models as in the discovery cohort, further adjusted for the cohort. The adjustment for cell counts (monocytes, granulocytes and lymphocytes) was based on lab measurements rather than houseman estimates. For the replication we applied a Bonferroni-corrected significance threshold of 6.25×10^{-3} (0.05/8 probes).

Meta-analyses

To combine results from the discovery and replication cohort, fixed effects meta-analyses were performed in METAL, using an inverse variance weighted method. In subsequent analyses, models were further adjusted for lipid-lowering medication use (model 2), waist circumference (model 3), and other lipids (model 4).

Additional analyses

Lipid levels may be affected by dietary fat intake, and this might be mediated through DNA methylation. Therefore, we explored associations between different types of fatty acid intake and DNA methylation at significantly replicated CpG sites. To account for potential measurement error and confounding by total energy intake, we used the residual method to adjust the fatty acid intake for total energy intake. Briefly, linear regression analyses were used with energy intake as independent variable and fatty acid intake as dependent variable to calculate energy-adjusted intake of individual fatty acids for each subject. We regressed out the estimated leukocyte proportions, age, sex, array number, and position on array on the beta-values of the CpG sites using linear mixed models. The associations between energy-adjusted fatty acid intakes and the residuals of the DNA methylation beta-values were examined using linear regression models. All models were adjusted for sex, age, total energy intake (kcal/d), and smoking. Furthermore, to test if there was an interaction between lipid-lowering medication or fat intake and methylation at novel loci on lipids, interaction terms were added to the regression model, with one of the lipid traits as the outcome. In *post-hoc* analyses, we adjusted our models on HDL-C for the *PCSK9* genotype, as *DHCR24* and *PCSK9* are located near each other. In these analyses, the top SNP from GWAS (rs17111503) was added this to our main model (model 1). As it is recommended to use M-values due to heteroscedasticity in beta-values, we performed additional analyses in which we replaced beta-values with M-values for comparison.[23]

Methylation risk score

A methylation risk score was calculated based on CpG sites that were associated with the phenotypes, using both newly identified CpG sites in the current study as well as the ones previously reported for the corresponding trait.[7,8] First, CpG sites were checked for correlation and CpGs were pruned giving priority to the most significant CpGs reported by the largest studies using a correlation coefficient cut-off of 0.6. Second, linear regression analyses were performed in the replication cohort, using the lipids as dependent variable and the included CpG sites as independent variables. Models were adjusted for age, sex, blood cell counts, and technical covariates. The effect estimates were used to build the methylation risk score using data from the discovery panel. With the use of linear regression models, we calculated the lipids variance explained by the methylation risk score.

Functional analyses

Considering that DNA methylation can have an effect on the expression of genes, we explored the association between DNA methylation at the statistically significant CpG sites identified by EWAS and mRNA expression of the corresponding genes. The DNA methylation proportions and mRNA expression levels of these genes were checked for association using linear mixed models, which were adjusted for age, sex, smoking, white blood cell proportions, and technical covariates (array number and position on array).

SUPPLEMENTARY MATERIAL

Supplementary Table 3.2.1. Statistically significant associations in meta-analyses from discovery and replication cohort between genome wide DNA methylation and lipid levels.

	ProbeID	chr	pos	Effect	P	Gene
Triglycerides	cg00574958	11	68607622	-0.0139	5.07×10^{-26}	*CPT1A*
	cg06500161	21	43656587	0.0173	1.42×10^{-25}	*ABCG1*
	cg11024682	17	17730094	0.0150	2.85×10^{-20}	*SREBF1*
	cg17901584	1	55353706	-0.0131	1.52×10^{-08}	*DHCR24*
	cg17058475	11	68607737	-0.0140	1.53×10^{-12}	*CPT1A*
	cg19693031	1	145441552	-0.0133	6.72×10^{-08}	*TXNIP*
	cg24174557	17	57903544	-0.0145	7.00×10^{-10}	*TMEM49*
	cg06690548	4	139162808	-0.0089	3.91×10^{-08}	*SLC7A11*
	cg27243685	21	43642366	0.0085	9.90×10^{-15}	*ABCG1*
	cg01176028	21	43653234	0.0123	1.22×10^{-08}	*ABCG1*
	cg13925011	1	111216387	-0.0097	1.26×10^{-08}	*KCNA3*
	cg06734985	11	26849501	-0.0129	5.19×10^{-08}	*NA*[3]
HDL-C	cg06500161	21	43656587	-0.0195	4.16×10^{-24}	*ABCG1*
	cg17901584	1	55353706	0.0170	5.78×10^{-10}	*DHCR24*
	cg06017212	17	1478463	-0.0109	1.67×10^{-09}	*SLC43A2*
	cg11969813	17	79816559	-0.0161	2.04×10^{-08}	*P4HB*
	cg21088259	17	81039990	-0.0121	4.92×10^{-09}	*METRNL*
	cg01101459	1	234871477	-0.0137	1.21×10^{-09}	*NA*[3]
	cg14468090	3	129032864	-0.0110	2.96×10^{-10}	*NA*[3]
	cg26313301	19	11219615	-0.0056	3.28×10^{-08}	*LDLR*
	cg08174890	4	9558077	0.0088	3.74×10^{-09}	*MIR548I2*
	cg03031932	16	81547138	-0.0106	2.82×10^{-11}	*CMIP*
	cg02578470	3	127320670	-0.0141	2.53×10^{-09}	*MCM2*
	cg07730360	3	128845626	-0.0110	2.91×10^{-09}	*NA*[3]
	cg08132940	7	1081526	-0.0160	2.46×10^{-08}	*C7orf50*
	cg27243685	21	43642366	-0.0105	2.79×10^{-16}	*ABCG1*
	cg11024682	17	17730094	-0.0137	1.10×10^{-12}	*SREBF1*
	cg07091481	10	82169149	-0.0131	5.26×10^{-08}	*C10orf58*
	cg07691624	3	9886198	-0.0083	7.04×10^{-09}	*RPUSD3*
	cg14939082	10	104535990	-0.0135	2.22×10^{-08}	*C10orf26*
	cg00521255	5	139726689	-0.0116	2.82×10^{-08}	*HBEGF*
	cg20214535	21	43619310	-0.0099	8.45×10^{-08}	*ABCG1*
	cg07565956	17	7381288	-0.0102	1.21×10^{-08}	*ZBTB4*
	cg20605134	6	15400462	-0.0104	8.01×10^{-08}	*JARID2*
	cg25217710	1	156609523	-0.0086	1.87×10^{-08}	*NA*[3]
	cg00805360	10	135091210	-0.0142	2.49×10^{-10}	*ADAM8*

Supplementary Table 3.2.1. Statistically significant associations in meta-analyses from discovery and replication cohort between genome wide DNA methylation and lipid levels. (continued)

	ProbeID	chr	pos	Effect	P	Gene
	cg13876650	1	26146005	-0.0085	$1.30E\times10^{-08}$	*FAM54B*
	cg02797539	17	72740524	-0.0109	4.02×10^{-08}	*RAB37*
	cg06372475	10	73534286	-0.0080	6.08×10^{-08}	*C10orf54*
	cg25877299	17	80273290	-0.0142	4.26×10^{-09}	*CD7*
	cg04262505	2	217850623	0.0088	8.55×10^{-08}	*NA*[3]
	cg08352115	17	66356057	-0.0097	7.93×10^{-08}	*ARSG*
	cg21113318	14	92983645	-0.0096	9.72×10^{-09}	*RIN3*
	cg26943120	4	5472116	-0.0213	3.60×10^{-09}	*STK32B*
	cg11849692	10	103875969	-0.0099	3.48×10^{-08}	*LDB1*
	cg03074946	12	120687694	-0.0122	2.74×10^{-09}	*PXN*
	cg27269962	7	127540997	-0.0082	4.50×10^{-08}	*SND1*
	cg09473207	6	155079453	-0.0118	8.72×10^{-10}	*RBM16*
	cg07175797	16	50317656	-0.0109	9.64×10^{-11}	*NA*[3]
	cg27366162	17	66375195	-0.0128	1.72×10^{-08}	*ARSG*
	cg12179380	8	142084624	-0.0104	2.45×10^{-08}	*NA*[3]
	cg26468878	5	112501418	0.0098	7.36×10^{-08}	*MCC*
	cg09256683	17	34313637	0.0077	9.50×10^{-08}	*CCL14*
	cg06151145	4	48346434	-0.0136	2.44×10^{-08}	*SLAIN2*
	cg06710464	17	79047695	-0.0096	5.39×10^{-08}	*BAIAP2*
	cg23799393	10	60588674	0.0069	6.29×10^{-09}	*BICC1*
	cg04061506	2	48137261	-0.0115	4.20×10^{-08}	*NA*[3]
	cg05457221	10	134272437	-0.0090	5.42×10^{-08}	*NA*[3]
	cg26405097	6	15428301	-0.0157	1.88×10^{-09}	*JARID2*
	cg06876354	2	121020189	-0.0100	3.57×10^{-10}	*RALB*
	cg00819078	5	16615081	-0.0152	2.35×10^{-18}	*FAM134B*
	cg10127660	8	21996234	-0.0107	2.57×10^{-08}	*REEP4*
	cg05045598	1	6418521	-0.0123	4.58×10^{-08}	*ACOT7*
	cg15483436	14	105857208	-0.0068	1.23×10^{-09}	*PACS2*
	cg18805734	12	31453901	-0.0072	4.23×10^{-08}	*FAM60A*
	cg26640901	10	104536035	-0.0104	4.82×10^{-08}	*C10orf26*
	cg13139542	2	8242815	-0.0071	1.73×10^{-08}	*NA*[3]
	cg08788930	8	142201685	-0.0116	8.11×10^{-08}	*DENND3*
Total cholesterol	cg17281677	12	6658557	-0.0035	5.57×10^{-08}	*IFFO1*
	cg25536676	1	55353327	0.0044	7.77×10^{-08}	*DHCR24*
	cg06500161	21	43656587	-0.0043	3.66×10^{-08}	*ABCG1*
	cg17475467	1	55316769	-0.0023	8.97×10^{-08}	*DHCR24*

[1]Models adjusted for age, gender, current smoking, houseman estimated leukocyte proportions, array number, and position on array (model 1).

[2]Only hits that are statistically significant ($p<1.08\times10^{-07}$) in the combined analyses are shown in the table

[3] Not annotated.

Supplementary Table 3.2.2. Associations between DNA methylation at significant CpG sites and fatty acid intake.[1]

	ProbeID	Gene	Effect	P
Total Fat	cg17901584	*DHCR24*	-0.00002	0.78
	cg11024682	*SREBF1*	-0.00001	0.89
	cg06500161	*ABCG1*	0.00003	0.61
	cg00574958	*CPT1A*	-0.00011	0.10
	cg17058475	*CPT1A*	-0.00005	0.60
SFA[2]	cg17901584	*DHCR24*	-0.00011	0.53
	cg11024682	*SREBF1*	0.00003	0.80
	cg06500161	*ABCG1*	0.00008	0.55
	cg00574958	*CPT1A*	-0.00019	0.22
	cg17058475	*CPT1A*	0.00024	0.28
MUFA[2]	cg17901584	*DHCR24*	-0.00011	0.53
	cg11024682	*SREBF1*	-0.00009	0.48
	cg06500161	*ABCG1*	0.00004	0.77
	cg00574958	*CPT1A*	-0.00023	0.13
	cg17058475	*CPT1A*	-0.00019	0.38
PUFA[2]	cg17901584	*DHCR24*	-0.00011	0.55
	cg11024682	*SREBF1*	0.00000	0.98
	cg06500161	*ABCG1*	0.00003	0.82
	cg00574958	*CPT1A*	-0.00018	0.29
	cg17058475	*CPT1A*	-0.00036	0.13

Values are regression coefficients based on linear mixed models and reflect differences in methylation beta values per increase in triglycerides or HDL-C unit.

[1]Models adjusted for age, gender, current smoking, and total energy intake.

[2]PUFA, poly-unsaturated fatty acids; MUFA, mono-unsaturated fatty acids; SFA, saturated fatty acids.

Supplementary Table 3.2.3. Statistically significant associations from the discovery cohort between lipid levels and genome wide DNA methylation Beta-values compared to M-values.

				M-value[1,2]	
	ProbeID	**chr**	**pos**	**Effect[3]**	**P[4]**
TG	cg00574958	11	68607622	-0.1925	$2.09E\times10^{-16}$
	cg06500161	21	43656587	0.1042	$1.51E\times10^{-10}$
	cg11024682	17	17730094	0.0799	$2.96E\times10^{-10}$
	cg17058475	11	68607737	-0.1630	$2.59E\times10^{-08}$
	cg17901584	1	55353706	-0.1061	$2.64E\times10^{-08}$
HDL-C	cg06500161	21	43656587	-0.1406	$7.78E\times10^{-12}$
	cg14816825	11	12128203	-0.1123	$1.48E\times10^{-07}$
	cg17901584	1	55353706	0.1368	$1.46E\times10^{-08}$

[1]Values are regression coefficients based on linear mixed models and reflect differences in methylation M-values per increase in HDL-C or log transformed triglycerides unit.
[2]M-value=log2(Beta/(1-Beta))
[3]Models are adjusted for age, gender, current smoking, leukocyte proportions, array number, and position on array.
[4]Level of significance: $p<1.08\times10^{-07}$

Supplementary Table 3.2.4. Correlations between CpG sites included in the triglycerides methylation risk score.

	cg00574958	**cg07504977**	**cg20544516**	**cg07397296**	**cg07815238**	**cg06500161**	**cg11024682**
cg00574958	1	-0.07	-0.03	-0.04	0.12	-0.24	-0.15
cg07504977	-	1	-0.35	0.17	-0.35	0.02	0.03
cg20544516	-	-	1	0.14	0.43	0.24	0.30
cg07397296	-	-	-	1	0.18	0.22	0.15
cg07815238	-	-	-	-	1	0.19	0.17
cg06500161	-	-	-	-	-	1	0.32
cg11024682	-	-	-	-	-	-	1

Correlation coefficients were based on Pearson correlation r^2

Supplementary Table 3.2.5. Correlations between CpG sites included in the HDL-C methylation risk score.

	cg06500161	cg19693031	cg11024682	cg00574958	cg27243685	cg14816825	cg17901584
cg06500161	1	-0.07	0.32	-0.24	0.48	0.05	-0.01
cg19693031	-	1	0.15	0.24	-0.01	0.26	0.01
cg11024682	-	-	1	-0.15	0.43	0.35	0.26
cg00574958	-	-	-	1	-0.09	-0.13	0.15
cg27243685	-	-	-	-	1	-0.01	0.27
cg14816825	-	-	-	-	-	1	-0.32
cg17901584	-	-	-	-	-	-	1

Correlation coefficients were based on Pearson correlation r^2

Supplementary Table 3.2.6. Lipid variation explained by methylation risk score, age, and sex.[1]

	Triglycerides	HDL-C
MRS[2]	0.09	0.05
Age	0.002	0.001
Sex	0.02	0.17
Age + sex	0.026	0.173
MRS + age + sex	0.10	0.21

[1]Variance explained by the methylation scores (multiple R^2, adjusting for age and sex) was calculated using a linear regression models
[2]MRS, methylation risk score.

Supplementary Table 3.2.7. Correlations between DNA methylation at significant CpG sites and expression probes.[1]

CpG site[1]	Gene	Location[2]	Expression probe	N	Correlation coefficients	P
cg00574958	*CPT1A*	5'UTR	ILMN_1696316	713	-0.12	2.0×10^{-03}
cg17058475	*CPT1A*	5'UTR	ILMN_1696316	712	-0.09	1.4×10^{-02}
cg06500161	*ABCG1*	Body	ILMN_1695968	712	-0.02	5.6×10^{-01}
cg06500161	*ABCG1*	Body	ILMN_1794782	712	-0.30	2.2×10^{-16}
cg06500161	*ABCG1*	Body	ILMN_2262362	712	-0.06	1.1×10^{-01}
cg06500161	*ABCG1*	Body	ILMN_2329927	712	-0.18	1.9×10^{-06}

[1]Correlation coefficients are based on pearson correlation test of methylation beta values and expression probe unit.
[2] Annotation according to genome coordinates provided by Illumina (GRCh37/hg19)

REFERENCES

1. Global Lipids Genetics C, Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, et al. Discovery and refinement of loci associated with lipid levels. Nat Genet. 2013;45(11):1274-83.
2. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nature Reviews Genetics. 2012;13(7):484-92.
3. Rakyan VK, Down TA, Balding DJ, Beck S. Epigenome-wide association studies for common human diseases. Nature Reviews Genetics. 2011;12(8):529-41.
4. Guay S-P, Brisson D, Lamarche B, Gaudet D, Bouchard L. Epipolymorphisms within lipoprotein genes contribute independently to plasma lipid levels in familial hypercholesterolemia. Epigenetics. 2014; 9(5):718-29.
5. Ma Y, Smith CE, Lai CQ, Irvin MR, Parnell LD, Lee YC, et al. Genetic variants modify the effect of age on APOE methylation in the Genetics of Lipid Lowering Drugs and Diet Network study. Aging Cell. 2015;14(1):49-59.
6. Guay S-P, Voisin G, Brisson D, Munger J, Lamarche B, Gaudet D, et al. Epigenome-wide analysis in familial hypercholesterolemia identified new loci associated with high-density lipoprotein cholesterol concentration. Epigenomics. 2012;4(6):623-39.
7. Irvin MR, Zhi D, Joehanes R, Mendelson M, Aslibekyan S, Claas SA, et al. Epigenome-wide association study of fasting blood lipids in the genetics of lipid-lowering drugs and diet network study. Circulation. 2014;130(7):565-72.
8. Pfeiffer L, Wahl S, Pilling LC, Reischl E, Sandling JK, Kunze S, et al. DNA methylation of lipid-related genes affects blood lipid levels. Circ Cardiovasc Genet. 2015;8(2):334-42.
9. Drzewinska J, Walczak-Drzewiecka A, Ratajewski M. Identification and analysis of the promoter region of the human DHCR24 gene: involvement of DNA methylation and histone acetylation. Molecular biology reports. 2011;38(2):1091-101.
10. Zerenturk EJ, Sharpe LJ, Ikonen E, Brown AJ. Desmosterol and DHCR24: unexpected new directions for a terminal step in cholesterol synthesis. Progress in lipid research. 2013;52(4):666-80.
11. Waterham HR, Koster J, Romeijn GJ, Hennekam RCM, Vreken P, Andersson HC, et al. Mutations in the 3β-hydroxysterol Δ 24-reductase gene cause desmosterolosis, an autosomal recessive disorder of cholesterol biosynthesis. The American Journal of Human Genetics. 2001;69(4):685-94.
12. Daimiel LA, Fernández-Suárez ME, Rodríguez-Acebes S, Crespo L, Lasunción MA, Gómez-Coronado D, et al. Promoter analysis of the DHCR24 (3β-hydroxysterol Δ24-reductase) gene: characterization of SREBP (sterol-regulatoryelement-binding protein)-mediated activation. Bioscience reports. 2013; 33(1):e00006.
13. Dekkers KF, van Iterson M, Slieker RC, Moed MH, Bonder MJ, van Galen M, et al. Blood lipids influence DNA methylation in circulating cells. Genome Biol. 2016;17(1):138.
14. Demerath EW, Guan W, Grove ML, Aslibekyan S, Mendelson M, Zhou Y-H, et al. Epigenome-wide Association Study (EWAS) of BMI, BMI Change, and Waist Circumference in African American Adults Identifies Multiple Replicated Loci. Human molecular genetics. 2015:ddv161.
15. Petersen A-K, Zeilinger S, Kastenmüller G, Römisch-Margl W, Brugger M, Peters A, et al. Epigenetics meets metabolomics: an epigenome-wide association study with blood serum metabolic traits. Human molecular genetics. 2014;23(2):534-45.
16. Demerath EW, Guan W, Grove ML, Aslibekyan S, Mendelson M, Zhou YH, et al. Epigenome-wide association study (EWAS) of BMI, BMI change and waist circumference in African American adults identifies multiple replicated loci. Hum Mol Genet. 2015;24(15):4464-79.

17. Hidalgo B, Irvin MR, Sha J, Zhi D, Aslibekyan S, Absher D, et al. Epigenome-wide association study of fasting measures of glucose, insulin, and HOMA-IR in the Genetics of Lipid Lowering Drugs and Diet Network study. Diabetes. 2014;63(2):801-7.
18. Hopkins PN, Toth PP, Ballantyne CM, Rader DJ. Familial hypercholesterolemias: prevalence, genetics, diagnosis and screening recommendations from the National Lipid Association Expert Panel on Familial Hypercholesterolemia. Journal of clinical lipidology. 2011;5(3):S9-S17.
19. Newbury DF, Monaco AP. Genetic advances in the study of speech and language disorders. Neuron. 2010;68(2):309-20.
20. Dick KJ, Nelson CP, Tsaprouni L, Sandling JK, Aissi D, Wahl S, et al. DNA methylation and body-mass index: a genome-wide analysis. Lancet. 2014;383(9933):1990-8.
21. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. Clinical chemistry. 1972;18(6): 499-502.
22. Frazier-Wood AC, Aslibekyan S, Absher DM, Hopkins PN, Sha J, Tsai MY, et al. Methylation at CPT1A locus is associated with lipoprotein subfraction profiles. J Lipid Res. 2014;55(7):1324-30.
23. Du P, Zhang X, Huang CC, Jafari N, Kibbe WA, Hou L, et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. BMC Bioinformatics. 2010;11:587.
24. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC bioinformatics. 2012;13(1): 86.
25. Hofman A, Brusselle GGO, Murad SD, van Duijn CM, Franco OH, Goedegebure A, et al. The Rotterdam Study: 2016 objectives and design update. European journal of epidemiology. 2015;30(8):661-708.
26. Sandoval J, Heyn H, Moran S, Serra-Musach J, Pujana MA, Bibikova M, et al. Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. Epigenetics. 2011;6(6):692-702.
27. Ligthart S, Steenaard RV, Peters MJ, van Meurs JBJ, Sijbrands EJG, Uitterlinden AG, et al. Tobacco smoking is associated with DNA methylation of diabetes susceptibility genes. Diabetologia. 2016:1-9.
28. Steenaard RV, Ligthart S, Stolk L, Peters MJ, van Meurs JB, Uitterlinden AG, et al. Tobacco smoking is associated with methylation of genes related to coronary artery disease. Clinical epigenetics. 2015;7(1): 1.
29. Ligthart S, Steenaard RV, Peters MJ, van Meurs JB, Sijbrands EJ, Uitterlinden AG, et al. Tobacco smoking is associated with DNA methylation of diabetes susceptibility genes. Diabetologia. 2016;59(5): 998-1006.
30. van Iterson M, Tobi EW, Slieker RC, den Hollander W, Luijk R, Slagboom PE, et al. MethylAid: visual and interactive quality control of large Illumina 450k datasets. Bioinformatics. 2014;30(23):3435-7.
31. Westra H-J, Peters MJ, Esko T, Yaghootkar H, Schurmann C, Kettunen J, et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. Nature genetics. 2013;45(10):1238-43.

3.3

Epigenome-wide association study (EWAS) on obesity-related traits

Dhana K*, **Braun KVE***, Nano J, Voortman T, Demerath EW, Guan W, Fornage M, van Meurs JBJ, Uitterlinden AG, Hofman A, Franco OH, Dehghan A. Epigenome-wide association study (EWAS) on obesity-related traits. *Accepted for publication in American Journal of Epidemiology.*

**Denotes equal contribution*

ABSTRACT

We conducted an epigenome-wide association study (EWAS) on obesity-related traits. We used data from two prospective, population-based cohort studies: the Rotterdam Study (RS) and the Atherosclerosis Risk in Communities (ARIC) Study. We used RS (n=1,454) as the discovery panel and ARIC (n=2,097) as replication panel. Linear mixed-effect models were used to assess the cross-sectional association between genome-wide DNA methylation in leukocytes with body mass index (BMI) and waist circumference (WC) adjusting for sex, age, smoking, leukocyte proportions, array number and position on array. The two latter were modelled as random effects. Fourteen CpGs were associated with BMI and 26 CpGs with WC in RS after Bonferroni-correction ($P < 1.07 \times 10^{-7}$), of which 12 and 14 CpGs replicated in ARIC Study, respectively. The most significant novel CpGs were located at *MSI2* (cg21139312) and *LARS2* (cg18030453) and were associated both with BMI and WC. CpGs at *BRDT*, *PSMD1*, *IFI44L*, *MAP1A*, and *MAP3K5* were associated with BMI. CpGs at *LGALS3BP*, *MAP2K3*, *DHCR24*, *CPSF4L*, and *TMEM49* were associated with WC. We report novel associations of methylation at *MSI2* and *LARS2* with obesity-related traits. These results provide further insight in mechanisms underlying obesity-related traits, which can enable identification of new biomarkers in obesity-related chronic diseases.

INTRODUCTION

Obesity is an important risk factor for cardiovascular disease, diabetes, some cancers, and musculoskeletal disorders.[1-3] Evidence suggests that obesity is not only dependent on lifestyle factors, but is a result of interactions between genes and lifestyle.[4, 5] Epigenetics has been proposed as a molecular mechanism that can affect the expression of genes by environmental influences and potentially could describe further the link between obesity and its complications.[6] Nevertheless, unlike genetics, DNA methylation is dynamic overtime, therefore change in DNA methylation could also be a consequence of obesity.

Epigenetics is the study of heritable variation in gene function that is not a result of a change in DNA sequence.[7] One of the best studied epigenetic mechanisms is DNA methylation, the attachment of a methyl group to a cytosine nucleotide of CpG dinucleotides. DNA methylation has varying functions at different locations in the human genome, including regulation of gene expression.[8] To date, epigenome-wide association studies (EWAS) have identified several differentially methylated CpG regions related to body mass index (BMI) - the most widely used measure of obesity - and waist circumference (WC).[9-11] These few studies were performed in either patient populations, specific ethnic groups, or young adults. However, information among the older adults from population-based studies are scarce. In older adults and elderly, biological mechanisms involved in body weight and body composition may be different compared to younger adults.[12] Therefore, it is crucial to explore the relationship of obesity to epigenetic variation in older adults.

We performed a cross-sectional EWAS of DNA methylation in blood leukocytes for BMI and WC in subjects from the Rotterdam Study (RS), and replicated our findings in the Atherosclerosis Risk in Communities (ARIC) Study.

METHODS

Study population

The RS is a large prospective, population-based cohort study aimed at assessing the occurrence of and risk factors for chronic diseases (cardiovascular, endocrine, hepatic, neurological, ophthalmic, psychiatric, dermatological, oncological, and respiratory) in the elderly.[13] The study comprises 14,926 subjects in total, living in the well-defined Ommoord district in the city of Rotterdam in the Netherlands. In 1989, the first cohort, Rotterdam Study-I (RS-I), was established and comprised of 7,983 subjects with age 55 years or above. In 2000, the second cohort, Rotterdam Study-II (RS-II) was included with 3,011 subjects who had reached an age of 55 years since 1989. In 2006, the third cohort, Rotterdam Study-III (RS-III) was further included with 3,932 subjects with age 45 years and above. The discovery panel for the current analysis consisted of a random sample of 1,454 participants from the first and second visit of the third cohort (RS-III-1, RS-III-2) and third visit of the second cohort (RS-II-3). We sought replication of the identified CpG sites in the ARIC Study, which is described in detail elsewhere.[14] Briefly, the ARIC Study is a prospective cohort study of cardiovascular disease in adults. Between 1987 and 1989, 7,082 men and 8,710 women aged 45-64 were recruited from four US communities. Methylation data was available in a subset of 2,097 African American participants.[10] RS and ARIC study protocols were approved by Institutional Review Boards at each participating university and all participants provided written informed consent.

Anthropometric measures and covariates

Height and weight were measured with the participants standing without shoes and heavy outer garments. WC was measured at the level midway between the lower rib margin and the iliac crest with participants in standing position without heavy outer garments and with emptied pockets, breathing out gently. Hip circumference was recorded as the maximum circumference over the buttocks. BMI was calculated as weight divided by height squared (kg/m^2), and WHR was calculated as WC divided by hip circumference.[15] Information on current and past smoking behavior was acquired from questionnaires.

DNA methylation data

DNA was extracted from whole peripheral blood (stored in EDTA tubes) by standardized salting out methods. Genome-wide DNA methylation levels were measured using the Illumina Human Methylation 450K array.[16] In short, samples (500ng of DNA per sample) were first bisulfite treated using the Zymo EZ-96 DNA-methylation kit (Zymo Research, Irvine, CA, USA). Next, samples were hybridized to the arrays according to the manufacturers' protocol. The methylation percentage of a CpG site was reported as a beta-value ranging between 0 (no methylation) and 1 (full methylation). The data preprocessing was additionally performed in both datasets using an R programming pipeline which is based on the pipeline developed by Tost & Toulemat,[17] which includes additional parameters and options to preprocess and normalize methylation data directly from idat files. We excluded probes which had a detection p-value >0.01 in >95% of samples. 11,648 probes at X and Y chromosomes were excluded to avoid gender bias. The raw beta values were then background corrected and normalized using the DASEN option of the WateRmelon R-package.[18] Per individual probe, participants with methylation levels higher than three times the inter-quartiles range (IQR) were excluded.

Statistical analyses

The characteristics of the discovery and replication population are presented as mean for continuous variables and proportion for the categorical variables. In the discovery stage, we modeled cross-sectional associations between Dasen normalized beta-values of the CpG sites as outcome and BMI or WC as exposure using linear mixed effect models adjusting for age, sex, smoking, white blood cell proportions, array number (65 arrays) and position on array (12 positions; a combination of row number and column number). We performed an independent analysis in individuals from RS-III-1, from RS-III-2, and from RS-II-3. We then performed a fixed effects meta-analysis on the estimates of these three cohorts using the inverse-variance weighted method implemented in METAL combining RS-III-1 with RS-III-2 and RS-II-3.[19] Technical covariates (array number and position on array) were modeled as random effects. For the RS-III-1 we estimated leukocyte proportions (B-cells, CD4+ T-cells, CD8+ T-cells, granulocytes, monocytes and NK-cells) by a formula developed by Houseman and implemented in the *minfi* package in R.[20,21] For RS-II-3 and RS-III-2 we used white blood cell counts (WBC), i.e. lymphocytes, monocytes, and granulocytes, which were assessed with a Coulter AcT diff2 Hematology Analyzer. We corrected for multiple testing using a robust Bonferroni corrected *P*-value of 1.07×10^{-7} as the threshold for significance (0.05/463,456 probes). The probes identified in the discovery analysis were tested for replication in the independent samples from the ARIC study. A Bonferroni corrected *P*-value of 0.05 divided by the number of significant findings in the discovery study was used as a threshold of significant replica-

tion. Finally, we checked all identified CpG sites for cross-reaction or polymorphism.[22] A CpG site was considered polymorphic when a SNP with a minor allele frequency of >0.01 resided at the position of the cytosine or guanine nucleotide, or within 10 bp from the CpG site within the probe binding site.[23]

Methylation risk score

A methylation risk score was calculated based on CpG sites that were associated with the phenotypes. The effect estimates were used to build the methylation risk score using data from the discovery panel. Linear regression analyses were performed in using BMI or WC as outcome variable and the included CpG sites as exposure variables. With the use of linear regression models we calculated the lipids variance explained by the methylation risk score.

RESULTS

Table 3.3.1 summarizes the characteristics of participants in the studies. RS is entirely comprised of Europeans, whereas the ARIC study included only African Americans. Compared to RS (mean age 63.7 (8.1)), the participants in ARIC were on average younger (mean age 56.2 (5.7)) and comprised more women (63% in ARIC vs. 55% in RS)). The respective mean values of BMI in RS and ARIC study were 27.6 kg/m^2 and 30.1 kg/m^2. The mean values of WC were 93.7 cm and 101.3 cm in RS and ARIC, respectively.

Table 3.3.1. Characteristics of study populations

	RS (N=1,450)	ARIC (N=2,097)
Age, years	63.7 (8.1)	56.2 (5.7)
Gender, women	55.9	63.6
Race, %		
European	100	0
African American	0	100
BMI (kg/m2)	27.7 (4.4)	30.1 (6.1)
BMI status		
Normal weight	28.9	17.6
Overweight	46.5	37.6
Obese	24.6	43.8
WC (cm)	93.7 (12.9)	101.3 (15.1)
Smoking status		
Current smoker	18.8	24.4
Current nonsmoker	81.2	75.6
Diabetes, %	11	26

Values are mean (SD) or percentage.
ARIC, Atherosclerosis Risk in Communities; BMI, body mass index; RS, Rotterdam Study; WC, waist circumference.

Table 3.3.2 and **Table 3.3.3** present the CpG sites associated with BMI and WC in both populations. Using the Bonferroni-corrected statistical significance level of 1.07×10^{-7} we identified 14 CpG sites associated with BMI (**Supplementary Table 3.3.1**) and 26 CpG sites with WC (**Supplementary Table 3.3.2**) in RS. The identified CpG sites for BMI and WC are presented in Manhattan plots (**Figure 3.3.1** and **Figure 3.3.2**). In the ARIC Study we successfully replicated 12 out of 14 BMI related CpG sites ($P < 3.57 \times 10^{-3}$) (**Table 3.3.2**) and 13 out of 26 WC related CpG sites ($P < 1.92 \times 10^{-3}$) (**Table 3.3.3**). Among these, eight BMI-related CpG sites and 11 WC-related CpG sites were novel. The most significant novel CpG sites were located at *MSI2* (cg21139312) and *LARS2* (cg18030453) both for BMI and WC. For *MSI2* methylation, we observed an increase of 0.0009 (P 4.5×10^{-10}) and 0.0004 (P 5.9×10^{-12}) with every increase of BMI (kg/m2) and WC (cm), respectively. For *LARS2* methylation, an increase of 0.0009 (P 4.5×10^{-10}) was observed with every unit increase of BMI, and an increase of 0.0003 (P 8.8×10^{-08}) with every unit increase of WC. Additionally, for BMI other novel CpGs were located in the *BRDT* (cg03421440) and *MAP1A* (cg15159104) genes. For WC, the other top novel CpG sites were located in *TMEM49* (cg24174557) and *LGALS3BP* (cg04927537) genes. In addition to these novel findings, we confirmed previous reported CpG sites including *CPT1A*, *ABCG1* and *SREBF1* associated with BMI and WC. The scatterplots of the association between the replicated CpG sites are shown in **Supplementary Figure 3.3.1** and **Supplementary Figure 3.3.2**. **Figure 3.3.3** summarizes successfully replicated findings for BMI and WC and highlights the overlapping loci

Table 3.3.2. CpG methylation sites associated with BMI in RS at level of genome-wide significance ($P < 1.08 \times 10^{-7}$) and successfully replicated at ARIC ($P < 3.57 \times 10^{-3}$)

				Discovery Panel (RS)		Replication Panel (ARIC)		
ProbeID	Chr	Gene	Mean (SD) methylation	Beta	*P*	Beta	*P*	%Variance explained
cg00574958	11	*CPT1A*	0.19 (0.04)	-0.0011	6.2×10^{-15}	-0.0029	3.2×10^{-12}	1.9
cg00851028	1	NA	0.72 (0.04)	0.0010	5.4×10^{-08}	0.0038	9.0×10^{-04}	0.4
cg03421440	1	*BRDT*	0.71 (0.07)	-0.0015	3.2×10^{-08}	-0.0043	1.3×10^{-03}	0.3
cg06096336	2	*PSMD1*	0.64 (0.05)	0.0016	4.3×10^{-08}	0.0058	5.5×10^{-04}	1.0
cg06500161	21	*ABCG1*	0.71 (0.03)	0.0011	1.7×10^{-09}	0.0081	1.5×10^{-13}	0.3
cg06872964	1	*IFI44L*	0.62 (0.06)	0.0015	4.8×10^{-08}	0.0100	4.3×10^{-07}	0.4
cg11024682	17	*SREBF1*	0.55 (0.04)	0.0013	6.6×10^{-15}	0.0068	9.6×10^{-09}	0.1
cg15159104	15	*MAP1A*	0.48 (0.05)	0.0010	3.2×10^{-08}	0.0048	5.1×10^{-06}	0.0
cg15903032	10	NA	0.57 (0.04)	0.0010	7.6×10^{-08}	0.0037	2.8×10^{-03}	0.2
cg18030453	3	*LARS2*	0.72 (0.04)	0.0009	4.5×10^{-09}	0.0028	1.7×10^{-03}	0.1
cg21139312	17	*MSI2*	0.89 (0.03)	0.0009	4.5×10^{-10}	0.0028	1.2×10^{-06}	2.0
cg21506299	6	*MAP3K5*	0.23 (0.06)	-0.0010	3.5×10^{-08}	-0.0019	2.8×10^{-03}	1.1

Betas shows the regression coefficients based on linear mixed models and reflect differences in methylation beta values per increase in BMI unit.

Models are adjusted for age, gender, current smoking, leukocyte proportions, array number, and position on array.

ARIC, Atherosclerosis Risk in Communities; BMI, Body Mass Index; NA, Not Annotated; RS, Rotterdam Study.

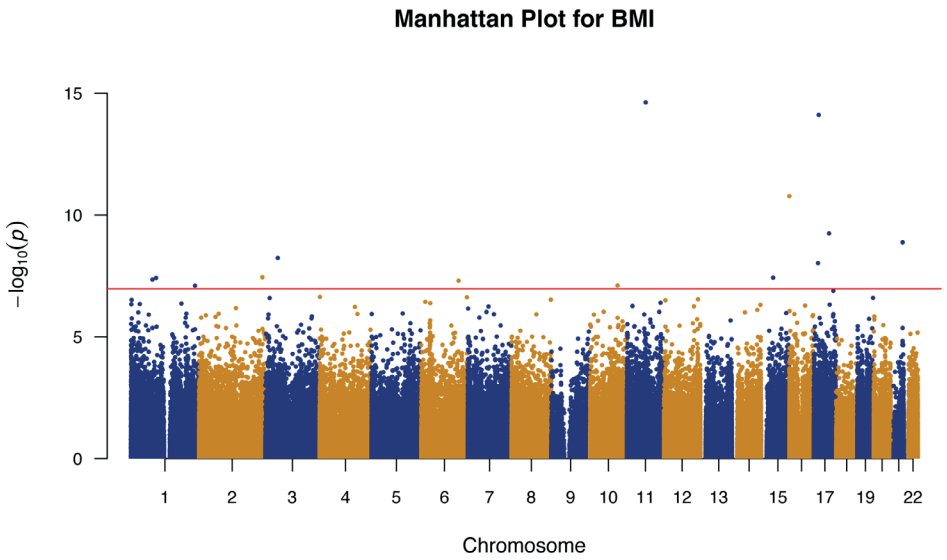


Figure 3.3.1. Manhattan plot epigenome-wide associations between genome wide DNA methylation and BMI.

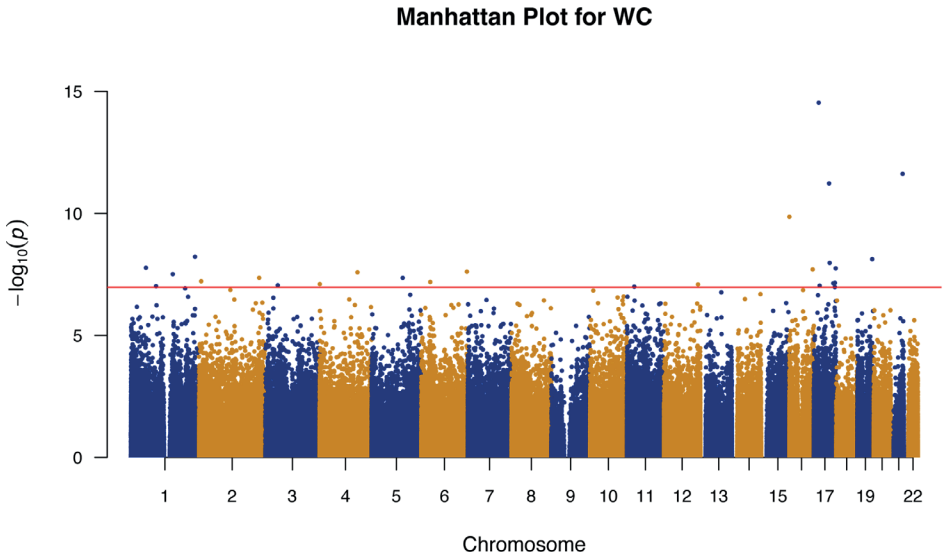


Figure 3.3.2. Manhattan plot epigenome-wide associations between genome wide DNA methylation and WC.

including *ABCG1*, *MSI2*, *LARS2*, *SREBF1*, and *CPT1A*. To test for genomic inflation we calculated the lambda for the EWAS on BMI and WC, and created QQ-plots. The lambda was 1.487 and 1.556 for the EWAS on BMI and WC, respectively. The QQ-plot are shown in **Supplementary Figure 3.3.3** and **Supplementary Figure 3.3.4**.

We calculated a methylation risk score based on the 12 CpG sites for BMI and 14 CpG sites for WC that were identified and replicated in the current study. For BMI, 2.0% of the variance was explained by the methylation risk score, whereas for WC the variance explained was 6.4%.

Table 3.3.3. CpG methylation sites associated with WC in RS at level of genome-wide significance ($P < 1.08 \times 10^{-7}$) and successfully replicated at ARIC ($P < 1.92 \times 10^{-3}$)

				Discovery Panel (RS)		Replication Panel (ARIC)		
ProbeID	Chr	Gene	Mean (SD) methylation	Beta	*P*	Beta	*P*	%Variance explained
cg00574958	11	*CPT1A*	0.19 (0.04)	-0.0005	1.2×10^{-17}	-0.0034	5.8×10^{-17}	3.3
cg00851028	1	*NA*	0.72 (0.04)	0.0004	6.0×10^{-09}	0.0043	1.2×10^{-04}	0
cg04927537	17	*LGALS3BP*	0.57 (0.05)	0.0006	7.0×10^{-08}	0.0093	7.0×10^{-08}	1.6
cg05899984	12	*NA*	0.84 (0.03)	0.0003	8.1×10^{-08}	0.0038	5.7×10^{-06}	2.9
cg06500161	21	*ABCG1*	0.71 (0.03)	0.0005	2.4×10^{-12}	0.0096	4.4×10^{-19}	0.8
cg11024682	17	*SREBF1*	0.55 (0.04)	0.0005	2.9×10^{-15}	0.0080	3.5×10^{-12}	1.2
cg13139542	2	*NA*	0.89 (0.02)	0.0002	6.0×10^{-08}	0.0029	4.7×10^{-06}	0
cg15416179	17	*MAP2K3*	0.14 (0.03)	-0.0002	9.1×10^{-08}	-0.0019	2.6×10^{-04}	3.6
cg17901584	1	*DHCR24*	0.68 (0.07)	-0.0005	1.7×10^{-08}	-0.0080	8.3×10^{-08}	2.0
cg18030453	3	*LARS2*	0.72 (0.04)	0.0003	8.8×10^{-08}	0.0029	8.6×10^{-04}	0
cg18772573	17	*CPSF4L*	0.85 (0.03)	0.0003	7.3×10^{-08}	0.0039	2.8×10^{-05}	0
cg21139312	17	*MSI2*	0.89 (0.03)	0.0004	5.9×10^{-12}	0.0028	6.1×10^{-07}	8.2
cg24174557	17	*TMEM49*	0.38 (0.07)	-0.0005	1.1×10^{-08}	-0.0059	5.3×10^{-05}	0

Betas shows the regression coefficients based on linear mixed models and reflect differences in methylation beta values per increase in WC unit.
Models are adjusted for age, gender, current smoking, leukocyte proportions, array number, and position on array.
ARIC, Atherosclerosis Risk in Communities; NA, Not Annotated; RS, Rotterdam Study; WC, waist circumference

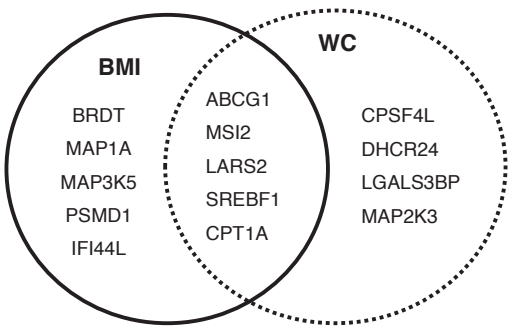


Figure 3.3.3. Successfully replicated CpGs for BMI and WC and their overlap

DISCUSSION

This study used an EWAS approach to identify novel differentially methylated genes for obesity-related traits in older adults. The EWAS analysis in RS provided numerous novel loci associated with BMI and WC, of which many findings successfully replicated in ARIC. Our most significant CpG sites associated with both BMI and WC were located at the *MSI2* and *LARS2* genes. Additionally, CpG sites at *BRDT* and *MAP1A* were associated with BMI, and CpG sites at *TMEM49* and *LGALS3BP* were associated with WC. Moreover, we confirmed previous findings that methylation at *CPT1A*, *ABCG1*, and *SREBF1* are associated with BMI and WC.

Previous EWAS on obesity traits were conducted in population-based studies including ARIC and Genetics of Lipid Lowering Drugs and Diet Network Study (GOLDN),[9, 10] and in individuals with history of myocardial infarction or healthy blood donors from the Cardiogenics Consortium.[11] Similar to our findings, ARIC and GOLDN reported an inverse association between CpG site at *CPT1A* and BMI,[9, 10] and positive associations of CpGs at *ABCG1* and *SREBF1* with BMI and WC.[10] The Cardiogenics Consortium, however, reported only a positive association between three CpG sites at *HIF3A* with BMI in both blood and adipose tissue DNA in European adults.[11] These CpGs sites at *HIF3A* did not achieve the threshold for statistical significance in our study. However, CpGs sites at *HIF3A* were replicated by ARIC in DNA blood.[10] This discrepancy may be due to difference in the prevalence of obesity and comorbidities in our study (25% obese, 11% diabetes, 7 % CHD) compared with ARIC (44% obese, 26% diabetes) and Cardiogenics Consortium (4% diabetes, 52% MI).

The known loci, *CPT1A*, *ABCG1*, and *SREBF1* are involved in regulation of lipids, lipoprotein metabolism and insulin sensitivity.[24-26] Specifically, the *CPT1A* gene encodes for carnitine palmitoy-transferase-1, which is a mitochondrial protein involved in fatty acid metabolism and lipoprotein subfraction.[25, 27] The *ABCG1* gene encodes for ATP-binding cassette sub-family G member 1 protein and is involved in the transport of cholesterol and phospholipids in macrophages.[28] Finally, the *SREBF1* gene encodes for sterol regulatory element-binding transcription factor 1, which is known to promote adipocyte differentiation and signaling of insulin action.[26] Although it has been shown previously that these loci are associated with obesity-related traits, it is still important to replicate these findings across different study population. Since the EWAS approach is hypothesis-free, findings are prone to be false-positive. By replicating previously reported results, we can say with more certainty that these CpG sites are true-positive findings.

In addition to confirming these previously identified loci, we have identified and replicated novel CpG sites located in the gene body of the *MSI2* (cg21139312) and *LARS2* (cg18030453) gene, which were associated with both BMI and WC. The CpG site at *MSI2* gene explained 2.2% of variation in BMI and 8.2% of variation in WC. *MSI2* encodes RNA-binding proteins and plays a central role in posttranscriptional gene regulation.[29] A genome-wide association study in pigs suggested that *MSI2* is associated with eating behaviors, including number of visits to feeder per day.[30] Moreover, another study performed in mice reported that *MSI2* is linked with the proliferation and maintenance of stem cells in the central nervous system.[31] This study suggested that during neurogenesis *MSI2* expression persisted in a subset of neuronal lineage cells, such as parvalbumin-containing GABA neurons in the neocortex.[29, 31] GABA receptors are involved in controlling feeding behavior, reinforcing the role of *MSI2* in obesity. The other novel locus associated with both BMI and WC, *LARS2*, encodes an

enzyme that catalyzes aminoacylation of mitochondrial tRNALeu.[32] A previous post-mortem study showed that *LARS2* expression (human leucyl-tRNA synthetase 2, mitochondrial NM015340) was increased in brain tissue of patients with bipolar disorder compared to controls.[33] Considering that bipolar disorder is associated with obesity, overweight, and abdominal obesity,[34] methylation of *MSI2* and *LARS2* could play a role in disturbances in eating behaviors, and consequently BMI and WC. However, further studies are warranted to establish the temporality and the pathway of the associations. Even though previous studies have investigated the association between DNA methylation and anthropometrics, this study is the first report an association between DNA methylation of several CpG sites, including *MSI2* and *LARS2*, with BMI and WC. One possible explanation of discrepancies between findings of our current study compared to previous similar studies, is the difference in population characteristics. Study populations from previous studies consisted of mixed ethnic groups, participants of younger age, or at high disease risk.[11, 35-37] Considering that our discovery cohort consisted of an ethnic homogenous group of older adults from the general population of Rotterdam, underlying mechanisms may differ from other population groups.

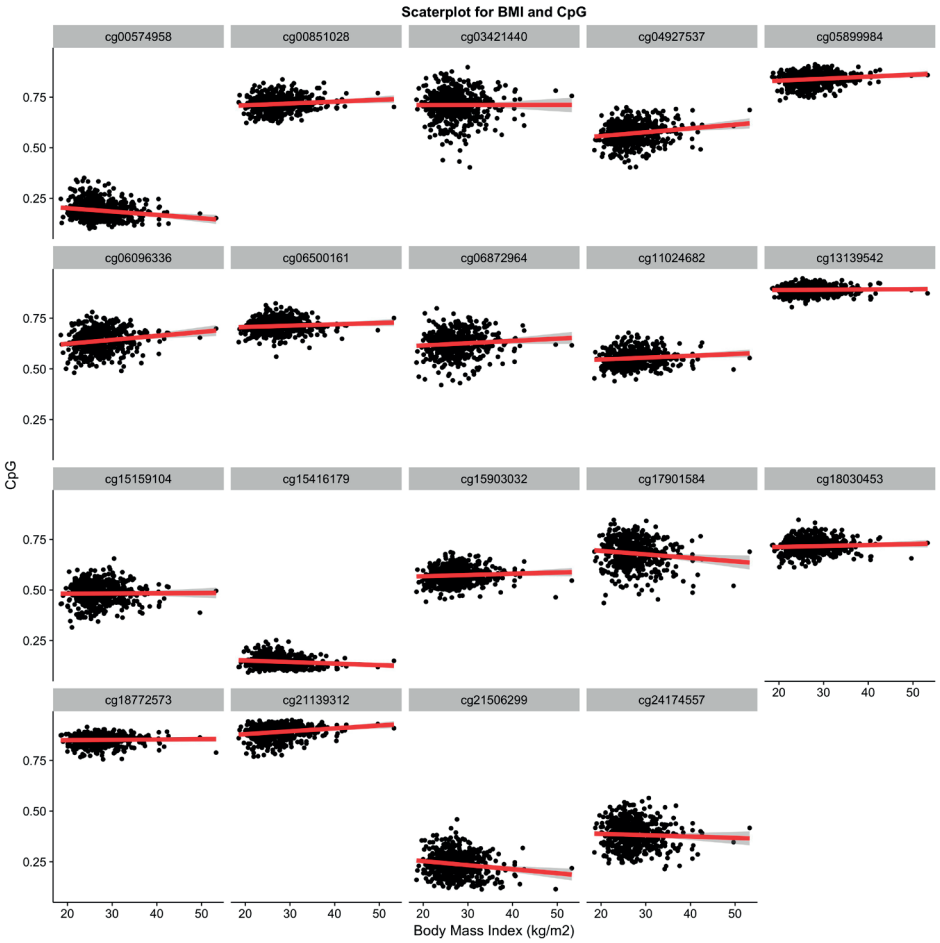
In this study we conducted an EWAS in a European population and replicated the findings in African Americans. Epidemiologic studies have reported large disparities across racial/ethnic groups in the development of obesity.[38] For example, in the current study the rates of obesity were significantly lower in Europeans (24.6%) than in African Americans (43.8%). However, despite the differences in ethnicity and prevalence of obesity between our studies, most of our CpGs sites (86.7%) successfully replicated in ARIC. This may indicate that, in contrast to genetic studies where replication across ethnic groups is challenging due to differences in LD pattern, epigenetic findings could more easily be translated across ethnic groups.

The strength of the current study includes the large sample size with available data on DNA methylation and the ability to replicate our findings in different ethnic population. However, the results of this study must be interpreted in light of several limitations. We used whole blood samples for the quantification of DNA methylation, whereas adipose tissue may be a more relevant tissue in examining obesity. In this case, important CpG sites may not have been identified in our study. Unlike in genetic studies, unraveling the direction of the association between DNA methylation and phenotypes in epigenetic epidemiology remains challenging. Due to the cross-sectional design and the nature of our variables, which are responsive to the environment and dynamic over time, a temporal direction of the association between DNA methylation and anthropometric measures cannot be determined. As previous studies have shown that change in DNA methylation is a consequence of BMI for the majority of CpG sites, this may be the most likely direction for the associations observed in the current study as well.[37] However, longitudinal studies are required to confirm the direction of the associations between DNA methylation and anthropometrics. Another possibility is that our findings could be explained by third common factors. For instance, associations may be confounded by differences in cell type proportion. In order to avoid this source of confounding, all analyses were adjusted for cell type proportions. However, as in any observational study residual confounding, due to various lifestyle factors, still remains an issue. Another possibility is that our findings could be explained by third common factors. For instance, associations may be confounded by differences in cell type proportion. In order to avoid this source of confounding, all analyses were adjusted for cell type proportions. However, as in any observational study residual confounding, due to various

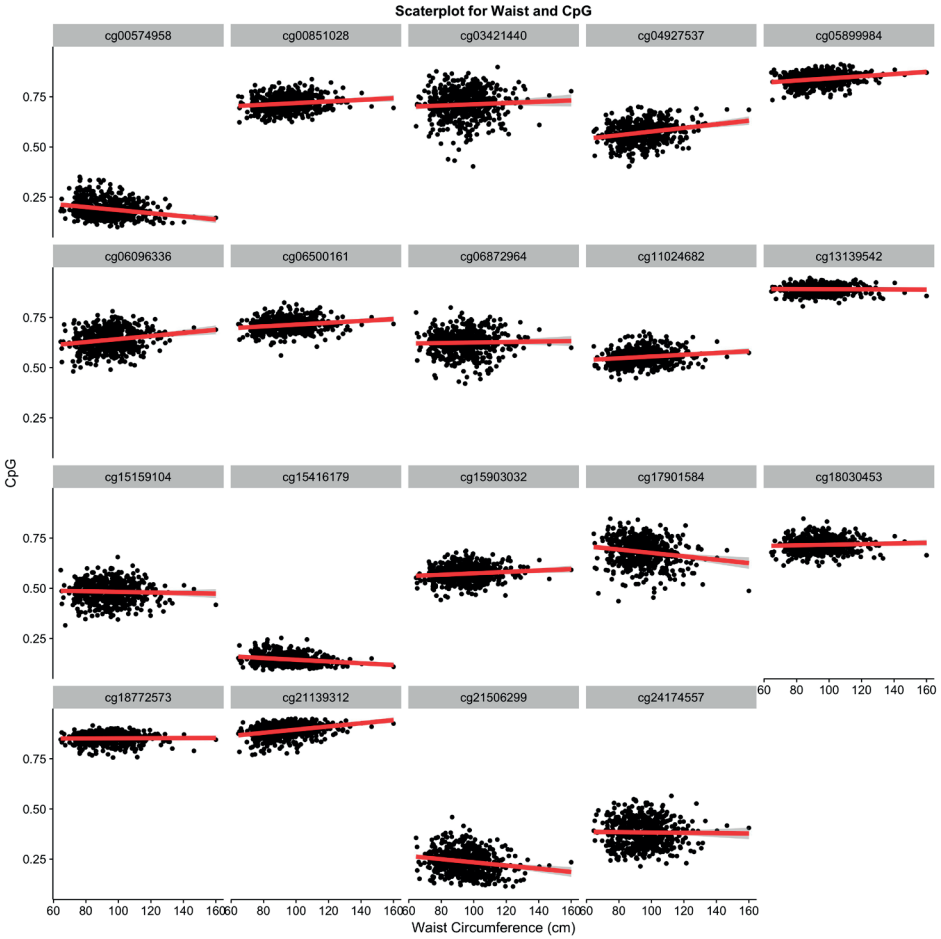
lifestyle factors, still remains an issue. Furthermore, the QQ plots showed a high genomic inflation. Many EWAS studies have reported high genomic inflation.[39] Adjustment for potential confounders such as technical covariates could decrease the inflation. The correlation between CpGs and the large number of findings in EWAS studies are suggested to explain the residual inflation.[40] In this study we did adequate adjustment for technical covariates. Moreover, the replication of our results in an independent population provides further evidence for the robustness of our findings.

In conclusion, we have reported a novel association of increased methylation at the *MSI2* and *LARS2* genes with increased BMI and WC in older adults. Moreover, we confirmed three previously identified methylation loci (*CPT1A*, *ABCG1* and *SREBF1*) suggested to be associated with obesity. Further investigations using repeatedly measured genome-wide DNA methylation and obesity-related traits are needed to assess causality and to further evolve the growing field of epigenetic epidemiology toward novel therapeutic and preventative approaches of obesity and non-communicable related disorders.

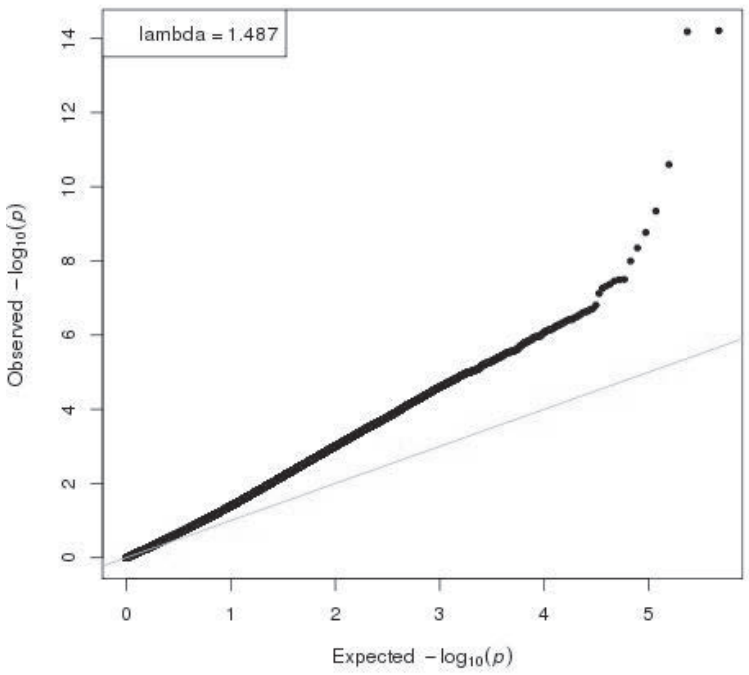
SUPPLEMENTARY MATERIAL



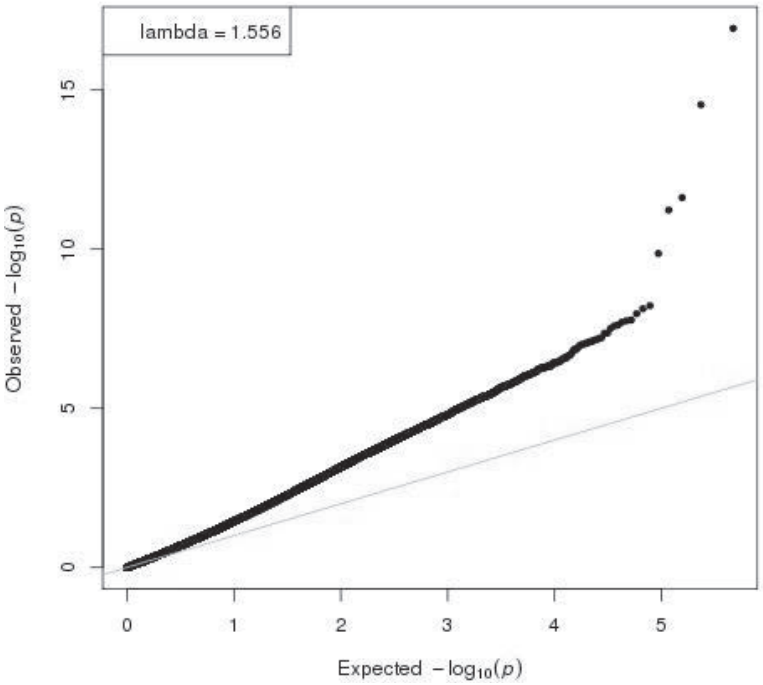
Supplementary Figure 3.3.1. Scatter plots and regression lines for significantly replicated CpG sites and BMI.



Supplementary Figure 3.3.2. Scatter plots and regression lines for significantly replicated CpG sites and WC.



Supplementary Figure 3.3.3. QQ plot for EWAS on BMI.



Supplementary Figure 3.3.4. QQ plot for EWAS on WC.

Supplementary Table 3.3.1. CpG methylation sites associated with BMI in RS (n=1,454).

ProbeID	chr	Gene	Effect	P^1
cg00574958	11	*CPT1A*	-0.0011	6.22×10^{-15}
cg11024682	17	*SREBF1*	0.0013	6.61×10^{-15}
cg16778405	16	*RAB40C*	0.0006	2.52×10^{-11}
cg21139312	17	*MSI2*	0.0009	4.54×10^{-10}
cg06500161	21	*ABCG1*	0.0011	1.71×10^{-09}
cg18030453	3	*LARS2*	0.0009	4.49×10^{-09}
cg08794157	17	*PMP22*	0.0010	1.01×10^{-08}
cg15159104	15	*MAP1A*	0.0010	3.15×10^{-08}
cg03421440	1	*BRDT*	-0.0015	3.23×10^{-08}
cg21506299	6	*MAP3K5*	-0.0010	3.49×10^{-08}
cg06096336	2	*PSMD1*	0.0016	4.26×10^{-08}
cg06872964	1	*IFI44L*	0.0015	4.81×10^{-08}
cg00851028	1	NA[2]	0.0010	5.50×10^{-08}
cg15903032	10	NA[2]	0.0010	7.55×10^{-08}

Betas shows the regression coefficients based on linear mixed models and reflect differences in methylation beta values per increase in BMI unit.

Models are adjusted for age, gender, current smoking, leukocyte proportions, array number, and position on array.

[1]*Level of significance:* $P < 1.07 \times 10^{-7}$

[2]*Not annotated.*

Supplementary Table 3.3.2. CpG methylation sites associated with waist circumference in RS (n=1,454).

ProbeID	chr	Gene	Effect	P^1
cg00574958	11	*CPT1A*	-0.0005	1.16×10^{-17}
cg11024682	17	*SREBF1*	0.0005	2.90×10^{-15}
cg06500161	21	*ABCG1*	0.0005	2.39×10^{-12}
cg21139312	17	*MSI2*	0.0004	5.92×10^{-12}
cg16778405	16	*RAB40C*	0.0002	1.37×10^{-10}
cg00851028	1	NA[2]	0.0004	5.97×10^{-09}
cg12970155	19	*MYADM*	0.0008	7.48×10^{-09}
cg24174557	17	*TMEM49*	-0.0005	1.07×10^{-08}
cg17901584	1	*DHCR24*	-0.0005	1.69×10^{-08}
cg11969813	17	*P4HB*	0.0005	1.78×10^{-08}
cg09780767	16	NA[2]	-0.0003	1.98×10^{-08}
cg03564497	6	*TCP10L2*	-0.0007	2.44×10^{-08}
cg06690548	4	*SLC7A11*	-0.0003	2.60×10^{-08}
cg01411912	1	*S100A4*	-0.0004	3.09×10^{-08}
cg20456243	2	*SPEG*	-0.0004	4.37×10^{-08}
cg05635169	5	NA[2]	-0.0005	4.39×10^{-08}
cg13139542	2	NA[2]	0.0002	6.05×10^{-08}
cg04927537	17	*LGALS3BP*	0.0006	6.99×10^{-08}
cg18772573	17	*CPSF4L*	0.0003	7.30×10^{-08}
cg03096285	4	*MAEA*	0.0003	7.90×10^{-08}
cg05899984	12	NA[2]	0.0003	8.11×10^{-08}
cg18030453	3	*LARS2*	0.0003	8.84×10^{-08}
cg15416179	17	*MAP2K3*	-0.0002	9.11×10^{-08}
cg03421440	1	*BRDT*	-0.0006	9.58×10^{-08}
cg06734985	11	NA[2]	-0.0005	1.00×10^{-07}
cg00945209	17	*USP36*	-0.0004	1.04×10^{-07}

Betas shows the regression coefficients based on linear mixed models and reflect differences in methylation beta values per increase in WC unit.
Models are adjusted for age, gender, current smoking, leukocyte proportions, array number, and position on array.
[1]*Level of significance:* $P < 1.07 \times 10^{-7}$
[2]*Not annotated.*

REFERENCES

1. Obesity: preventing and managing the global epidemic. Report of a WHO consultation. World Health Organ Tech Rep Ser. 2000;894:i-xii, 1-253.
2. Berrington de Gonzalez A, Hartge P, Cerhan JR, Flint AJ, Hannan L, MacInnis RJ, et al. Body-mass index and mortality among 1.46 million white adults. The New England journal of medicine. 2010; 363(23):2211-9.
3. Collaboration NCDRF. Trends in adult body-mass index in 200 countries from 1975 to 2014: a pooled analysis of 1698 population-based measurement studies with 19.2 million participants. Lancet (London, England). 2016;387(10026):1377-96.
4. Marti A, Martinez-Gonzalez MA, Martinez JA. Interaction between genes and lifestyle factors on obesity. Proc Nutr Soc. 2008;67(1):1-8.
5. Vattikuti S, Guo J, Chow CC. Heritability and genetic correlations explained by common SNPs for metabolic syndrome traits. PLoS Genet. 2012;8(3):e1002637.
6. Bollati V, Baccarelli A. Environmental epigenetics. Heredity. 2010;105(1):105-12.
7. Mill J, Dempster E, Caspi A, Williams B, Moffitt T, Craig I. Evidence for monozygotic twin (MZ) discordance in methylation level at two CpG sites in the promoter region of the catechol-O-methyltransferase (COMT) gene. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics. 2006;141(4): 421-5.
8. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet. 2012;13(7):484-92.
9. Aslibekyan S, Demerath EW, Mendelson M, Zhi D, Guan W, Liang L, et al. Epigenome-wide study identifies novel methylation loci associated with body mass index and waist circumference. Obesity (Silver Spring). 2015;23(7):1493-501.
10. Demerath EW, Guan W, Grove ML, Aslibekyan S, Mendelson M, Zhou YH, et al. Epigenome-wide association study (EWAS) of BMI, BMI change and waist circumference in African American adults identifies multiple replicated loci. Hum Mol Genet. 2015;24(15):4464-79.
11. Dick KJ, Nelson CP, Tsaprouni L, Sandling JK, Aissi D, Wahl S, et al. DNA methylation and body-mass index: a genome-wide analysis. Lancet. 2014;383(9933):1990-8.
12. Mendelson MM, Marioni RE, Joehanes R, Liu C, Hedman AK, Aslibekyan S, et al. Association of Body Mass Index with DNA Methylation and Gene Expression in Blood Cells and Relations to Cardiometabolic Disease: A Mendelian Randomization Approach. PLoS Med. 2017;14(1):e1002215.
13. Hofman A, Murad SD, van Duijn CM, Franco OH, Goedegebure A, Ikram MA, et al. The Rotterdam Study: 2014 objectives and design update. Eur J Epidemiol. 2013;28(11):889-926.
14. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. Am J Epidemiol. 1989;129(4):687-702.
15. Eveleth PB. Physical status: The use and interpretation of anthropometry. Report of a WHO Expert Committee - WHO. Am J Hum Biol. 1996;8(6):786-7.
16. Sandoval J, Heyn H, Moran S, Serra-Musach J, Pujana MA, Bibikova M, et al. Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. Epigenetics. 2011;6(6):692-702.
17. Touleimat N, Tost J. Complete pipeline for Infinium((R)) Human Methylation 450K BeadChip data processing using subset quantile normalization for accurate DNA methylation estimation. Epigenomics. 2012;4(3):325-41.
18. Pidsley R, CC YW, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to preprocessing Illumina 450K methylation array data. BMC Genomics. 2013;14:293.

19. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics. 2010;26(17):2190-1.
20. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics. 2014;30(10):1363-9.
21. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics. 2012;13:86.
22. Chen YA, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. Epigenetics. 2013;8(2):203-9.
23. Barfield RT, Almli LM, Kilaru V, Smith AK, Mercer KB, Duncan R, et al. Accounting for population stratification in DNA methylation studies. Genet Epidemiol. 2014;38(3):231-41.
24. Braun KV, Voortman T, Dhana K, Troup J, Bramer WM, Troup J, et al. The role of DNA methylation in dyslipidaemia: A systematic review. Prog Lipid Res. 2016;64:178-91.
25. Frazier-Wood AC, Aslibekyan S, Absher DM, Hopkins PN, Sha J, Tsai MY, et al. Methylation at CPT1A locus is associated with lipoprotein subfraction profiles. J Lipid Res. 2014;55(7):1324-30.
26. Kim JB, Spiegelman BM. ADD1/SREBP1 promotes adipocyte differentiation and gene expression linked to fatty acid metabolism. Genes Dev. 1996;10(9):1096-107.
27. Gobin S, Thuillier L, Jogl G, Faye A, Tong L, Chi M, et al. Functional and structural basis of carnitine palmitoyltransferase 1A deficiency. J Biol Chem. 2003;278(50):50428-34.
28. Klucken J, Buchler C, Orso E, Kaminski WE, Porsch-Ozcurumez M, Liebisch G, et al. ABCG1 (ABC8), the human homolog of the Drosophila white gene, is a regulator of macrophage cholesterol and phospholipid transport. Proc Natl Acad Sci U S A. 2000;97(2):817-22.
29. Sakakibara S, Nakamura Y, Satoh H, Okano H. Rna-binding protein Musashi2: developmentally regulated expression in neural precursor cells and subpopulations of neurons in mammalian CNS. J Neurosci. 2001;21(20):8091-107.
30. Do DN, Strathe AB, Ostersen T, Jensen J, Mark T, Kadarmideen HN. Genome-wide association study reveals genetic architecture of eating behavior in pigs and its implications for humans obesity by comparative mapping. PLoS One. 2013;8(8):e71509.
31. Sakakibara S, Nakamura Y, Yoshida T, Shibata S, Koike M, Takano H, et al. RNA-binding protein Musashi family: roles for CNS stem cells and a subpopulation of ependymal cells revealed by targeted disruption and antisense ablation. Proc Natl Acad Sci U S A. 2002;99(23):15194-9.
32. Sohm B, Sissler M, Park H, King MP, Florentz C. Recognition of human mitochondrial tRNALeu(UUR) by its cognate leucyl-tRNA synthetase. Journal of Molecular Biology. 2004;339(1):17-29.
33. Munakata K, Iwamoto K, Bundo M, Kato T. Mitochondrial DNA 3243A>G mutation and increased expression of LARS2 gene in the brains of patients with bipolar disorder and schizophrenia. Biol Psychiatry. 2005;57(5):525-32.
34. McElroy SL, Kotwal R, Malhotra S, Nelson EB, Keck PE, Nemeroff CB. Are mood disorders and obesity related? A review for the mental health professional. J Clin Psychiatry. 2004;65(5):634-51, quiz 730.
35. Aslibekyan S, Demerath EW, Mendelson M, Zhi D, Guan W, Liang L, et al. Epigenome-wide study identifies novel methylation loci associated with body mass index and waist circumference. Obesity. 2015;23(7):1493-501.
36. Demerath EW, Guan W, Grove ML, Aslibekyan S, Mendelson M, Zhou Y-H, et al. Epigenome-wide Association Study (EWAS) of BMI, BMI Change, and Waist Circumference in African American Adults Identifies Multiple Replicated Loci. Human molecular genetics. 2015:ddv161.

37. Wahl S, Drong A, Lehne B, Loh M, Scott WR, Kunze S, et al. Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. Nature. 2017;541(7635):81-6.
38. Zhang Q, Wang Y, Huang ES. Changes in racial/ethnic disparities in the prevalence of Type 2 diabetes by obesity level among US adults. Ethnicity & Health. 2009;14(5):439-57.
39. Joubert BR, Herman T, Felix JF, Bohlin J, Ligthart S, Beckett E, et al. Maternal plasma folate impacts differential DNA methylation in an epigenome-wide meta-analysis of newborns. Nature communications. 2016;7.
40. Lehne B, Drong AW, Loh M, Zhang W, Scott WR, Tan ST, et al. A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. Genome Biol. 2015;16:37.

Chapter 4

Nutrition & cardiometabolic health in early life

4.1

Methyl donor nutrient intake in early childhood and body composition at the age of 6 years: The Generation R Study

Based on:

Braun KVE*, Voortman T*, Kiefte-de Jong JC, Jaddoe VWV, Hofman A, Franco OH, van den Hooven EH. Dietary Intakes of Folic Acid and Methionine in Early Childhood Are Associated with Body Composition at School Age. *Journal of Nutrition. 2015;145(9):2123-9.*

**Denotes equal contribution*

ABSTRACT

Background: Deficiency of vitamin B-6, vitamin B-12, folate, folic acid and methionine may lead to dysregulation of DNA methylation, which might lead to a disturbed energy and lipid metabolism.
Objective: We aimed to explore whether intake of vitamin B-6, vitamin B-12, folate, folic acid and methionine at 1 year is associated with measures of growth and body composition at the age of 6 years.
Methods: This study was performed in 2,922 children participating in The Generation R Study, a population-based prospective cohort study. Dietary intakes of vitamin B-6, B-12, folate, folic acid, and methionine were assessed at a median age of 12.9 months using a validated food frequency questionnaire. At the age of 6 years, height and weight were measured and BMI was calculated. Body fat was measured using DXA and body fat percentage and the android:gynoid ratio were calculated. Multivariable linear regression analyses were used to study associations between intakes of vitamin B-6, vitamin B-12, folate, folic acid and methionine with body composition.
Results: In models adjusted for maternal and child characteristics, children with folic acid intakes in the highest tertile had a 0.16 SD lower weight (95%CI -0.31, -0.02) and a 0.14 SD lower BMI (95%CI -0.26, -0.01) compared to those in the lowest tertile. Children with vitamin B-12 intakes in the highest tertile had a 0.13 SD higher android:gynoid ratio (95%CI 0.00, 0.25) compared to those in the lowest tertile. In addition, children with intakes in the highest tertile of methionine had a 0.09 SD higher BMI (95%CI 0.01, 0.17), and a 0.12 SD android:gynoid ratio (95%CI 0.02, 0.22). Vitamin B-6 and folate intakes were not associated with any of the body composition outcomes measured.
Conclusions: In this population of children, high folic acid intakes were associated with a lower body weight and BMI. In contrast, higher methionine intakes were associated with unfavorable body composition at the age of 6 years. Future studies should investigate long term consequences of these outcomes on health.

INTRODUCTION

The prevalence of overweight is rising among children, increasing the risk of metabolic syndrome, type 2 diabetes, and obesity in adulthood.[1, 2] For adequate prevention, it is important to identify factors that influence the development of overweight in early life. One of the mechanisms that might be involved in the early development of overweight is DNA methylation. DNA methylation is a mechanism in which gene expression is altered in response to environmental influences, such as nutrition and physical activity.[3] For example, there are several nutrients, including vitamin B-6, vitamin B-12 and folate, which can directly or indirectly act as methyl donors in DNA methylation.[4-6] Deficiency of these methyl donor nutrients could lead to dysregulation of DNA methylation and might generate metabolic disturbances, including disturbed energy and lipid metabolism, and a higher future risk of cardiovascular disease.[7-9]

Considering that DNA methylation might influence the development of overweight, this might be particularly important during early life, when epigenetic changes are actively occurring.[10-12] Therefore, it is possible that methyl donor nutrient intake during critical periods of growth, for instance during early childhood, could have an influence on the development of overweight in children. For instance, Perng et al. observed that a lower DNA methylation was associated with the development of adiposity in 5-12 year old boys, but not in girls.[13]

Several animal studies have examined the effects of methyl donor nutrient supplementation on body composition. Results of these studies suggest that high intakes of methyl donor nutrients could have a protective effect on liver fat accumulation,[14] and that methyl donor nutrient supplementation during pregnancy has a protective effect on the development of obesity in the offspring.[15, 16] A recent study by Gunanti et al. reported that low serum levels of folate and vitamin B-12 were associated with higher BMI and fat mass in Mexican American children aged 8-15 years.[17] Furthermore, a few other cross-sectional studies investigated the associations between folate and vitamin B-12 and body composition in children and adolescents. Results of these studies show that lower levels or intakes of vitamin B-12 or folate were associated with unfavorable body composition outcomes, such as a higher BMI and waist circumference.[18-21] In addition, a few studies suggest that a higher homocysteine level, which could be the result of a vitamin B-12 and/or folate deficiency, is associated with unfavourable body composition in children.[22-24] The aim of this study was to investigate the associations between dietary intakes of vitamin B-6, vitamin B-12, folate, folic acid, and methionine in early childhood and growth and body composition outcomes at school-age in 2,922 children participating in a prospective cohort study.

METHODS

Design and subjects

This study was embedded in the Generation R Study, a population-based prospective cohort study from early fetal life onward in Rotterdam, the Netherlands.[25] The study was approved by the Medical Ethical Committee of Erasmus Medical Center, Rotterdam. Written informed consent was obtained from all participating mothers. All children were born between 2002 and 2006 and data on follow-up

in early childhood was available for 7,893 children (Figure 4.1.1). A Food Frequency Questionnaire (FFQ) to assess infant diet was sent to 5,088 mothers who provided informed consent for postnatal follow-up and mastered the Dutch language sufficiently.[26] This FFQ was returned by 3,650 mothers (72%). After exclusion of subjects with invalid dietary data and withdrawn consent, information on nutrition was available for 3,629 children.[26] Of these children, growth and body composition data at the age of 6 years were available for 2,922 singleton born children (**Figure 4.1.1**).

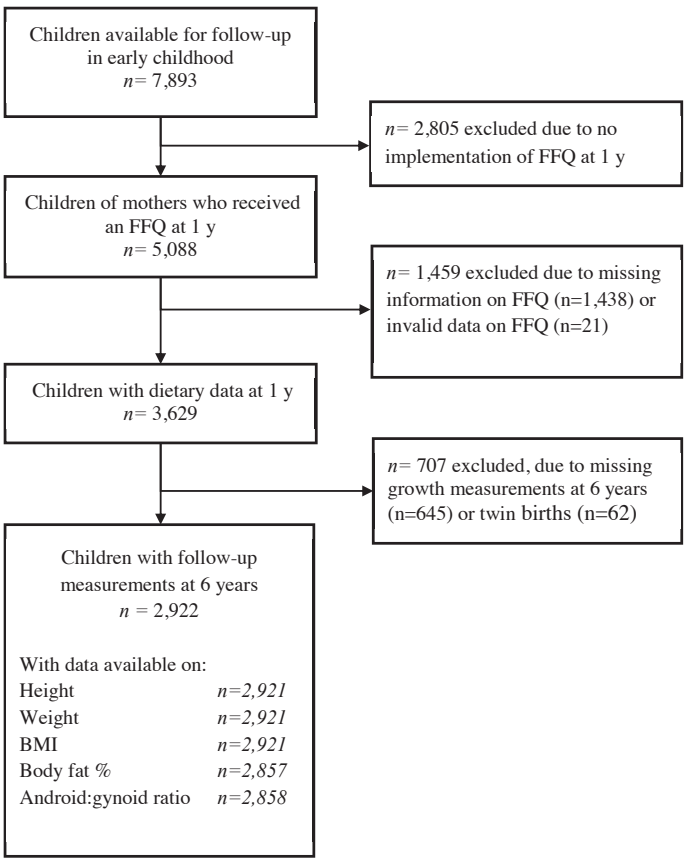


Figure 4.1.1. Flowchart: Population for analysis

Dietary assessment

Dietary intake of the children was assessed at a median age of 12.9 months (95% range 12.2-18.9) using a semi-quantitative FFQ filled in by one of the parents, covering the previous month.[26] This FFQ developed for our study population and was modified to include only foods that are frequently consumed by children aged 9 to 18 months according to a Dutch national food consumption survey in 2002.[27] The final FFQ consisted of 211 food items and included questions on frequency, quantity, type and preparation methods. This FFQ was validated against three 24h-recalls, obtained by trained

nutritionists, in a representative sample of 32 Dutch children aged 14 months.[26] Intraclass correlation coefficients for nutrient intakes ranged from 0.36 to 0.74.[28] Food frequencies and quantities were converted into grams per day using standardized portion sizes. Vitamin B-6, vitamin B-12, folate and methionine intakes were calculated using the Dutch Food Composition Table 2006.[29] For folate intakes we calculated both naturally occurring dietary folate and synthetically produced folic acid, which is added to fortified foods. Total dietary folate equivalents (DFE) were calculated taking into account the differences in bioavailability of dietary folate and synthetic folic acid (1 DFE = 1 μg dietary folate or 0.6 μg folic acid added to food), according to a commonly used formula for DFE.[30]

Growth and body composition outcomes

At a median age of 5.9 years (95% range 5.7-6.6) children visited our research center in the Erasmus Medical Center for a detailed physical examination. Height was determined in standing position to the nearest millimeter by a Harpenden stadiometer (Holtain Limited, Dyfed, U.K.). Weight was measured using a mechanical personal scale (SECA, Almere, the Netherlands). Height and weight were measured without shoes and heavy clothing and body mass index (BMI) was calculated (kg/m^2). Age- and sex-specific standard deviation (SD) scores for height, weight, and BMI were obtained using Dutch reference growth charts.[31] Total body, android, and gynoid fat mass were measured using dual-energy X-ray absorptiometry (DXA) (iDXA, General Electrics – Lunar, 2008, Madison, WI) using enCORE software v.13.6, as described in detail previously.[32] Body fat percentage was calculated and android fat mass was divided by gynoid fat mass to obtain the ratio. Age- and sex-specific SD scores for body fat percentage and android:gynoid ratio were calculated based on data from our study population.

Covariates

Maternal height and weight were measured at enrolment and BMI (kg/m^2) was calculated. Data on maternal age, educational level, household income, parity, and folic acid supplement use were available from self-administered questionnaires at enrolment. Maternal educational level was categorized as no/primary; secondary; or higher education. Net monthly household income was categorized as lower than 1,400 euro; 1,400 to 2,200 euro; or higher than 2,200 euro. Parity was categorized as nulliparous or multiparous. Smoking and alcohol consumption during pregnancy were assessed using questionnaires in each trimester and categorized as never; until pregnancy was known; or continued during pregnancy. Folic acid supplement use during pregnancy was categorized as none during early pregnancy; during the first 10 weeks of pregnancy; or periconceptional. Child's methylenetetrahydrofolate reductase (*MTHFR)* variants C677T (rs1801133) and A1298C (rs1801131) were genotyped in cord blood.[33] Information on child's sex, ethnicity, gestational age at birth, and birth weight were obtained from hospital medical records. Ethnicity was defined as Dutch or non-Dutch according to Statistics Netherlands.[34] Information on breastfeeding during infancy was obtained with postnatal questionnaires and categorized as exclusively for at least 4 months; partially in the first 4 months; or never. Infant formula intake at 1 year was obtained from the previously mentioned FFQ. A previously defined child diet score was used to quantify overall diet quality at the age of 1 year.[35] The diet score was calculated for each child using data from the FFQ. Information on child's screen time (television watching and computer use, h/d) and sports participation (yes/no) at the age of 6 years was obtained using a questionnaire.

Statistical analysis

To control for confounding by total energy intake, we used the residual method to adjust the intakes of vitamin B-6, vitamin B-12, folate, folic acid and methionine for total energy intake.[36] In short, linear regression analyses were used to calculate energy adjusted intake of vitamin B-6, vitamin B-12, folate, folic acid and methionine for each subject, with energy intake as independent variable and intake of vitamin B-6, vitamin B-12, folate, folic acid and methionine as dependent variables. Since methionine is an amino acid and therefore part of dietary protein, methionine intake was additionally adjusted for total protein intake with the residual method. After visual inspection, all outcome variables were considered to be normally distributed. Multivariable linear regression models were used to assess associations between intake of vitamin B-6, vitamin B-12, folate, folic acid, DFEs and methionine with age- and sex-adjusted SD scores for growth and body composition outcomes (height, weight, BMI, body fat percentage and android:gynoid ratio). As recommended for dietary intake variables, nutrient intakes were analyzed both as continuous variables and categorized into tertiles, with the lowest tertile as reference category.[36] Potential covariates were selected based on literature or a ≥ 10% change in effect estimates.[37] The adjusted models included maternal age, BMI, education, parity, folic acid supplement use, smoking and alcohol consumption during pregnancy; household income; child's ethnicity, birth weight Z-score, breastfeeding during infancy, intake of formula, total energy intake, diet quality score, screen time, and participation in sports. Considering that vitamin B-12 only occurs in animal foods, models including vitamin B-12 were additionally adjusted for animal protein intake. To reduce potential bias due to attrition, missing values of covariates were multiple imputed based on the correlation between each variable with missing values and other subject characteristics (**Supplementary Table 4.1.1**). Five independent datasets were generated. Because we found similar results, we present results based on the imputed datasets.[38] Because the FFQ was developed and validated for Dutch children, we performed sensitivity analyses in Dutch children only. Previous studies suggest that the influence of several nutrients in early life might differ between boys and girls.[39, 40] Furthermore, it is suggested that nutrients involved in the one-carbon metabolism, such as folate, might interact with *MTHFR* polymorphisms.[41] To assess if associations between intakes of vitamin B-6, vitamin B-12, folate, folic acid and methionine and body composition outcomes differed by sex or by *MTHFR* polymorphisms, we analyzed interaction terms. A statistical interaction was analyzed by adding a product of each of the nutrients (e.g. folic acid) and the potential effect modifier (e.g. sex) as an independent variable to the linear regression model. The linear regression model in which the interaction was tested thus included one of the nutrients, the effect modifier and the interaction term of these two variables as independent variables, and one of the outcomes as dependent variables. Stratified analyses were performed in case the interaction term was significant ($P<0.05$). All statistical analyses were conducted using SPSS version 21.0 (SPSS Inc., Chicago, IL, USA). Results were considered statistically significant with a P value < 0.05. Values are presented as mean (± SD), median [95%range], n(%) and regression coefficients (95%CI lower limit, upper limit).

RESULTS

Subject characteristics

Characteristics of the children and their mothers are presented in **Table 4.1.1**. The majority of the children had a Dutch ethnicity (68.7%) and 49.1 % were boys. Mean (± SD) total energy intake at the median age of 12.9 months (95% range 12.2-18.9) was 1,310 ± 393 kcal/d. Median intakes of vitamin B-6 and vitamin B-12 were comparable to the median intakes among Dutch children aged 2-3 years from the general population,[42] but intake of DFEs were considerably higher (i.e. 182 µg/d vs. 106 µg/d).[42] This high DFEs intake in our population was primarily due to added folic acid in infant formula. Main sources of vitamin B-12 and methionine in our population were meat, fish, and eggs. Various food groups contributed to vitamin B-6 and naturally present folate intake. In non-response analyses, we observed that children without FFQ data had a higher BMI and body fat percentage at the age of 6 years than children with FFQ data (**Supplementary Table 4.1.3**).

Table 4.1.1. Population characteristics (*n=2,922*)

Characteristics	Values[1]
Child characteristics	
Sex, male	1,434 (49.1)
Ethnicity	
Dutch	2,007 (68.7)
Non Dutch	915 (31.3)
Gestational age at birth (wk)	40.1 [36.0-42.3]
Birth weight (g)	3470 ± 552
Breastfeeding	
Exclusively ≥ 4 months	864 (29.6)
Partially ≥ 4 months	1,708 (58.4)
Never	350 (12.0)
Characteristics dietary measurement (1y)	
Age filling out FFQ (months)	12.9 [12.2-18.9]
Total energy intake (kcal/d)	1,310 ± 393
Vitamin B-6 intake (mg/d)	1.2 [0.5-2.2]
Vitamin B-12 intake (µg/d)	2.5 [0.7-5.6]
Folate intake	
Dietary folate (µg/d)	91 [32-209]
Added folic acid (µg/d)	430 [0-1420]
Folate equivalents (µg/d)	182 [67-783]
Methionine intake (mg/d)	642 [197-1390]

Table 4.1.1. Population characteristics (*n=2,922*) (continued)

Characteristics	Values[1]
Characteristics growth measurements (6y)	
Age (years)	6.0 ± 0.2
Height (cm)	118 ± 5
Weight (kg)	22.4 ± 3.4
BMI (kg/m^2)	16.0 ± 1.6
Body fat percentage (%)	24.2 ± 5.1
Android-gynoid fat ratio	0.24 ± 0.06
Sports participation (yes %)	1291 (44.2)
Screen time (h/d)	1.6 [0.3-4.4]
Maternal characteristics	
Age at enrolment (years)	31.5 ± 4.5
BMI at enrolment (kg/m^2)	24.4 [18.8-35.6]
Educational level	
No or primary education	302 (10.3)
Secondary	911 (31.2)
Higher	1,709 (58.5)
Household income per month	
<1,400 euro	414 (14.2)
1,400-2,200 euro	556 (19.0)
>2,200 euro	1,952 (66.8)
Parity	
0	1,761 (60.3)
1	842 (28.8)
2≤	319 (10.9)
Folic acid supplement use	
None	469 (16.0)
First 10 weeks of pregnancy	890 (30.5)
Periconceptional	1,563 (53.5)
Smoking during pregnancy	
Never	2,277 (78.0)
Until pregnancy was known	296 (10.1)
Continued	349 (11.9)
Alcohol consumption during pregnancy	
Never	1,119 (38.3)
Until pregnancy was known	443 (15.2)
Continued	1,359 (46.5)

[1]*Mean ± SD, Median [95% range], or N (%)*

Intake of vitamin B-6, vitamin B-12, folate, folic acid and methionine and growth and body composition outcomes

Table 4.1.2 shows the covariate-adjusted associations between intake of vitamin B-6, vitamin B-12, folate, folic acid and methionine and growth and body composition outcomes. Children with intake of vitamin B-12 in the highest tertile had a 0.13 SD higher android:gynoid ratio (95%CI 0.00;0.25) compared to children in the lowest tertile. In addition, for the association between vitamin B-12 with android:gynoid ratio *P* for trend was significant (*P*=0.03). Additional adjustment for saturated fat intake did not change this result (data not shown). In contrast, children with intakes of folic acid in the highest tertile had a 0.16 SD lower weight (95% CI -0.31; -0.02) and a 0.14 SD lower BMI (95% CI -0.26; -0.01) compared to children with intakes in the lowest tertile. In line with this, children with intakes of DFEs in the second and highest tertile had a lower weight, -0.10 SD (95%CI -0.19, -0.01) and -0.11 SD (95%CI -0.00, -0.22) respectively, compared to children with intakes in the lowest tertile. However, intake of DFEs in the third tertile was also associated with a lower height (-0.12 SD (95%CI -0.23,-0.01)) compared to intake in the lowest tertile. Children with intakes in the highest tertile of methionine had a 0.09 SD higher BMI (95%CI 0.01, 0.17), and a 0.12 SD android:gynoid ratio (95%CI 0.02, 0.22). In addition, P for trend was significant for body fat percentage (*P*=0.03) and android:gynoid ratio (*P*=0.01). In the covariate-adjusted models, vitamin B-6 and folate intakes were not associated with body composition outcomes.

Results of the crude analyses are presented in **Supplementary Table 4.1.4**, showing associations in similar directions to the adjusted models. However, in the crude models higher folate and vitamin B-12 intakes were associated with a higher weight and BMI. These results were no longer significant in the adjusted model.

Additional analyses

Analyses restricted to Dutch children (*n=2,002*) were similar to those in the whole study population (**Supplementary Table 4.1.5**). Only two out of 30 interaction terms for *MTHFR* polymorphisms were significant (folate and folic acid on height), however, stratification on *MTHFR* for these two models revealed no clear differences in associations (data not shown). Interactions of sex with vitamin B-12, folic acid and methionine intake were significant for height and weight ($p<0.05$), but not for BMI, body fat percentage, or android:gynoid ratio. We stratified all analyses for sex (**Supplementary Tables 4.1.6 and 4.1.7**), but analysis stratified for sex showed effect estimates in the same direction as the whole group. Only for vitamin B-6 intake, results differed between girls and boys. Girls with higher vitamin B-6 intake had a lower BMI, while boys had a higher body fat percentage.

Table 4.1.2. Covariate-adjusted associations between intake of vitamin B-6, vitamin B-12, folate, folic acid and methionine and growth and body composition outcomes in children (*n=2,922*).[1]

	Height (SDS)	Weight (SDS)	BMI (SDS)	Body fat % (SDS)	Android:gynoid ratio (SDS)
Vitamin B-6, mg/d					
Tertile 1 (<1.1)	Reference	Reference	Reference	Reference	Reference
Tertile 2 (1.1-1.3)	-0.01 [-0.09;0.09]	0.04 [-0.05;0.12]	0.05 [-0.03;0.12]	0.09 [-0.00;0.18]	-0.00 [-0.10;0.09]
Tertile 3 (>1.3)	-0.04 [-0.14;0.06]	-0.05 [-0.15;0.04]	-0.04 [-0.12;0.05]	0.02 [-0.07;0.12]	-0.01 [-0.11;0.09]
P-trend	*0.63*	*0.20*	*0.15*	*0.19*	*0.36*
Vitamin B-12[2], µg/d					
Tertile 1 (<2.2)	Reference	Reference	Reference	Reference	Reference
Tertile 2 (2.2-2.9)	-0.01 [-0.11;0.08]	-0.01 [-0.10;0.08]	-0.00 [-0.08;0.08]	-0.06 [-0.15;0.04]	0.00 [-0.10;0.10]
Tertile 3 (>2.9)	-0.01 [-0.12;0.11]	0.06 [-0.06;0.17]	0.08 [-0.02;0.18]	0.04 [-0.08;0.16]	0.13 [0.00;0.25]*
P-trend	*0.35*	*0.51*	*0.66*	*0.15*	*0.03**
Folate, µg/d					
Tertile 1 (<80.9)	Reference	Reference	Reference	Reference	Reference
Tertile 2 (80.9-110.3)	0.03 [-0.07;0.13]	0.08 [-0.02;0.18]	0.08 [-0.01;0.16]	0.05 [-0.05;0.15]	-0.01 [-0.15;0.13]
Tertile 3 (>110.3)	-0.02 [-0.15;0.10]	0.06 [-0.06;0.18]	0.09 [-0.01;0.20]	0.05 [-0.08;0.17]	0.03 [-0.15;0.20]
P-trend	*0.12*	*0.58*	*0.59*	*0.74*	*0.63*
Folic acid, µg/d					
Tertile 1 (<304.3)	Reference	Reference	Reference	Reference	Reference
Tertile 2 (304.3-669.5)	-0.03 [-0.13;0.07]	-0.06 [-0.16;0.03]	-0.06 [-0.15 ;0.03]	-0.04 [-0.14;0.06]	-0.09 [-0.19;0.02]
Tertile 3 (>669.5)	-0.10 [-0.25;0.05]	-0.16 [-0.31;-0.02]*	-0.14 [-0.26;-0.01]*	-0.06 [-0.21;0.09]	-0.09 [-0.24;0.07]
P-trend	*0.60*	*0.13*	*0.06*	*0.12*	*0.93*
Folate equivalent, µg/d					
Tertile 1 (<167.8)	Reference	Reference	Reference	Reference	Reference

Table 4.1.2. Covariate-adjusted associations between intake of vitamin B-6, vitamin B-12, folate, folic acid and methionine and growth and body composition outcomes in children (*n=2,922*).[1] (continued)

	Height (SDS)	Weight (SDS)	BMI (SDS)	Body fat % (SDS)	Android:gynoid ratio (SDS)
Tertile 2 (167.8-214.4)	-0.07 [-0.17;0.02]	-0.10 [-0.19;-0.01]*	-0.08 [-0.16;0.00]	0.03 [-0.07;0.12]	-0.02 [-0.12;0.08]
Tertile 3 (>241.5)	-0.12[-0.23;-0.01]*	-0.11 [-0.22;-0.00]*	-0.06 [-0.15;0.03]	0.08 [-0.03;0.19]	-0.03 [-0.15;0.09]
P-trend	*0.75*	*0.14*	*0.05*	*0.67*	*0.61*
Methionine[3], mg/d					
Tertile 1 (<681.5)	Reference	Reference	Reference	Reference	Reference
Tertile 2 (681.5-682.2)	0.01[-0.08;0.09]	0.00[-0.09;0.09]	0.00 [-0.08;0.08]	-0.04 [-0.13;0.05]	-0.00 [-0.10;0.09]
Tertile 3 (>682.2)	0.03[-0.06;0.12]	0.08[-0.00;0.17]	0.09 [0.01;0.17]*	0.08 [-0.02;0.17]	0.12 [0.02;0.22]*
P-trend	*0.87*	*0.70*	*0.65*	*0.03**	*0.01**

[1]Values are regression coefficients [95% confidence interval] from linear regression models and reflect the difference in body composition for tertiles of vitamin B-6, vitamin B-12, folate, folic acid and methionine intake, as compared to the lowest tertile. Tertiles were based on a division of 33.3% of the total study population per tertile of nutrient intake. Tests for trend were conducted with nutrient intake as a continuous variable in the regression model. All outcomes are age- and sex-specific SD scores. Nutrient intakes are adjusted for total energy intake using the residual method.

Models are adjusted for child ethnicity, birth weight Z-score, breastfeeding during infancy, infant formula intake (energy adjusted), total energy intake, diet quality score, participation in sports, screen-time; and maternal age, maternal BMI, maternal education, household income, parity, folic acid supplement use during pregnancy, smoking during pregnancy, and alcohol use during pregnancy.

[2] Vitamin B-12 is additionally adjusted for animal protein intake

[3] Methionine is additionally adjusted for total protein intake using the residual method.

** P <0.05*

DISCUSSION

In this population-based prospective cohort study, we investigated associations between dietary intake of vitamin B-6, vitamin B-12, folate, folic acid and methionine in early childhood and growth and body composition outcomes at the age of 6 years. We hypothesized that higher intake of vitamin B-6, vitamin B-12, folate, folic acid, and methionine were associated with favorable body composition outcomes, since these nutrients play a role in DNA methylation, which may be involved in obesity risk.[13] We observed that folic acid intake in the highest tertile was associated with a lower body weight and BMI, which is in line with our hypothesis. However, in contrast to our hypothesis, a higher methionine intake was associated with a higher BMI, a higher body fat percentage and a higher android:gynoid ratio, and a higher vitamin B-12 intake was associated with a higher android:gynoid ratio. Finally, there were no consistent associations between vitamin B-6 intake and any of the body composition outcomes in the total group. However, when analyses were stratified for sex, a higher vitamin B-6 intake was associated with a higher body fat percentage in boys, but with a lower BMI in girls.

Our findings regarding vitamin B-6 are in line with findings of the study by Gunanti et al. in Mexican American children, which also did not find an association between vitamin B-6 intake and body composition.[17] However, two cross-sectional studies in adults reported that vitamin B-6 levels were lower in obese subjects or patients with metabolic syndrome compared to controls.[43, 44] These findings are in concurrence with our finding that a higher vitamin B-6 intake was associated with a lower BMI in girls. However, in boys a higher intake was associated with a higher body fat percentage. There was no difference in vitamin B-6 intake between boys and girls, suggesting that divergent results might be due to sex-specific differences.

In our study, we observed that a folic acid intake in the highest tertile was associated with a lower weight and a lower BMI, but not with body fat percentage or android:gynoid ratio. Similar patterns of associations were observed for intake of DFEs, probably because folic acid intake was the main contributor to the intake of DFEs.[30] In line with this, a previous cross-sectional study in Brazilian children aged 10-19 years reported that a higher total folate intake was associated with a lower waist circumference.[19] However, in this latter study no distinction was made between natural folate and added folate (folic acid). Moreover, Gunanti et al.,[17] Huemer et al.,[20] and Gallistl et al.[21] observed unfavorable body composition outcomes in relation to lower folate levels in children and adolescents aged 8-15, 2-17, and 4-17 years, respectively. However, two cross-sectional studies did not observe associations between folate levels and overweight in children and adolescents aged 6-17 years.[45, 46]

In contrast to our results for folic acid, intakes of folate were not associated with any of the body composition outcomes. And in sensitivity analyses including Dutch children only, a higher folate intake was even associated with a higher BMI (*P* for trend = 0.03). Only one study reported similar results. In this cross-sectional study levels of folate were significantly higher in overweight and obese boys compared to boys with a normal weight.[18] Potential mechanisms behind these associations remain unclear, but may involve DNA synthesis induced by folate. Intake of folate and vitamin B-12 has been associated with increased DNA synthesis,[47] which stimulates growth and might also increase BMI.

In line with this latter mechanism, vitamin B-12 intakes were associated with a higher android:gynoid ratio. However, these findings are not in agreement with findings from three previous studies, which reported higher vitamin B-12 levels among children and adolescents aged 8-15, 2-17, and 10-19 years with a lower BMI compared to those with higher BMI.[17, 18, 20] These conflicting results might be explained by the fact that previous studies used blood levels of vitamin B-12, rather than dietary vitamin B-12 intakes. Another reason for our unexpected findings might be residual confounding. Vitamin B-12 only occurs in animal products, which are usually also high in energy, protein and fat. However, we adjusted our models for total energy intake, animal protein intake, diet quality and saturated fat. Other possible explanations for this unexpected finding might be other mechanisms involved, such as DNA synthesis. As well as folate, vitamin B-12 is also associated with an increased DNA synthesis, and this could lead to a stimulated growth, which might explain our a higher android:gynoid ratio among children with a higher vitamin B-12 intake. Potential mechanisms underlying these associations should be further investigated in future studies.

We observed that a higher methionine intake, independent of total protein intake, was associated with a higher BMI, body fat percentage and android:gynoid ratio. Although this is in contrast with our hypothesis focusing on DNA methylation, it is in line with a number of previous studies. Results from a randomized controlled trial in adults with metabolic syndrome conducted by Plaisance et al. suggested that a methionine-restricted diet led to an increase in fat oxidation and decrease in intrahepatic lipid content.[48] Previous animal studies performed by Hasek et al. and Malloy et al. showed that rats fed methionine-restricted diets had limited fat deposition and reduced visceral fat mass.[49, 50] These results suggest that a potential beneficial effect on body composition of higher DNA methylation due to a higher methionine intake might be overruled by other mechanisms, such as decreased energy expenditure.[49] For methionine, as well as other nutrients, we observed larger effect estimates in girls compared to boys. A possible explanation for these different findings may be due to differences in adiposity rebound between boys and girls. This onset of rapid growth in body fat occurs around the age of 6 years[51] and seems to occur earlier in girls than in boys.[52]

The current study has several strengths and limitations. An important strength is the size of the population included and the prospective, population-based design. There was large amount of information available on potentially confounding sociodemographic and lifestyle factors of the children and their parents. However, residual confounding may still be an issue, as in any observational study. For instance, detailed information on physical activity levels of children was not available. However, we adjusted for proxies of physical activity, including participation in sports and screen time. In addition, a diet score was used to control for overall healthy diet. Detailed body fat measurements were available by using DXA, which is a more objective measure as compared to several anthropometric measures, such as skinfold thickness. We aimed to reduce attrition bias as much as possible. Therefore, we used a multiple imputation procedure, which is an appropriate method to deal with missing data.[38]

A limitation of our study was the fact that the FFQ was only validated for Dutch children, while we had a multi-ethnic study population. Furthermore, the FFQ was validated for several nutrients, but not for the nutrients used in the current study.[26] This might have led to measurement error in the assessment of vitamin B-6, vitamin B-12, folate, folic acid and methionine intakes, which may have led to random misclassification, and consequently attenuation of effect estimates.[53] However, we adjusted all the nutrients for energy intake using the nutrient residual method, which has been de-

scribed by Kipnis *et al.* as an accurate approach for reducing attenuated effect estimates in nutritional epidemiology.[54] In addition, we analyzed the nutrient intakes in tertiles to avoid misspecification and to decrease the influence of outliers.[36] Unfortunately, as the intake tertiles show, the variation of vitamin B-6 and vitamin B-12 intake in our population is quite low, which made it less likely to find a significant association with body composition outcomes. Furthermore, there might be some distortion in calculating the amount of DFEs in human milk. Folate in human milk is considered naturally present, but might have the same bioavailability as added folic acid. However, the number of children still being breastfed at the age of one year was very small in our population (*n=267*), therefore the impact this error might have on our results is minor. The FFQ was only sent to mothers of participating children who mastered the Dutch language sufficiently, of whom 72% returned the questionnaire, which may have led to selection bias. We observed that children without FFQ data were taller and heavier than children with FFQ data (**Supplementary Table 4.1.3**). However, this would only result in selection bias if the associations between intake of vitamin B-6, vitamin B-12, folate, folic acid and methionine and body composition outcomes were different between included and excluded participants in the final analysis. Unfortunately, blood levels of nutrients were not available, which would be a more objective measurement of nutritional status than dietary intake. Also, we did not have information on choline intake, which is another important nutrient involved in DNA methylation. Finally, an important limitation of our study is that we had no information available on DNA methylation. We could therefore not examine whether intake of vitamin B-6, vitamin B-12, folate, folic acid and methionine were associated with child body composition via changes in DNA methylation, as suggested by animal studies. Future studies in population-based samples of children are needed to investigate this effect.

CONCLUSIONS

Results of this prospective cohort study suggest that high folic acid intake during early childhood may be associated with a lower weight and BMI at the age of 6 years. In contrast, a higher methionine intake was associated with a higher BMI, body fat percentage and android:gynoid ratio. Future studies are needed to confirm these findings, to investigate the underlying mechanisms and examine long term consequences of these outcomes on health.

SUPPLEMENTARY MATERIAL

Supplementary Table 4.1.1. Details of the multiple imputation procedure.

Software used:	IBM SPSS Statistics 21
Imputation method and key settings:	Markov Chain Monte Carlo (MCMC) method
Number of imputed data sets created:	5
Variable included in the imputation procedure (imputed or used as predictors of missing data):	Age at DXA measurement; gender; ethnicity; vitamin B-6 intake; vitamin B-12 intake; folate intake; folic acid intake; folate equivalents intake; methionine intake; height-for-age at age 6; weight-for-age at age 6; BMI-for-age at age 6; body fat percentage at age 6; android:gynoid ratio at age 6; SDS weight at 2, 3,4, 5-10, 10-13, 13-17, 17-23, 23-29, 29-35, 23-35, 35-44 and 44-56 months; SDS height at 2, 3,4, 5-10, 10-13, 13-17, 17-23, 23-29, 29-35, 23-35, 35-44 and 44-56 months; SDS BMI at 2, 3,4, 5-10, 10-13, 13-17, 17-23, 23-29, 29-35, 23-35, 35-44 and 44-56 months; diet score; total protein intake; animal protein intake; breastfeeding during infancy; formula intake; energy intake; sports participation; screen time; maternal age; maternal BMI at enrolment; maternal education; family income; parity; folic acid supplementation during pregnancy; smoking during pregnancy; alcohol consumption during pregnancy.
Variables additionally added as predictors of missing data to increase plausibility of missing at random assumption:	Maternal prepregnancy BM;, maternal BMI in second trimester; maternal ethnicity; gestational age at intake; paternal education; paternal ethnicity; paternal BMI; age biological father; gestational age at birth; TV watching at 2, 3 and 4 years; marital status; age until child was breastfed; age introduction fruit/vegetables; walking to school; biking to school; playing outside.
Treatment of non-normally distributed variables:	Predictive mean matching
Treatment of binary/categorical variables:	Logistic regression models

Supplementary Table 4.1.2. Population characteristics before and after imputation (*n=2,922*).

Characteristics	Values before imputation[1]	Values after imputation[1]
Child characteristics		
Sex, male	1,434 (49.1)	NI
Missing	-	-
Ethnicity		
Dutch	2,003	2,007 (68.7)
Non Dutch	908	915 (31.3)
Missing	*11 (0.4)*	-
Gestational age at delivery (wk)	40.1 [36.0-42.3]	NI
Missing (%)	4 (0.001)	-
Birth weight (g)	3,470 ± 552	NI
Missing	2 (0.0006)	-
Breastfeeding		
Exclusively ≥ 4 months	777 (26.6)	864 (29.6)
Partially ≥ 4 months	1,639 (56.1)	1,708 (58.4)
Never	210 (7.2)	350 (12.0)
Missing	*296 (10.1)*	-
Characteristics dietary measurement (1y)		
Missing	-	
Age filling out FFQ (months)	12.9 [12.2-18.9]	NI
Missing	-	-
Total energy intake (kcal/d)	1,310 ± 393	NI
Vitamin B-6 (mg/d)	1.2 [0.5-2.2]	NI
Vitamin B-12 intake (µg/d)	2.5 [0.7-5.6]	NI
Folate intake		
Dietary folate (µg/d)	91 [32-209]	NI
Added folic acid (µg/d)	430 [0-1,420]	NI
Folate equivalents (µg/d)	182 [67-783]	NI
Methionine (mg/d)	642 [197-1,390]	NI
Characteristics growth measurements (6y)		
Age (years)	6.0 ± 0.2	NI
Missing	-	-
Height (cm)	118 ± 5	NI
Missing	1	-
Weight (kg)	22.4 ± 3.4	NI
Missing	1	-
BMI (kg/m^2)	16.0 ± 1.6	NI
Missing	1	-
Body fat percentage (%)	24.2 ± 5.1	NI
Missing	64 (0.02)	-

Supplementary Table 4.1.2. Population characteristics before and after imputation (*n=2,922*). (continued)

Characteristics	Values before imputation[1]	Values after imputation[1]
Android-gynoid fat ratio	0.24 ± 0.06	NI
Missing	64 (0.02)	-
Sports participation (yes %)	1,206 (44.6)	1,291 (44.2)
Missing	215 (0.07)	-
Screen time (h/d)	1.2 [0.3-4.4]	1.6 [0.3-4.4]
Missing	533 (18.2)	-
Maternal characteristics		
Age at enrolment (years)	31.5 ± 4.5	NI
Missing	-	-
BMI at enrolment (kg/m^2)	23.5 [18.8-35.6]	24.4 [18.8-35.6]
Missing	238 (0.08)	-
Highest completed education level		
No or primary education	38 (1.3)	302 (10.3)
Secondary	757 (25.9)	911 (31.2)
Higher	1,549 (53.0)	1,709 (58.5)
Missing	*578 (19.8)*	-
Household income per month		
<1,400 euro	315 (10.8)	414 (14.2)
1,400-2,200 euro	459 (15.7)	556 (19.0)
>2,200 euro	1,715 (58.7)	1,952 (66.8)
Missing	*433 (14.8)*	-
Parity		
0	1,724 (59.0)	1,761 (60.3)
1	815 (27.9)	842 (28.8)
2≤	305 (10.4)	319 (10.9)
Missing	*78 (2.7)*	-
Folic acid supplement use		
None	310 (10.6)	469 (16.0)
First 10 weeks of pregnancy	647 (22.1)	890 (30.5)
Periconceptional	1,182 (40.5)	1,563 (53.5)
Missing	*783 (26.8)*	-
Smoking during pregnancy		
Never	2,056 (70.4)	2,277 (78.0)
Until pregnancy was known	263 (9.0)	296 (10.1)
Continued	312 (10.7)	349 (11.9)
Missing	*291 (10.0)*	-
Alcohol consumption during pregnancy		
Never	924 (31.6)	1,119 (38.3)

Supplementary Table 4.1.2. Population characteristics before and after imputation (*n=2,922*). (continued)

Characteristics	Values before imputation[1]	Values after imputation[1]
Until pregnancy was known	367 (12.6)	443 (15.2)
Continued	1,116 (38.2)	1,359 (46.5)
Missing	*515 (17.6)*	-

[1]*Mean ± SD,*
Median [95% range],
or N (%)
Abbreviations: NI, not imputed.

Supplementary Table 4.1.3. Non response analysis.

	With FFQ [12] (*n=2,922*)	Without FFQ [12] (*n=3,757*)	*P value* [3]
Age (y)	6.0 ± 0.2	6.4 ± 0.6	*<0.01*
Height (cm)	118 ± 5	120 ± 7	*<0.01*
Weight (kg)	22.4 ± 3.4	24.0 ± 4.7	*<0.01*
BMI (kg/m^2)	16.0 ± 1.6	16.4 ± 2.1	*<0.01*
Body fat percentage(%)	23.9 ± 5.0	25.2 ± 6.0	*<0.01*
Android:gynoid ratio	0.24 ± 0.06	0.26 ± 0.07	*<0.01*

[1] *Differences in growth and body composition outcomes at the age of 6 years between children with and without dietary data available at the age of 13 months.*
[2] *Values are means ± SD*
[3] *P values are obtained from independent sample t-tests.*

Supplementary Table 4.1.4. Crude associations between intake of vitamin B-6, vitamin B-12, folate, folic acid and methionine and growth and body composition outcomes in children (*n=2,922*). [1]

	Height (SDS)	Weight (SDS)	BMI (SDS)	Body fat % (SDS)	Android:gynoid ratio (SDS)
Vitamin B-6, mg/d					
Tertile 1 (<1.1)	Reference	Reference	Reference	Reference	Reference
Tertile 2 (1.1-1.3)	0.04 [-0.04;0.13]	0.07 [-0.02;0.16]	0.06 [-0.02;0.14]	0.07 [-0.03;0.15]	-0.02 [-0.11;0.07]
Tertile 3 (>1.3)	0.02 [-0.07;0.10]	0.04 [-0.05;0.13]	0.04 [-0.03;0.12]	0.08 [-0.01;0.17]	0.02 [-0.07;0.11]
P-trend	*0.78*	*0.45*	*0.38*	*<0.01**	*0.08*
Vitamin B-12, µg/d					
Tertile 1 (<2.2)	Reference	Reference	Reference	Reference	Reference
Tertile 2 (2.2-2.9)	0.01 [-0.08;0.10]	-0.01 [-0.10;0.08]	-0.01 [-0.09;0.07]	-0.04 [-0.13;0.05]	-0.01 [-0.10;0.08]
Tertile 3 (>2.9)	0.06 [-0.02;0.15]	0.12 [0.04;0.21]*	0.11 [0.03;0.19]*	0.08 [-0.01;0.17]	0.11 [0.02;0.20]*
P-trend	*0.56*	*0.10*	*0.08*	*<0.01**	*<0.01**
Folate, µg/d					
Tertile 1 (<80.9)	Reference	Reference	Reference	Reference	Reference
Tertile 2 (80.9-110.3)	0.07 [-0.02;0.15]	0.11 [0.02;0.20]*	0.09 [0.01;0.17]*	0.05 [-0.04;0.14]	0.02 [-0.07;0.11]
Tertile 3 (>110.3)	0.09 [0.00;0.18]	0.18 [0.09;0.26]*	0.16 [0.09;0.24]*	0.08 [-0.01;0.17]	0.02 [-0.07;0.11]
P-trend	*0.32*	*<0.01**	*<0.01**	*0.06*	*0.40*
Folic acid, µg/d					
Tertile 1 (<304.3)	Reference	Reference	Reference	Reference	Reference
Tertile 2 (304.3-669.5)	-0.08 [-0.17;0.01]	-0.12 [-0.21;-0.04]*	-0.10 [-0.18;-0.03]*	-0.04 [-0.13;0.05]	-0.07 [-0.16;0.02]
Tertile 3 (>669.5)	-0.11 [-0.20;-0.03]*	-0.18 [-0.26;-0.09]*	-0.15 [-0.22;-0.07]*	-0.05 [-0.14;0.04]	-0.04 [-0.13;0.05]
P-trend	*0.04**	*<0.01**	*<0.01**	*0.86*	*0.80*
Folate equivalent, µg/d					
Tertile 1 (<167.8)	Reference	Reference	Reference	Reference	Reference

Supplementary Table 4.1.4. Crude associations between intake of vitamin B-6, vitamin B-12, folate, folic acid and methionine and growth and body composition outcomes in children (*n=2,922*). [1] (continued)

	Height (SDS)	Weight (SDS)	BMI (SDS)	Body fat % (SDS)	Android:gynoid ratio (SDS)
Tertile 2 (167.8-214.4)	-0.07 [-0.16;0.02]	-0.13 [-0.22;-0.04]*	-0.12 [-0.20;-0.05]*	-0.02 [-0.11;0.07]	-0.04 [-0.13;0.05]
Tertile 3 (>241.5)	-0.14 [-0.22;-0.05]*	-0.12 [-0.21;-0.04]*	-0.07 [-0.14;0.01]	0.08 [-0.01;0.17]	-0.01 [-0.10;0.08]
P-trend	*0.79*	*0.43*	*0.32*	*0.50*	*0.21*
Methionine[2], mg/d					
Tertile 1 (<681.5)	Reference	Reference	Reference	Reference	Reference
Tertile 2 (681.5-682.2)	0.01[-0.8;0.10]	-0.00[-0.09;0.08]	-0.01[-0.09;0.07]	-0.04[-0.13;0.05]	-0.01[-0.10;0.08]
Tertile 3 (>682.2)	0.07[-.02;0.16]	0.13 [0.04;0.21]*	0.11[0.03;0.19]*	0.08[-0.01;0.17]	0.11[0.02;0.20]
P-trend	*0.56*	*0.10*	*0.09*	*<0.01**	*0.01**

[1] Values are regression coefficients [95% confidence interval] from linear regression models and reflect the difference in body composition for tertiles of vitamin B-6, vitamin B-12, folate, folic acid and methionine intake, as compared to the lowest tertile. Tertiles were based on a division of 33.3% of the total study population per tertile of nutrient intake. Nutrient intakes are adjusted for total energy intake using the residual method. Tests for trend were conducted with nutrient intake as a continuous variable in the regression model. All outcomes are age- and sex-specific SD scores.

[2] Methionine is additionally adjusted for total protein intake using the residual method.

** P <0.05*

Supplementary Table 4.1.5. Covariate-adjusted associations between intake of vitamin B-6, vitamin B-12, folate, folic acid and methionine and growth and body composition outcomes in Dutch children only (*n=2,007*).[1]

	Height (SDS)	Weight (SDS)	BMI (SDS)	Body fat % (SDS)	Android:gynoid ratio (SDS)
Vitamin B-6, mg/d					
Tertile 1 (<1.1)	Reference	Reference	Reference	Reference	Reference
Tertile 2 (1.1-1.3)	-0.04[-0.15;0.07]	0.03[-0.07;0.13]	0.07[-0.02;0.15]	0.11[0.02;0.21]*	0.03[-0.08;0.13]
Tertile 3 (>1.3)	-0.07[-0.19;0.04]	-0.05[-0.16;0.05]	-0.02[-0.12;0.08]	0.04[-0.06;0.15]	0.02[-0.10;0.13]
P-trend	*0.49*	*0.32*	*0.34*	*0.09*	*0.22*
Vitamin B-12[2], µg/d					
Tertile 1 (<2.2)	Reference	Reference	Reference	Reference	Reference
Tertile 2 (2.2-2.9)	-0.01[-0.12;0.10]	-0.00[-0.10;0.10]	0.01[-0.09;0.10]	-0.03[-0.13;0.08]	0.05[-0.06;0.16]
Tertile 3 (>2.9)	-0.01[-0.15;0.12]	0.05 [-0.07;0.18]	0.07[-0.04;0.19]	0.07[-0.07;0.19]	0.15[0.01;0.29]*
P-trend	*0.75*	*0.45*	*0.23*	*0.01**	*0.01**
Folate, µg/d					
Tertile 1 (<80.9)	Reference	Reference	Reference	Reference	Reference
Tertile 2 (80.9-110.3)	0.13[0.01;0.25]*	0.14[0.03;0.25]*	0.08[-0.02;0.18]	0.05[-0.07;0.16]	0.04[-0.08;0.16]
Tertile 3 (>110.3)	0.02[-0.13;0.16]	0.11[-0.02;0.24]	0.13[0.00;0.25]*	0.00[-0.14;0.14]	0.00[0.15;0.15]
P-trend	*0.31*	*0.35*	*0.03**	*0.36*	*0.39*
Folic acid, µg/d					
Tertile 1 (<304.3)	Reference	Reference	Reference	Reference	Reference
Tertile 2 (304.3-669.5)	-0.09[-0.20;0.03]	-0.10[-0.21;0.01]	-0.07 [-0.17;0.03]	0.03[-0.08;0.14]	-0.04 [-0.16;0.08]
Tertile 3 (>669.5)	-0.16[-0.36;0.02]	-0.16[-0.33;0.00]	-0.10[-0.25;0.05]	0.04 [-0.12;0.21]	0.03 [-0.15;0.21]
P-trend	*0.09*	*0.01**	*0.01**	*0.15*	*0.60*
Folate equivalent, µg/d					
Tertile 1 (<167.8)	Reference	Reference	Reference	Reference	Reference

Supplementary Table 4.1.5. Covariate-adjusted associations between intake of vitamin B-6, vitamin B-12, folate, folic acid and methionine and growth and body composition outcomes in Dutch children only (*n=2,007*).[1] (continued)

	Height (SDS)	Weight (SDS)	BMI (SDS)	Body fat % (SDS)	Android:gynoid ratio (SDS)
Tertile 2 (167.8-214.4)	-0.06 [-0.17;0.05[	-0.06[-0.16;0.04]	-0.04[-0.14;0.05]	0.04[-0.07;0.014]	0.03[-0.09;0.44]
Tertile 3 (>241.5)	-0.13[-0.62;0.00]	-0.08[-0.20;0.04]	-0.02[-0.13;0.10]	0.07[-0.06;0.19]	-0.03 [-0.16;0.11]
P-trend	*0.74*	*0.57*	*0.05**	*0.82*	*0.27*
Methionine[3], mg/d					
Tertile 1 (<681.5)	Reference	Reference	Reference	Reference	Reference
Tertile 2 (681.5-682.2)	0.01[-0.09;0.12]	0.01[-0.08;0.11]	0.01[-0.07;0.10]	-0.02[-0.12;0.08]	0.04[-0.06;0.15]
Tertile 3 (>682.2)	0.03[-0.08;0.14]	0.09[-0.02;0.19]	0.09[-0.00;0.18]	0.08[-0.02;0.18]	0.13[0.02;0.24]*
P-trend	*0.63*	*0.10*	*0.06*	*0.01**	*0.02**

[1]Values are regression coefficients [95% confidence interval] from linear regression models and reflect the difference in body composition for tertiles of vitamin B-6, vitamin B-12, folate, folic acid and methionine intake, as compared to the lowest tertile. Tertiles were based on a division of 33.3% of the total study population per tertile of nutrient intake. Tests for trend were conducted with nutrient intake as a continuous variable in the regression model. All outcomes are age- and sex-specific SD scores. Nutrient intakes are adjusted for total energy intake using the residual method.

Models are adjusted for child ethnicity, birth weight Z-score, breastfeeding during infancy, infant formula intake (energy adjusted), total energy intake, diet quality score, participation in sports, screen-time; and maternal age, maternal BMI, maternal education, household income, parity, folic acid supplement use during pregnancy, smoking during pregnancy, and alcohol use during pregnancy.

[2] Vitamin B-12 is additionally adjusted for animal protein intake

[3] Methionine is additionally adjusted for total protein intake using the residual method.

** P <0.05*

Supplementary Table 4.1.6. Covariate-adjusted associations between intake of vitamin B-6, vitamin B-12, folate, folic acid and methionine and growth and body composition outcomes in girls (*n=1,488*).[1]

	Height (SDS)	Weight (SDS)	BMI (SDS)	Body fat % (SDS)	Android:gynoid ratio (SDS)
Vitamin B-6, mg/d					
Tertile 1 (<1.1)	Reference	Reference	Reference	Reference	Reference
Tertile 2 (1.1-1.3)	-0.04[-0.17;0.09]	-0.01[-0.13;0.11]	0.02[-0.09;0.13]	0.05[-0.08;0.17]	-0.01[-0.13;0.12]
Tertile 3 (>1.3)	0.01[-0.13;0.15]	-0.09[-0.22;0.04]	-0.13[-0.25;-0.02]*	-0.13[-0.27;0.00]	-0.10[0.24;0.05]
P-trend	*0.91*	*0.10*	*0.01**	*0.17*	*0.30*
Vitamin B-12[2], µg/d					
Tertile 1 (<2.2)	Reference	Reference	Reference	Reference	Reference
Tertile 2 (2.2-2.9)	-0.07[-0.20;0.07]	-0.02[-0.15;0.11]	0.03[-0.08;0.14]	0.00[-0.13;0.13]	-0.00[-0.14;0.14]
Tertile 3 (>2.9)	0.01[-0.15;0.18]	0.09[-0.06;0.25]	0.11[-0.03;0.25]	0.08[-0.09;0.24]	0.17[-0.00;0.34]
P-trend	*0.79*	*0.50*	*0.29*	*0.31*	*0.10*
Folate, µg/d					
Tertile 1 (<80.9)	Reference	Reference	Reference	Reference	Reference
Tertile 2 (80.9-110.3)	0.02[-0.13;0.16]	0.08[-0.06;0.22]	0.08[-0.04;0.20]	0.11[-0.03;0.25]	0.12[-0.03;0.27]
Tertile 3 (>110.3)	-0.05[-0.23;0.13]	0.07[-0.10;0.24]	0.12[-0.03;0.27]	0.10[-0.07;0.28]	0.07[-0.11;0.25]
P-trend	*0.26*	*0.97*	*0.42*	*0.27*	*0.48*
Folic acid, µg/d					
Tertile 1 (<304.3)	Reference	Reference	Reference	Reference	Reference
Tertile 2 (304.3-669.5)	-0.05[-0.19;0.09]	-0.10[-0.24;0.03]	-0.11[-0.23;0.01]	-0.05[-0.19;0.08]	-0.06[-0.20;0.08]
Tertile 3 (>669.5)	-0.12[-0.33;0.10]	-0.25[-0.47;-0.04]*	-0.25[-0.43;-0.07]*	-0.17[-0.38;0.04]	-0.16[-0.38;0.06]
P-trend	*0.16*	*<0.01**	*<0.01**	*0.16*	*0.11*
Folate equivalent, µg/d					
Tertile 1 (<167.8)	Reference	Reference	Reference	Reference	Reference

Supplementary Table 4.1.6. Covariate-adjusted associations between intake of vitamin B-6, vitamin B-12, folate, folic acid and methionine and growth and body composition outcomes in girls (*n=1,488*).[1] (continued)

	Height (SDS)	Weight (SDS)	BMI (SDS)	Body fat % (SDS)	Android:gynoid ratio (SDS)
Tertile 2 (167.8-214.4)	-0.05[-0.18;0.09]	-0.10[-0.23;0.03]	-0.11[-0.22;0.00]	0.04[-0.09;0.17]	-0.03[-0.17;0.11]
Tertile 3 (>241.5)	-0.17[-0.32;-0.01]*	-0.17[-0.32;-0.03]*	-0.11[-0.24;0.02]	0.08[-0.08;0.23]	-0.02[-0.18;0.13]
P-trend	*0.53*	*0.80*	*0.34*	*0.96*	*0.92*
Methionine[3], mg/d					
Tertile 1 (<681.5)	Reference	Reference	Reference	Reference	Reference
Tertile 2 (681.5-682.2)	-0.04[-0.17;0.08]	-0.01[-0.13;0.11]	0.03[-0.07;0.14]	0.03[-0.09;0.15]	0.02[-0.11;0.14]
Tertile 3 (>682.2)	0.07[-0.65;0.20]	0.14[0.10;0.26]*	0.13[0.01;0.24]*	0.14[0.01;0.27]*	0.19[0.05;0.32]*
P-trend	*0.57*	*0.58*	*0.89*	*0.02**	*0.01**

[1] *Values are regression coefficients [95% confidence interval] from linear regression models and reflect the difference in body composition for tertiles of vitamin B-6, vitamin B-12, folate, folic acid and methionine intake, as compared to the lowest tertile. Tertiles were based on a division of 33.3% of the total study population per tertile of nutrient intake. Tests for trend were conducted with nutrient intake as a continuous variable in the regression model. All outcomes are age- and sex-specific SD scores. Nutrient intakes are adjusted for total energy intake using the residual method.*

Models are adjusted for child ethnicity, birth weight Z-score, breastfeeding during infancy, infant formula intake (energy adjusted), total energy intake, diet quality score, participation in sports, screen-time; and maternal age, maternal BMI, maternal education, household income, parity, folic acid supplement use during pregnancy, smoking during pregnancy, and alcohol use during pregnancy.

[2] *Vitamin B-12 is additionally adjusted for animal protein intake*

[3] *Methionine is additionally adjusted for total protein intake using the residual method.*

** P <0.05*

Supplementary Table 4.1.7. Covariate-adjusted associations between intake of vitamin B-6, vitamin B-12, folate, folic acid and methionine and growth and body composition outcomes in boys (*n=1,434*).[1]

	Height (SDS)	Weight (SDS)	BMI (SDS)	Body fat % (SDS)	Android:gynoid ratio (SDS)
Vitamin B-6, mg/d					
Tertile 1 (<1.1)	Reference	Reference	Reference	Reference	Reference
Tertile 2 (1.1-1.3)	0.04[-0.09;0.17]	0.09[-0.04;0.21]	0.08[-0.03;0.19]	0.14[0.01;0.27]*	0.00[-0.14;0.14]
Tertile 3 (>1.3)	-0.07[-0.21;0.06]	-0.00[-0.14;0.13]	0.06[-0.07;0.18]	0.18[0.04;0.32]*	0.07[-0.08;0.21]
P-trend	*0.53*	*0.93*	*0.65*	*<0.01**	*0.04**
Vitamin B-12[2], µg/d					
Tertile 1 (<2.2)	Reference	Reference	Reference	Reference	Reference
Tertile 2 (2.2-2.9)	0.04[-0.09;0.18]	0.00[-0.13;0.13]	-0.03[-0.15;0.09]	-0.12[-0.26;0.01]	-0.00[-0.15;0.14]
Tertile 3 (>2.9)	-0.02[-0.18;0.14]	0.03[-0.13;0.19]	0.05[-0.09;0.20]	0.00[-0.17;0.17]	-0.00[-0.00;0.00]
P-trend	*0.43*	*0.88*	*0.76*	*0.41*	*0.20*
Folate, µg/d					
Tertile 1 (<80.9)	Reference	Reference	Reference	Reference	Reference
Tertile 2 (80.9-110.3)	0.03[-0.11;0.17]	0.07[-0.06;0.21]	0.08[-0.05;0.20]	-0.01[-0.15;0.14]	-0.06[-0.21;0.09]
Tertile 3 (>110.3)	-0.02[-0.18;0.15]	0.05[-0.12;0.21]	0.07[-0.07;0.22]	0.01[-0.16;0.18]	-0.02[-0.20;0.16]
P-trend	*0.26*	*0.52*	*0.95*	*0.65*	*0.90*
Folic acid, µg/d					
Tertile 1 (<304.3)	Reference	Reference	Reference	Reference	Reference
Tertile 2 (304.3-669.5)	-0.02[-0.16;0.12]	-0.03[-0.17;0.11]	-0.01[-0.14;0.11]	-0.03[-0.18;0.11]	-0.13[-0.28;0.03]
Tertile 3 (>669.5)	-0.07[-0.29;0.14]	-0.08[-0.29;0.13]	-0.03[-0.22;0.16]	0.03[-0.19;0.25]	-0.03[-0.26;0.20]
P-trend	*0.42*	*0.39*	*0.64*	*<0.01**	*0.16*
Folate equivalent, µg/d					
Tertile 1 (<167.8)	Reference	Reference	Reference	Reference	Reference

SupplementaryTable 4.1.7. Covariate-adjusted associations between intake of vitamin B-6, vitamin B-12, folate, folic acid and methionine and growth and body composition outcomes in boys (*n=1,434*).[1] (continued)

	Height (SDS)	Weight (SDS)	BMI (SDS)	Body fat % (SDS)	Android:gynoid ratio (SDS)
Tertile 2 (167.8-214.4)	-0.10[-0.23;0.03]	-0.10[-0.22;0.03]	-0.05[-0.17;0.07]	0.02[-0.12;0.16]	-0.02[-0.16;0.13]
Tertile 3 (>241.5)	-0.08[-0.23;0.07]	-0.06[-0.21;0.10]	-0.01[-0.15;0.13]	0.07[-0.09;0.23]	-0.04[-0.21;0.13]
P-trend	*0.31*	*0.07*	*0.08*	*0.47*	*0.48*
Methionine[3], mg/d					
Tertile 1 (<681.5)	Reference	Reference	Reference	Reference	Reference
Tertile 2 (681.5-682.2)	0.06[-0.07;0.18]	0.01[-0.12;0.13]	-0.03[-0.14;0.08]	-0.12[-0.25;0.01]	-0.02[-0.16;0.11]
Tertile 3 (>682.2)	-0.01[-0.13;0.12]	0.04[-0.09;0.16]	0.05[-0.07;0.16]	0.01[-0.12;0.14]	0.05[-0.09;0.19]
P-trend	*0.56*	*0.93*	*0.65*	*0.44*	*0.35*

[1]Values are regression coefficients [95% confidence interval] from linear regression models and reflect the difference in body composition for tertiles of vitamin B-6, vitamin B-12, folate, folic acid and methionine intake, as compared to the lowest tertile. Tertiles were based on a division of 33.3% of the total study population per tertile of nutrient intake. Tests for trend were conducted with nutrient intake as a continuous variable in the regression model. All outcomes are age- and sex-specific SD scores. Nutrient intakes are adjusted for total energy intake using the residual method.

Models are adjusted for child ethnicity, birth weight Z-score, breastfeeding during infancy, infant formula intake (energy adjusted), total energy intake, diet quality score, participation in sports, screen-time; and maternal age, maternal BMI, maternal education, household income, parity, folic acid supplement use during pregnancy, smoking during pregnancy, and alcohol use during pregnancy.

[2] Vitamin B-12 is additionally adjusted for animal protein intake

[3] Methionine is additionally adjusted for total protein intake using the residual method.

** P <0.05*

REFERENCES

1. Weiss R, Dziura J, Burgert TS, Tamborlane WV, Taksali SE, Yeckel CW, et al. Obesity and the metabolic syndrome in children and adolescents. New England Journal of Medicine. 2004;350(23):2362-74.
2. Crocker MK, Yanovski JA. Pediatric obesity: etiology and treatment. Pediatric Clinics of North America. 2011;58(5):1217-40.
3. Barres R, Zierath JR. DNA methylation in metabolic disorders. The American journal of clinical nutrition. 2011;93(4):897S-900S.
4. Niculescu MD, Zeisel SH. Diet, methyl donors and DNA methylation: interactions between dietary folate, methionine and choline. The Journal of nutrition. 2002;132(8):2333S-5S.
5. Harnack L, Jacobs DR, Nicodemus K, Lazovich D, Anderson K, Folsom AR. Relationship of Folate, Vitamin B-6, Vitamin B-12, and Methionine Intake to Incidence of Colorectal Cancers. Nutrition and Cancer. 2002;43(2):152-8.
6. Friso S, Choi S-W. Gene-nutrient interactions and DNA methylation. The Journal of nutrition. 2002; 132(8):2382S-7S.
7. Obeid R. The Metabolic Burden of Methyl Donor Deficiency with Focus on the Betaine Homocysteine Methyltransferase Pathway. Nutrients. 2013;5(9):3481-95.
8. Glier MB, Green TJ, Devlin AM. Methyl nutrients, DNA methylation, and cardiovascular disease. Molecular Nutrition & Food Research. 2014;58(1):172-82.
9. Fardet A. New hypotheses for the health-protective mechanisms of whole-grain cereals: what is beyond fibre? Nutrition Research Reviews. 2010;23(01):65-134.
10. Hussain N. Epigenetic Influences That Modulate Infant Growth, Development, and Disease. Antioxidants & Redox Signaling. 2012;17(2):224-36.
11. Bertolo RF, McBreairty LE. The nutritional burden of methylation reactions. Current Opinion in Clinical Nutrition and Metabolic Care. 2013;16(1):102-8.
12. Kotsopoulos J, Sohn K-J, Kim Y-I. Postweaning dietary folate deficiency provided through childhood to puberty permanently increases genomic DNA methylation in adult rat liver. The Journal of nutrition. 2008;138(4):703-9.
13. Perng W, Mora-Plazas M, Marín C, Rozek LS, Baylin A, Villamor E. A Prospective Study of LINE-1DNA Methylation and Development of Adiposity in School-Age Children. PLoS ONE. 2013;8(4): e62587.
14. Cordero P, Campion J, Milagro FI, Martinez JA. Transcriptomic and epigenetic changes in early liver steatosis associated to obesity: effect of dietary methyl donor supplementation. Molecular genetics and metabolism. 2013;110(3):388-95.
15. Waterland RA, Travisano M, Tahiliani KG, Rached MT, Mirza S. Methyl donor supplementation prevents transgenerational amplification of obesity. International Journal of Obesity. 2008;32(9):1373-9.
16. Carlin J, George R, Reyes TM. Methyl donor supplementation blocks the adverse effects of maternal high fat diet on offspring physiology. PLoS ONE. 2013;8(5):e63549.
17. Gunanti IR, Marks GC, Al-Mamun A, Long KZ. Low Serum Vitamin B-12 and Folate Concentrations and Low Thiamin and Riboflavin Intakes Are Inversely Associated with Greater Adiposity in Mexican American Children. The Journal of nutrition. 2014;144(12):2027-33.
18. Akanji AO, Thalib L, Al-Isa AN. Folate, vitamin B12 and total homocysteine levels in Arab adolescent subjects: Reference ranges and potential determinants. Nutrition, Metabolism and Cardiovascular Diseases. 2012;22(10):900-6.
19. Vitolo MR, Canal Q, Campagnolo PDB, Gama CM. Factors associated with risk of low folate intake among adolescents. Jornal de pediatria. 2006;82(2):121-6.

20. Huemer M, Vonblon K, Födinger M, Krumpholz R, Hubmann M, Ulmer H, et al. Total homocysteine, folate, and cobalamin, and their relation to genetic polymorphisms, lifestyle and body mass index in healthy children and adolescents. Pediatric research. 2006;60(6):764-9.
21. Gallistl S, Sudi K, Mangge H, Erwa W, Borkenstein M. Insulin is an independent correlate of plasma homocysteine levels in obese children and adolescents. Diabetes Care. 2000;23(9):1348-52.
22. Pereira da Silva N, Suano de Souza FI, Ifanger Pendezza A, Luiz Affonso Fonseca F, Hix S, Oliveira AC, et al. Homocysteine and cysteine levels in prepubertal children: Association with waist circumference and lipid profile. Nutrition. 2013;29(1):166-71.
23. Papandreou D, Mavromichalis I, Makedou A, Rousso I, Arvanitidou M. Reference range of total serum homocysteine level and dietary indexes in healthy Greek schoolchildren aged 6–15 years. British Journal of Nutrition. 2006;96(04):719-24.
24. Shen M-H, Chu N-F, Wu D-M, Chang J-B. Plasma homocyst (e) ine, folate and vitamin B12 levels among school children in Taiwan: The Taipei Children Heart Study. Clinical biochemistry. 2002;35(6): 495-8.
25. Jaddoe VWV, van Duijn CM, Franco OH, van der Heijden AJ, van Iizendoorn MH, de Jongste JC, et al. The Generation R Study: design and cohort update 2012. European journal of epidemiology. 2012; 27(9):739-56.
26. Kiefte-de Jong JC, de Vries JH, Bleeker SE, Jaddoe VWV, Hofman A, Raat H, et al. Socio-demographic and lifestyle determinants of 'Western-like'and 'Health conscious' dietary patterns in toddlers. British Journal of Nutrition. 2013;109(01):137-47.
27. Hulshof K, Breedveld B. Results of the study on nutrient intake in young toddlers 2002. TNO Nutrition: Zeist, The Netherlands. 2002.
28. Voortman TV, Kiefte-de Jong JC, Geelen A, Villamor E, Hofman A, Jaddoe VWV, et al. A diet score to assess diet quality in preschool children: Development and determinants in the Generation R study. Submitted.
29. Center NN. Nevo: Dutch food composition database 2006. The Hague, The Netherlands: Netherlands Nutrition Centre. 2006.
30. Suitor CW, Bailey LB. Dietary folate equivalents: interpretation and application. Journal of the American Dietetic Association. 2000;100(1):88-94.
31. Fredriks AM, Van Buuren S, Burgmeijer RJF, Meulmeester JF, Beuker RJ, Brugman E, et al. Continuing positive secular growth change in The Netherlands 1955–1997. Pediatric research. 2000;47(3):316-23.
32. Heppe DHM, Medina-Gomez C, Hofman A, Franco OH, Rivadeneira F, Jaddoe VWV. Maternal first-trimester diet and childhood bone mass: the Generation R Study. The American journal of clinical nutrition. 2013;98(1):224-32.
33. Jaddoe VWV, Bakker R, van Duijn CM, van der Heijden AJ, Lindemans J, Mackenbach JP, et al. The Generation R Study Biobank: a resource for epidemiological studies in children and their parents. European journal of epidemiology. 2007;22(12):917-23.
34. Statistics Netherlands. Immigrants in the Netherlands 2004 (Allochtonen in Nederland 2004). . Den Haag/Heerlen: Statistics Netherlands (Centraal Bureau voor de Statistiek), 2004.
35. Voortman T, Kiefte-de Jong JC, Geelen A, Villamor E, Moll HA, de Jongste JC, et al. The Development of a Diet Quality Score for Preschool Children and Its Validation and Determinants in the Generation R Study. The Journal of nutrition. 2015:jn. 114.199349.
36. Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. The American journal of clinical nutrition. 1997;65(4):1220S-8S.
37. Mickey RM, Greenland S. The impact of confounder selection criteria on effect estimation. American Journal of Epidemiology. 1989;129(1):125-37.

38. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ: British Medical Journal. 2009;338.
39. Leermakers ETM, Felix JF, Erler NS, Ćerimagić A, Wijtzes AI, Hofman A, et al. Sugar-containing beverage intake in toddlers and body composition up to age 6 years: The Generation R Study. European journal of clinical nutrition. 2015.
40. Simen-Kapeu A, Veugelers PJ. Should public health interventions aimed at reducing childhood overweight and obesity be gender-focused? BMC Public Health. 2010;10(1):340.
41. Friso S, Choi S-W, Girelli D, Mason JB, Dolnikowski GG, Bagley PJ, et al. A common mutation in the 5, 10-methylenetetrahydrofolate reductase gene affects genomic DNA methylation through an interaction with folate status. Proceedings of the National Academy of Sciences. 2002;99(8):5606-11.
42. Ocké MC, Van Rossum CTM, Fransen HP, Buurma EM, De Boer EJ, Brants HAM, et al. Dutch national food consumption survey: young children 2005/2006: RIVM; 2008.
43. Aasheim ET, Hofsø D, Hjelmesæth J, Birkeland KI, Bøhmer T. Vitamin status in morbidly obese patients: a cross-sectional study. The American journal of clinical nutrition. 2008;87(2):362-9.
44. Odum EP, Wakwe VC. Plasma concentrations of water. soluble vitamins in metabolic syndrome subjects. Nigerian journal of clinical practice. 2013;15(4):442-7.
45. Papandreou D, Rousso I, Makedou A, Arvanitidou M, Mavromichalis I. Association of blood pressure, obesity and serum homocysteine levels in healthy children. Acta Paediatrica. 2007;96(12):1819-23.
46. Narin F. leptin and apolipoprotein B levels in childhood obesity. Ann Saudi Med. 2005;25(3):209-14.
47. Obeid R, Herrmann W. Homocysteine, folic acid and vitamin B12 in relation to pre-and postnatal health aspects. Clinical Chemical Laboratory Medicine. 2005;43(10):1052-7.
48. Plaisance EP, Greenway FL, Boudreau A, Hill KL, Johnson WD, Krajcik RA, et al. Dietary methionine restriction increases fat oxidation in obese adults with metabolic syndrome. The Journal of Clinical Endocrinology & Metabolism. 2011;96(5):E836-E40.
49. Hasek BE, Stewart LK, Henagan TM, Boudreau A, Lenard NR, Black C, et al. Dietary methionine restriction enhances metabolic flexibility and increases uncoupled respiration in both fed and fasted states. American Journal of Physiology-Regulatory, Integrative and Comparative Physiology. 2010; 299(3):R728-R39.
50. Molloy AM, Kirke PN, Brody LC, Scott JM, Mills JL. Effects of folate and vitamin B 12 deficiencies during pregnancy on fetal, infant, and child development. Food & Nutrition Bulletin. 2008;29(Supplement 1):101-11.
51. Rolland-Cachera M-F, Deheeger M, Bellisle F, Sempe M, Guilloud-Bataille M, Patois E. Adiposity rebound in children: a simple indicator for predicting obesity. The American journal of clinical nutrition. 1984;39(1):129-35.
52. Boonpleng W, Park CG, Gallo AM. Timing of adiposity rebound: a step toward preventing obesity. Pediatric nursing. 2011;38(1):37-42.
53. Freedman LS, Schatzkin A, Midthune D, Kipnis V. Dealing with dietary measurement error in nutritional cohort studies. Journal of the National Cancer Institute. 2011;103(14):1086-92.
54. Kipnis V, Subar AF, Midthune D, Freedman LS, Ballard-Barbash R, Troiano RP, et al. Structure of dietary measurement error: results of the OPEN biomarker study. American Journal of Epidemiology. 2003;158(1):14-21.

4.2

Dietary intake of protein in early childhood and growth trajectories between 1 and 9 years of age: The Generation R Study

Based on:

Braun KVE, Erler NS, Kiefte-de Jong JC, Jaddoe VW, van den Hooven EH, Franco OH, Voortman T. Dietary Intake of Protein in Early Childhood Is Associated with Growth Trajectories between 1 and 9 Years of Age. *Journal of Nutrition. 2016;146(11):2361-7.*

ABSTRACT

Background: High protein intake in infancy might lead to a higher BMI in childhood. However, whether these associations differ between different sources of protein is unclear.

Objective: We investigated associations between the intake of total protein, protein from different sources, and individual amino acids in early childhood, and repeatedly measured height, weight, and BMI up to the age of 9 years.

Methods: This study was performed in 3,564 children participating in the Generation R Study, a population-based prospective cohort study in Rotterdam, the Netherlands. Intake of total protein, animal protein, vegetable protein and individual amino acids (including methionine, arginine, lysine, threonine, valine, leucine, isoleucine, phenylalanine, tryptophan, histidine, cysteine, tyrosine, alanine, asparagine, glutamine, glycine, proline, and serine) at 1 year was assessed using an FFQ. Height and weight were measured around the ages of 14, 18, 24, 30, 36, 45 months, 6 years, and 9 years, and BMI was calculated.

Results: After adjustment for confounders, linear mixed models showed that a 10 g/d higher total protein intake at 1 year was significantly associated with a 0.03 SD higher height (95%CI 0.00; 0.06), a 0.06 SD higher weight (95%CI 0.03; 0.09), and a 0.05 SD higher BMI (95%CI 0.03; 0.08) up to the age of 9 years. Associations were stronger for animal than for vegetable protein intake, but did not differ between dairy and non-dairy animal protein, or between specific amino acids.

Conclusion: A higher intake of protein, especially animal protein, at 1 year of age was associated with a higher height, weight, and BMI in childhood up to 9 years of age. Future studies should explore the role of growth hormones, and investigate if protein intake in early childhood affects health later in life.

INTRODUCTION

Diet in early childhood plays an important role in development and growth.[1] In particular protein intake is important, because it provides essential amino acids required for protein synthesis, which are necessary for growth.[2, 3] When intake is too low, growth during childhood is restricted.[4] However, in most Western countries protein intake exceeds recommendations and a high protein intake in early childhood might lead to increased growth and a higher risk of obesity. This accelerated growth may be mediated through hormonal responses.[5, 6]

Several observational studies have reported that a higher total protein intake in early childhood is associated with higher height, weight, and BMI.[7, 8] This has also been confirmed in a randomized controlled trial with dairy protein-based infant formula. Results of this study showed that children who received a lower-protein formula during infancy had a lower weight and BMI up to 6 years of age than children who had received a higher-protein infant formula, but that these children were not different with respect to height.[9, 10] Some studies suggest that it is not the total amount of protein intake, but specific sources, such as animal or dairy protein, that may cause accelerated growth.[11, 12] A difference in effect of different types of protein might be due to a different amino acid composition in different sources of protein. It has been suggested that certain amino acids, such as arginine and lysine, may induce a release of growth hormone and insulin, which may lead to accelerated growth.[13, 14]

Even though previous studies showed an association between early protein intake and growth, so far there is only limited evidence on whether associations differ between different sources of protein.[15] Our aim was to explore associations between intake of protein, measured as total protein, different sources of dietary protein, and individual amino acids, and repeatedly measured height, weight, and BMI up to the age of 9 years in a large prospective cohort within a multi-ethnic study population.

METHODS

Design and subjects

This study was embedded in the Generation R Study, a population-based prospective cohort study from fetal life onward in Rotterdam, the Netherlands.[16] The study was approved by the Medical Ethical Committee of Erasmus Medical Center, Rotterdam. Written informed consent was obtained from all participating mothers. All children were born between 2002 and 2006, and data on follow-up in early childhood was available for 7,893 children. A Food Frequency Questionnaire (FFQ) to assess infant diet, including protein intake, was sent to 5,088 mothers who provided informed consent for postnatal follow-up and mastered the Dutch language sufficiently.[17, 18] This FFQ was returned by 3,650 mothers (72%). After exclusion of subjects with invalid dietary data (n=21), information on nutrition was available for 3,629 children.[18] Of these children, growth data, i.e. height, weight and BMI between 13 months and 9 years was available for 3,564 children (**Figure 4.2.1**.).

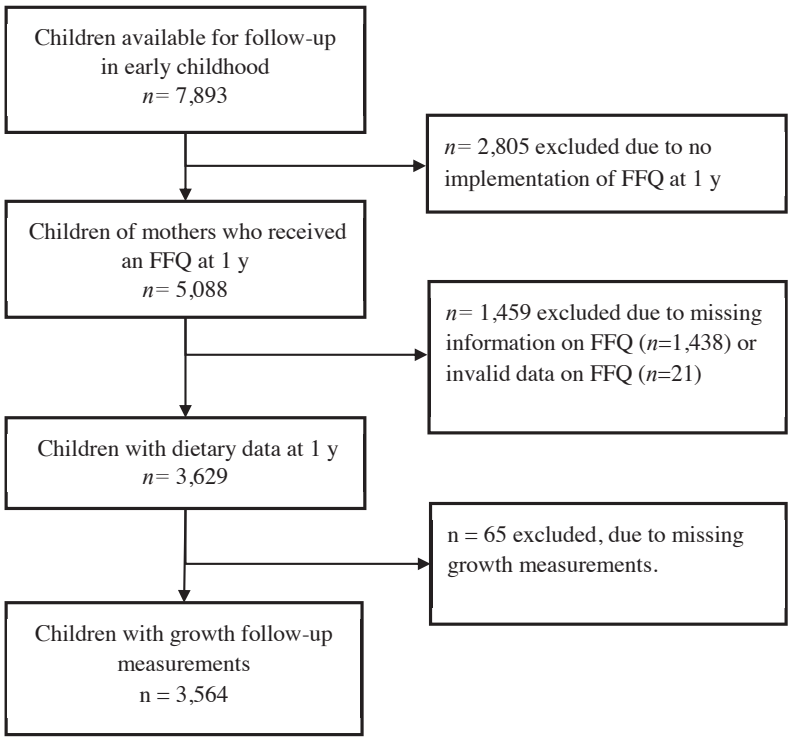


Figure 4.2.1. Flowchart: Population for analysis

Dietary assessment

Dietary intake of the children was assessed at a median age of 12.9 months (95% range 12.2-18.9) using a semi-quantitative FFQ.[17] This FFQ was modified to include only foods that are frequently consumed during the second year of life according to a Dutch national food consumption survey in 2002.[19] The final FFQ consisted of 211 food items and included questions on frequency, quantity, type of foods consumed, and on preparation methods. Food frequencies and quantities were converted into grams per day using standardized portion sizes. Intakes of total protein, animal protein (further divided into dairy and non-dairy), and vegetable protein were calculated using the Dutch Food Composition Table 2006.[20] Amino acid content was calculated from McCance and Widdowson's food composition table, which included chemically analyzed amino acid composition of 150 foods.[21] The FFQ was evaluated against three 24h-recalls, obtained by trained nutritionists, in a representative sample of 32 Dutch children aged 14 months. This evaluation showed intraclass correlation coefficients of 0.4 for total energy, 0.7 for total protein, 0.4 for total fat, 0.4 for carbohydrates, and 0.7 for dietary fiber intake.[17, 18]

Growth and body composition outcomes

Data on height, weight, and BMI was collected at eight different time point between the ages of 13 months and 9 years. At the Child Health Centers, height and weight were measured without shoes

and heavy clothing during routine visits at median ages of 14 (95% range 13-16), 18 (95% range 17-21), 24 (95% range 23-28), 30 (95% range 29-34), 36 (95% range 35-40), and 45 (95% range 44-48) months. At median ages of 5.9 years (95% range 5.7-6.6) and 9.7 years (95% range 9.4-10.8), children visited our research center in the Erasmus Medical Center for a detailed physical examination. Height was determined in standing position to the nearest millimeter by a Harpenden stadiometer (Holtain Limited, Dyfed, U.K.). Weight was measured using a mechanical personal scale (SECA, Almere, the Netherlands). For all time points, body mass index (BMI) was calculated (kg/m^2). Age- and sex-specific standard deviation (SD) scores for height, weight, and BMI were calculated. Overweight was determined using cut-off values for BMI according to the Cole criteria.[22]

Covariates

Information on maternal age, educational level (no/primary, secondary or higher), household income (<1400, 1400-2200 or >2200 euros/month), parity (0, 1 or ≥2), and folic acid supplement use (start preconceptional, start in the first 10 weeks of pregnancy, or never) was obtained through self-administered questionnaires at enrollment. Smoking and alcohol consumption during pregnancy were assessed with questionnaires in each trimester and categorized as never, until pregnancy was known, or continued during pregnancy. At enrollment maternal height and weight were measured during early pregnancy and BMI (kg/m^2) was calculated. Information on child's sex, ethnicity, gestational age at birth, and birth weight were obtained from hospital medical records. Birth weight Z-scores for gestational age were calculated using reference data.[23] Ethnicity was defined as Dutch or non-Dutch, on the basis of the countries of birth of the parents, according to Statistics Netherlands.[24] Information on breastfeeding (exclusively for at least 4 months, partially in the first 4 months, or never) was obtained by postnatal questionnaires. A previously defined child diet score was used to quantify overall diet quality at the age of 1 year,[18] which we previously found to be associated with body composition.[25] The diet score is composed of ten food groups (**Supplementary Table 4.2.1**), for which the ratio of the reported intake and the recommended intake was calculated. For each component the score was truncated at 1. The scores for the single components were added together, resulting in an overall score that ranged from 0 to 10 on a continuous scale, with a higher score representing a healthier diet.[18] Information on child's sports participation (yes/no) and screen time (television watching and computer use, h/d) was obtained using a questionnaire at the age of 6 years.

Statistical analysis

To account for potential measurement error and confounding by total energy intake, we used the residual method to adjust the intake of protein for total energy intake.[26] In short, linear regression analyses were used to calculate energy adjusted intake of protein for each subject, with energy intake as independent variable and protein intake as dependent variable. Similarly, intakes of amino acids and the diet score were adjusted for both energy and total protein intake.

To reduce potential bias due to missing values, incomplete data on covariates were multiple imputed (m=5 imputations).[27] In order to establish covariates to include in our models, we used linear regression analyses to examine the associations of total protein, animal protein, and vegetable protein intake at the age of 1 year with age- and sex-adjusted SD scores of growth measurements (height, weight and BMI) at the age of 9 years, with protein intake as independent variable and growth

outcomes as dependent variables. Covariates were added to the model and were selected based a 10% change in effect estimates and on previous literature. Three multivariable models were used: Model 1 (crude) included child's sex, ethnicity, age at FFQ and energy intake; model 2 (confounders) was additionally adjusted for birthweight Z score, breastfeeding, playing sports, household income, maternal BMI at enrollment, maternal education, folic acid use during pregnancy and smoking during pregnancy; and model 3 (confounders + diet score) was additionally adjusted for the diet score to examine if protein intake was related to growth independent of overall diet quality.

Associations between protein intake and total trajectories of growth between the age of 1 and 9 years were analyzed using linear mixed models. This method allows to use all available repeated measurements of either height, weight or BMI simultaneously and takes into account that these measurements are correlated within subjects. The fixed effects structure was specified according to the previously mentioned models 1, 2 and 3. Likelihood ratio tests were used to determine a suitable random effects structure, random intercepts and slopes, which were used in each of the longitudinal models. To examine whether protein intake modified growth trajectories, i.e. velocity, we included interaction terms between protein intake and age at measurement of the outcome in the fixed effects structure. Furthermore, to assess if associations between protein intake and growth outcomes differed by sex, ethnicity, age at dietary measurement, or birth year we added interaction terms between the potential effect modifier and protein intake to the crude and confounder models. Considering that the FFQ was developed for Dutch children, we performed sensitivity analyses in children with a Dutch ethnic background only. In addition, sensitivity analyses were performed by stratifying for sex, since previous studies suggest that the influence of diet in early life might differ by sex.[28, 29] To test whether associations differed when protein intake was replaced by fat or carbohydrate intake, macronutrient substitution models were performed.

Results were considered statistically significant with a $P < 0.05$. Statistical analyses were conducted using SPSS version 21.0 (IBM Inc., Armonk, NY, USA) and R version 3.1.2 (The R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

Subject characteristics

Characteristics of the children and their mothers are presented in **Table 4.2.1**. The majority of the study population had a Dutch ethnicity (67.7%). Total protein intake at the age of 12.9 months (95% range 12.2-18.9) was 42.1 ± 13.4 g/d. This is higher than the recommended intake for this age group,[30] but comparable to observed in other Western populations.[7, 8, 31]

Mean animal protein intake was 26.5 ± 10.6 g/d, and mean vegetable protein intake was 15.1 ± 5.9 g/d. At the age of 9 years mean height (cm) was 141.6 ± 6.4, mean weight (kg) was 34.7 ± 6.7, and mean BMI (kg/m^2) was 17.2 ± 2.5.

Table 4.2.1. Population characteristics (*n* =3564)[1]

Characteristics	Values[1]
Child characteristics	
Sex, female	1816 (51.0)
Ethnicity	
Dutch	2415 (67.8)
Non Dutch	1149 (32.2)
Gestational age at birth (wk)	40.1 [35.6 - 42.3]
Birth weight (g)	3446 ± 580
Breastfeeding	
Exclusively ≥ 4 months	1005 (28.2)
Partially ≥ 4 months	2263 (63.5)
Never	296 (8.3)
Characteristics at dietary measurement (1y)	
Age filling out FFQ (months)	12.9 [12.2 - 18.9]
BMI (kg/m^2)	17.4 ± 1.4
Total energy intake (kcal)	1321 ± 413
Total protein intake (g/d)	42.1 ± 13.4
Animal protein intake (g/d)	26.5 ± 10.6
Dairy animal protein (g/d)	17.7 ± 8.3
Non-dairy animal protein intake (g/d)	8.8 ± 5.9
Vegetable protein intake (g/d)	15.1 ± 5.9
Amino acids	
Methionine (g/d)	0.68 ± 0.33
Lysine (g/d)	2.04 ± 0.99
Arginine (g/d)	1.64 ± 0.72
Isoleucine (g/d)	1.49 ± 0.67
Leucine (g/d)	2.83 ± 1.12
Valine (g/d)	1.83 ± 0.82
Cysteine (g/d)	0.50 ± 0.19
Phenylalanine (g/d)	1.56 ± 0.67
Tyrosine (g/d)	1.17 ± 0.52
Threonine (g/d)	1.31 ± 0.59
Tryptophan (g/d)	4.12 ± 0.18
Histidine (g/d)	0.93 ± 0.40
Alanine (g/d)	1.45 ± 0.64
Asparagine (g/d)	2.84 ± 1.26
Glutamine (g/d)	6.88 ± 2.85
Glycine (g/d)	1.22 ± 0.52
Proline (g/d)	2.50 ± 1.03

Table 4.2.1. Population characteristics (n =3564)[1] (continued)

Characteristics	Values[1]
Serine (g/d)	1.87 ± 0.74
Diet quality score	4.2 ± 1.3
Characteristics at growth measurements (6y)	
Age (years)	6.0 ± 0.2
Height (cm)	118.2 ± 5.2
Weight (kg)	22.4 ± 3.4
BMI (kg/m^2)	16.0 ± 1.6
Sport participation	1557 (43.7)
Characteristics at growth measurements (9y)	
Age (years)	9.7 ± 0.3
Height (cm)	141.6 ± 6.4
Weight (kg)	34.7 ± 6.7
BMI (kg/m^2)	17.2 ± 2.5
Maternal characteristics	
Maternal age at enrollment (years)	31.4 ± 4.6
Maternal BMI at enrollment (kg/m^2)	23.5 [18.8 - 35.4]
Maternal education	
No or primary education	238 (6.7)
Secondary	1294 (36.3)
Higher	2033 (57.0)
Household income	
<1,400 euro/mo	625 (17.5)
1,400-2,200 euro/mo	686 (19.2)
>2,200 euro/mo	2252 (63.2)
Folic acid supplement use during pregnancy	
None	609 (17.1)
First 10 weeks	1101 (30.9)
Periconceptional	1854 (52.0)
Smoking during pregnancy	
Never	2764 (77.5)
Until pregnancy was known	356 (10.0)
Continued	444 (12.5)
Alcohol consumption during pregnancy	
Never	1417 (39.8)
Until pregnancy was known	561 (15.7)
Continued	1585 (44.5)

[1] *Values represent mean ± SD, Median [95% range], or N(%)*
BMI, body mass index; FFQ, food frequency questionnaire.

Protein intake and growth

The results of the linear mixed models are presented in **Table 4.2.2**. The adjusted models (model 2) show that a 10 g/d higher protein intake was associated with an increase in trajectory of height with 0.03 SDS (95% CI 0.00; 0.06), of weight with 0.06 SDS (95%CI 0.03; 0.09), and of BMI with 0.05 SDS (95%CI 0.03; 0.08) up to the age of 9 years. This association did not attenuate and remained statistically significant after additional adjustment for the diet score (model 3). To illustrate, these results correspond to actual units of growth measurements of approximately 0.2 cm for height, 0.4 kg for weight, and 0.1 kg/m^2 for BMI at the age of 9 years. When analyses were performed for protein from animal and vegetable sources, associations between animal protein and growth outcomes were stronger than for vegetable protein intake (Table 4.2.2). The results of the analysis in which animal protein was further divided into intake from dairy and non-dairy food sources are presented in **Table 4.2.3**. Effect estimates for both dairy protein and non-dairy protein were very similar. **Supplementary Table 4.2.2** shows the results from the linear mixed models for intake of amino acids. Except for

Table 4.2.2. Associations of protein intake at the age of 1 year with childhood growth measures between the age of 1 and 9 years (*n*=3564).[1]

	Height (SDS)	Weight (SDS)	BMI (SDS)
Total protein intake (10 g/d)[2]			
Model 1 (crude)[4]	0.04 (0.01; 0.07)*	0.06 (0.03; 0.09)*	0.06 (0.03; 0.08)*
Model 2 (confounders)[5]	0.03 (0.00; 0.06)*	0.06 (0.03; 0.09)*	0.05 (0.03; 0.08)*
Model 3 (diet score)[6]	0.03 (0.01; 0.06)*	0.06 (0.03; 0.09)*	0.05 (0.03; 0.08)*
Animal protein intake (10 g/d)[2,3]			
Model 1 (crude)[4]	0.03 (0.00; 0.07)*	0.06 (0.03; 0.09)*	0.06 (0.03; 0.09)*
Model 2 (confounders)[5]	0.03 (0.00; 0.06)*	0.06 (0.03; 0.09)*	0.06 (0.03; 0.08)*
Model 3 (diet score)[6]	0.04 (0.01; 0.07)*	0.07 (0.04; 0.10)*	0.06 (0.03; 0.09)*
Vegetable protein intake (10 g/d)[2,3]			
Model 1 (crude)[4]	0.03 (0.00; 0.06)*	0.04 (0.00; 0.07)*	0.02 (-0.01; 0.05)
Model 2 (confounders)[5]	0.02 (-0.01; 0.05)	0.03 (-0.00; 0.05)	0.02 (-0.01; 0.05)
Model 3 (diet score)[6]	0.01 (-0.02; 0.04)	0.01 (-0.02; 0.04)	0.01 (-0.02; 0.04)

[1]Values are regression coefficients and 95% confidence intervals based on linear mixed models and reflect differences in growth measures (age- and sex-specific SD scores) per 10 g/d increase in energy-adjusted protein intake.
[2]Total, animal and vegetable protein intake were adjusted for total energy intake using the residual method.
[3]Models including animal protein were adjusted for vegetable protein and vice versa.
[4]Model 1 is adjusted for child sex , ethnicity, age at dietary measurement and total energy intake at 1 y;
[5]Model 2 is additionally adjusted for birthweight Z score, breastfeeding, playing sports; household income; and maternal BMI at enrollment, maternal education, folic acid use during pregnancy and smoking during pregnancy.
[6]Model 3 is additionally adjusted for diet score
** statistically significant effect estimates.*
BMI, body mass index; SDS, standard deviation score.

Table 4.2.3. Associations of dairy and non-dairy animal protein intake at the age of 1 year with childhood growth measures between the age of 1 and 9 years (*n*= 3564).[1]

	Height (SDS)	Weight (SDS)	BMI (SDS)
Dairy protein intake (10 g/d)[2,3]			
	0.04 (-0.01; 0.08)	0.07 (0.03; 0.12)*	0.07 (0.02; 0.11)*
Non-dairy protein intake (10 g/d)[2,3]			
	0.04 (-0.02; 0.09)	0.07 (0.02; 0.12)*	0.07 (0.02; 0.12)*

[1]Values are regression coefficients and 95% confidence intervals based on linear mixed models and reflect differences in growth measures (age- and sex-specific SD scores) per 10 g/d increase in energy-adjusted protein intake.
[2] Dairy and non-dairy protein intake were adjusted for total energy intake using the residual method.
[3]Models are adjusted for birthweight Z score, breastfeeding, playing sports; household income; and maternal BMI at enrollment, maternal education, folic acid use during pregnancy and smoking during pregnancy. Models including dairy animal protein were adjusted for non-dairy animal protein and vice versa.
** statistically significant effect estimates.*
BMI, body mass index; SDS, standard deviation score.

serine, higher intakes of all of the individual amino acids were significantly associated with a higher weight and BMI. For most amino acids, strengths of the associations were fairly similar, only for cysteine and tryptophan the effect estimates were slightly larger compared to the other amino acids: a 100 mg/d higher cysteine intake was associated with a 0.038 (95% CI 0.016; 0.060) SDS increase in BMI and a 100 mg/d higher intake in tryptophan was associated with a 0.052 (95% CI 0.029; 0.074) SDS increase in BMI. Overall for weight and BMI, associations did not differ substantially between the essential and non-essential amino acids. Higher intakes of all of the essential amino acids were significantly associated with an increase in height. However, some of the non-essential amino acids, including cysteine, arginine, glycine and serine, were not significantly associated with height.

Additional analyses

Interactions of protein intake with age at outcome measurement were not statistically significant. This suggests that protein intake does not affect the velocity height/weight/BMI. Furthermore, interactions of protein intake with sex, ethnicity, age at FFQ, and birth year were not significant. Sensitivity analyses in children with a Dutch ethnic background showed that effect estimates for total and animal protein intake in this group were similar to those in the total population, but had slightly wider confidence intervals (**Supplementary Table 4.2.3**). Food sources of total, animal and vegetable protein intake were similar between Dutch only and the total study population (data not shown). Results were similar between boys and girls, but effect estimates were slightly stronger for girls compared to boys (**Supplementary Tables 4.2.4 and 4.2.5**). Macronutrient substitution models in which protein intake was replaced either with fat or carbohydrate intakes showed similar results (data no shown).

DISCUSSION

In our multi-ethnic population-based study, a higher protein intake at the age of 1 year was associated with a higher height, weight, and BMI up to the age of 9 years. This association was stronger for animal protein compared to vegetable protein, but did not differ between dairy and non-dairy animal protein. Furthermore, associations were not driven by intake of specific amino acids.

One of the suggested mechanisms that may explain the associations we observed is that protein intake stimulates the release of insulin like growth factor (IGF-I) and other growth hormones, which in turn results in rapid growth and increased muscle and fat mass. However, we have previously shown that this increased BMI at age 6 years is fully driven by an increase in fat mass and not fat-free mass.[32]

Our results are in line with a few previous studies, which also showed that a higher total protein intake in early childhood is associated with higher height, weight and BMI in later childhood.[7, 8, 31] In addition, the causal effect of a higher protein intake in early childhood on a higher weight and BMI has been confirmed in a randomized controlled trial among 1138 children comparing the effect of higher- and lower-protein formula during the first year of life on growth up to 6 years of age.[9, 10] However, in contrast to our results, no difference in height was observed between the lower- and higher-protein formula groups in this trial. In our study we found a significant increase in height with higher protein intake, although effect estimates were not as strong as those for weight and BMI. A possible explanation might be that height is more genetically determined and is less prone to modification by external exposures such as diet.[33]

Our results suggest that the associations between protein intake and growth might be stronger for intake of animal protein rather than vegetable protein. This corresponds with findings from two previous studies which found that a higher animal protein intake in early life was associated with higher BMI later in childhood.[11, 34] These different results for animal and vegetable protein intake might be due to the higher intake of animal protein (8.0 E%) compared to vegetable protein (4.6 E%) in our population and in previous studies.[11, 34] Nevertheless, it has also been suggested that protein from animal sources specifically stimulate growth. These effects of animal protein intake have been ascribed to protein from dairy sources in particular, which has been suggested to have specific growth stimulating effects by raising hormonal responses.[12, 35] However, when we further separated animal protein into intake from dairy and non-dairy animal food sources, effect sizes for both were similar. These results indicate that in our population the associations observed between animal protein intake and growth were not specifically driven by dairy or non-dairy animal protein.

Another possible explanation for the difference in results regarding animal versus vegetable protein might be due to the high proportion of essential amino acids present in animal protein. Axelsson et al. suggested that certain specific amino acids (i.e., arginine, lysine, threonine, valine, leucine, isoleucine, phenylalanine) might stimulate secretion of insulin and growth hormone, which might mediate the association between protein and growth.[36] Furthermore, Vught et al. reported that higher arginine intake was associated with growth velocity in children aged 7-13 years after 3 years of follow up,[14] and that a high arginine intake in combination with a high lysine intake was associated with a higher fat mass index among 8-10 year old girls after 6 years of follow up.[37] However, our results suggested that

associations did not differ substantially between essential and non-essential amino acids and were not driven by any of the individual amino acids.

The current study has several strengths and limitations. Important strengths of this study are the large sample size of the population included, the prospective, population-based study design, and the repeated growth measurements obtained by professional staff. Another strength is the large amount of information available on potential sociodemographic and lifestyle confounders of the children and their parents. We aimed to reduce attrition bias as much as possible. Therefore, we used a multiple imputation procedure, which is an appropriate method to deal with missing data because it requires the least assumptions and reduces bias when the missing data are not completely at random.[27] A limitation of this study is the observational nature of the study design, therefore, no causal relation can be concluded. As in all observational studies, residual confounding may still be an issue, for example by physical activity and pubertal status. The FFQ was only sent to mothers who mastered the Dutch language sufficiently, of which 72% filled in and returned the questionnaire. This may have led to selection bias. As we previously reported, children without FFQ data had a higher height, weight, and BMI compared to the children with FFQ data.[38] However, this difference would only result in selection bias if associations between protein intake and growth differs between the included and excluded participants. Another limitation of this study is that the FFQ was developed for Dutch children, whereas our study was conducted in a multi-ethnic study population. However, when we performed sensitivity analyses restricted to Dutch children only, results were similar to those in the whole group. Furthermore, the FFQ was evaluated for total protein, showing a relatively high intraclass correlation coefficient (0.7), but protein from specific sources or amino acids were not evaluated. Not all amino acids could be calculated from the FFQ, 18 amino acids were available for analysis. Moreover, with the calculation of the amount of amino acids, intakes from breast milk were not taken into account because no amino acid composition data was available. However, the number of children still being breast fed at the age of 1 year was relatively small (10%), therefore the error due to this missing information is probably minimal. Unfortunately, we had no repeated dietary data or dietary measurements after the age of 1 year. In addition, data on IGF-I or other growth factors that might mediate the association between protein intake and growth were not available in this study.

CONCLUSION

Results of this prospective cohort study support findings from previous studies that a higher intake of protein during early childhood is associated with a higher height, weight, and BMI in later childhood. Furthermore, our results add to the literature that in particular animal protein is associated with growth, but that this is not specifically driven by dairy protein or specific amino acids. Future studies should explore the role of growth hormones, and investigate if protein intake in early childhood affects health later in life.

SUPPLEMENTARY MATERIAL

Supplementary Table 4.2.1. Diet quality score.

Food group	Cut-off level	Summary of included items	Summary of excluded items
Vegetables	≥100 g	Fresh vegetables, frozen or canned vegetables	Pickles
Fruit	≥150 g	Fresh fruit, canned fruit without added sugar	Canned fruit with added sugar, fruit juice
Bread and cereals	≥70 g	Whole-wheat bread or crackers, oatmeal, muesli without added sugar	White bread or crackers, breakfast cereals with added sugar
Rice, pasta, potatoes, and legumes	≥70 g	Boiled or steamed potatoes, whole-wheat pasta, couscous, whole-grain rice, legumes	Fried potatoes, French fries, white pasta, white rice
Dairy	≥370 g	Whole, semi-skimmed and skimmed milk and yogurt without added sugars, soy milk without added sugars, low-fat and reduced-fat cheeses (20+ and 30+)	Milk products with added sugars, full-fat cheeses
Meat, eggs and meat substitutes	≥35 g	Low-fat meat, eggs, tofu, tempeh	Fat and processed meat
Fish	≥15 g	Fresh or canned fish	Fish fingers
Oils and fats	≥25 g	Low-fat margarine products(≤16g saturated fat and ≤1g trans fat per 100g), oils, liquid cooking or frying fat	Butter, solid cooking or frying fats
Candy and snacks	≤ 20 g/d	Ice cream, potato chips, cookies, candy bars, fried snacks, cakes	-
Sugar-sweetened beverages	≤ 100 g/d	Soft drinks, lemonade	-

We used a food-based diet score to overall diet quality in the children. The score is composed of eight food groups. For each food group, we calculated the ratio of the reported intake and the recommended intake (see table below). For each component the score was truncated at 1. The scores for the single components were added together, resulting in an overall score that ranged from 0 to 8 on a continuous scale, with a higher score representing a healthier diet. The score was adjusted for energy intake using the residuals method.

Supplementary Table 4.2.2. Associations of amino acid intake at the age of 1 year with childhood growth measures between the age of 1 and 9 years (*n*=3564).

	Height (SDS)	Weight (SDS)	BMI (SDS)
Essential amino acids			
Methionine (100 mg/d)	0.013 (0.002; 0.023)*	0.024 (0.015; 0.034)*	0.023 (0.014; 0.032)*
Lysine (100/mg)	0.006 (0.001; 0.010)*	0.010 (0.006; 0.014)*	0.009 (0.005; 0.013)*
Valine (100 mg/d)	0.006 (0.002; 0.010)*	0.010 (0.006; 0.014)*	0.009 (0.006; 0.013)*
Leucine (100 mg/d)	0.006 (0.002; 0.010)*	0.010 (0.006; 0.014)*	0.009 (0.005; 0.013)*
Isoleucine (100 mg/d)	0.008 (0.002; 0.015)*	0.015 (0.009; 0.021)*	0.014 (0.008; 0.019)*
Phenylalanine (100 mg/d)	0.008 (0.002; 0.014)*	0.014 (0.008; 0.020)*	0.013 (0.007; 0.018)*
Threonine (100 mg/d)	0.010 (0.002; 0.017)*	0.017 (0.010; 0.024)*	0.016 (0.009; 0.022)*
Tryptophan (100 mg/d)	0.031 (0.006; 0.056)*	0.057 (0.033; 0.080)*	0.052 (0.029; 0.074)*
Histidine (100 mg/d)	0.014 (0.003; 0.025)*	0.024 (0.014; 0.034)*	0.021 (0.011; 0.031)*
Non-essential amino acids			
Cysteine (100 mg/d)	0.020 (-0.004; 0.044)	0.040 (0.017; 0.063)*	0.038 (0.016; 0.060)*
Tyrosine (100 mg/d)	0.011 (0.003; 0.019)*	0.019 (0.012; 0.027)*	0.017 (0.010; 0.025)*
Alanine (100 mg/d)	0.007 (0.001; 0.014)*	0.014 (0.008; 0.021)*	0.013 (0.007; 0.019)*
Arginine (100 mg/d)	0.005 (-0.000; 0.009)	0.009 (0.004; 0.014)*	0.009 (0.004; 0.013)*
Asparagine (100 mg/d)	0.004 (0.000; 0.007)*	0.007 (0.004; 0.010)*	0.006 (0.003; 0.010)*
Glutamine (100 mg/d)	0.002 (0.000; 0.003)*	0.003 (0.002; 0.005)*	0.003 (0.001; 0.004)*
Glycine (100 mg/d))	0.005 (-0.002; 0.015)	0.013 (0.005; 0.021)*	0.013 (0.005; 0.021)*
Proline (100 mg/d)	0.006 (0.002; 0.009)*	0.009 (0.005; 0.013)*	0.008 (0.004; 0.011)*
Serine (100 mg/d)	0.002 (-0.003; 0.007)	0.003 (-0.002; 0.008)	0.003 (-0.002; 0.007)

[1] Values are regression coefficients and 95% confidence intervals based on linear mixed models and reflect differences in growth measures (age- and sex-specific SD scores) per 100 mg/d increase in intake.
[2] Amino acids were adjusted for energy and protein using the residual method.
[3]Models are adjusted for birthweight Z score, breastfeeding, playing sports; household income; and maternal BMI at enrollment, maternal education, folic acid use during pregnancy and smoking during pregnancy.
** statistically significant effect estimates.*
BMI, body mass index; SDS, standard deviation score.

Supplementary Table 4.2.3. Associations of protein intake at the age of 1 year with childhood growth measures between the age of 1 and 9 years in Dutch children only (*n*=2433).[1]

	Height (SDS)	Weight (SDS)	BMI (SDS)
Total protein intake* (10 g/d)[2]			
Model 1 (crude)[4]	0.03 (-0.01; 0.07)	0.07 (0.03; 0.10)*	0.07 (0.04; 0.10)*
Model 2 (confounders)[5]	0.03 (-0.00; 0.07)	0.07 (0.04; 0.10)*	0.07 (0.04; 0.10)*
Model 3 (diet score)[6]	0.03 (-0.03; 0.07)	0.07 (0.04; 0.10)*	0.07 (0.04; 0.10)*
Animal protein intake* (10 g/d)[2,3]			
Model 1 (crude)[4]	0.03 (-0.01; 0.07)	0.06 (0.02; 0.10)*	0.06 (0.03; 0.10)*
Model 2 (confounders)[5]	0.04 (-0.00; 0.07)	0.07 (0.03; 0.10)*	0.07 (0.03; 0.10)*
Model 3 (diet score)[6]	0.04 (0.00; 0.07)*	0.07 (0.04; 0.10)*	0.06 (0.03; 0.10)*
Vegetable protein intake* (10 g/d)[2,3]			
Model 1 (crude)[4]	0.02 (-0.02; 0.07)	0.06 (0.02; 0.10)*	0.07 (0.03; 0.10)*
Model 2 (confounders)[5]	-0.01 (-0.03; 0.05)	0.05 (0.01; 0.08)*	0.06 (0.02; 0.09)*
Model 3 (diet score)[6]	0.00 (-0.04; 0.04)	0.04 (0.00; 0.07)*	0.05 (0.02; 0.09)*

[1]Values are regression coefficients and 95% confidence intervals based on linear mixed models and reflect differences in growth measures (age- and sex-specific SD scores) per 10 g/d increase in energy-adjusted protein intake.
[2]Total, animal and vegetable protein intake were adjusted for total energy intake using the residual method.
[3]Models including animal protein were adjusted for vegetable protein and vice versa.
[4]Model 1 is adjusted for child sex , ethnicity, age at dietary measurement and total energy intake at 1 y;
[5]Model 2 is additionally adjusted for birthweight Z score, breastfeeding, playing sports; household income; and maternal BMI at enrollment, maternal education, folic acid use during pregnancy and smoking during pregnancy.
[6]Model 3 is additionally adjusted for diet score
** statistically significant effect estimates.*
BMI, body mass index; SDS, standard deviation score.

Supplementary Table 4.2.4. Associations of protein intake at the age of 1 year with childhood growth measures between the age of 1 and 9 years in girls (*n*=1816). [1]

	Height (SDS)	Weight (SDS)	BMI (SDS)
Total protein intake* (10 g/d)[2]			
Model 1 (crude)[4]	0.05 (0.00; 0.09)*	0.07 (0.03; 0.11)*	0.06 (0.02; 0.10)*
Model 2 (confounders)[5]	0.04 (-0.00; 0.08)	0.07 (0.02; 0.11)*	0.06 (0.02; 0.09)*
Model 3 (diet score)[6]	0.04 (-0.00; 0.08)	0.07 (0.03; 0.10)*	0.06 (0.02; 0.09)*
Animal protein intake* (10 g/d)[2,3]			
Model 1 (crude)[4]	0.05 (0.00; 0.10)*	0.08 (0.04; 0.13)*	0.07 (0.03; 0.11)*
Model 2 (confounders)[5]	0.05 (0.01; 0.09)*	0.08 (0.04; 0.12)*	0.07 (0.03; 0.11)*
Model 3 (diet score)[6]	0.05 (0.01; 0.10)*	0.08 (0.04; 0.13)*	0.07 (0.03; 0.11)*
Vegetable protein intake* (10 g/d)[2,3]			
Model 1 (crude)[4]	0.04 (-0.01; 0.08)	0.04 (-0.00; 0.08)	0.02 (-0.02; 0.06)
Model 2 (confounders)[5]	0.02 (-0.02; 0.06)	0.03 (-0.01; 0.07)	0.02 (-0.02; 0.06)
Model 3 (diet score)[6]	0.02 (-0.03; 0.06)	0.01 (-0.03; 0.05)	0.01 (-0.03; 0.04)

[1]Values are regression coefficients and 95% confidence intervals based on linear mixed models and reflect differences in growth measures (age- and sex-specific SD scores) per 10 g/d increase in energy-adjusted protein intake.
[2]Total, animal and vegetable protein intake were adjusted for total energy intake using the residual method.
[3]Models including animal protein were adjusted for vegetable protein and vice versa.
[4]Model 1 is adjusted for child sex , ethnicity, age at dietary measurement and total energy intake at 1 y;
[5]Model 2 is additionally adjusted for birthweight Z score, breastfeeding, playing sports; household income; and maternal BMI at enrollment, maternal education, folic acid use during pregnancy and smoking during pregnancy.
[6]Model 3 is additionally adjusted for diet score
** statistically significant effect estimates.*
BMI, body mass index; SDS, standard deviation score.

Supplementary Table 4.2.5. Associations of protein intake at the age of 1 year with childhood growth measures between the age of 1 and 9 years in boys (*n*=1748).[1]

	Height (SDS)	Weight (SDS)	BMI (SDS)
Total protein intake* (10 g/d)[2]			
Model 1 (crude)[4]	0.03 (-0.02; 0.07)	0.05 (0.01; 0.10)*	0.05 (0.01; 0.09)*
Model 2 (confounders)[5]	0.02 (-0.02; 0.07)	0.05 (0.01; 0.09)*	0.05 (0.01; 0.09)*
Model 3 (diet score)[6]	0.03 (-0.02; 0.07)	0.05 (0.01; 0.09)*	0.05 (0.01; 0.09)*
Animal protein intake* (10 g/d)[2,3]			
Model 1 (crude)[4]	0.02 (-0.02; 0.06)	0.04 (-0.00; 0.09)	0.04 (0.00; 0.08)*
Model 2 (confounders)[5]	0.02 (-0.02; 0.06)	0.04 (0.00; 0.08)*	0.04 (0.00; 0.08)*
Model 3 (diet score)[6]	0.02 (-0.02; 0.06)	0.05 (0.01; 0.09)*	0.04 (0.01; 0.08)*
Vegetable protein intake* (10 g/d)[2,3]			
Model 1 (crude)[4]	0.02 (-0.02; 0.07)	0.04 (-0.01; 0.08)	0.02 (-0.02; 0.07)
Model 2 (confounders)[5]	0.02 (-0.03; 0.06)	0.03 (-0.02; 0.07)	0.02 (-0.02; 0.06)
Model 3 (diet score)[6]	0.01 (-0.04; 0.06)	0.02 (-0.03; 0.06)	0.02 (-0.03; 0.06)

[1]Values are regression coefficients and 95% confidence intervals based on linear mixed models and reflect differences in growth measures (age- and sex-specific SD scores) per 10 g/d increase in energy-adjusted protein intake.
[2]Total, animal and vegetable protein intake were adjusted for total energy intake using the residual method.
[3]Models including animal protein were adjusted for vegetable protein and vice versa.
[4]Model 1 is adjusted for child sex , ethnicity, age at dietary measurement and total energy intake at 1 y;
[5]Model 2 is additionally adjusted for birthweight Z score, breastfeeding, playing sports; household income; and maternal BMI at enrollment, maternal education, folic acid use during pregnancy and smoking during pregnancy.
[6]Model 3 is additionally adjusted for diet score
** statistically significant effect estimates.*
BMI, body mass index; SDS, standard deviation score.

REFERENCES

1. Walker SP, Wachs TD, Grantham-McGregor S, Black MM, Nelson CA, Huffman SL, et al. Inequality in early childhood: risk and protective factors for early child development. The Lancet. 2011;378(9799): 1325-38.
2. Michaelsen KF, Greer FR. Protein needs early in life and long-term health. Am J Clin Nutr. 2014;99(3): 718S-22S.
3. Joint WHO. Protein and amino acid requirements in human nutrition. World health organization technical report series. 2007(935):1.
4. Allen LH. Nutritional influences on linear growth: a general review. Eur J Clin Nutr. 1994;48 Suppl 1: S75-89.
5. Rolland-Cachera MF, Deheeger M, Akrout M, Bellisle F. Influence of macronutrients on adiposity development: A follow up study of nutrition and growth from 10 months to 8 years of age. Int J Obes. 1995;19(8):573-8.
6. Scaglioni S, Agostoni C, De Notaris R, Radaelli G, Radice N, Valenti M, et al. Early macronutrient intake and overweight at five years of age. Int J Obes. 2000;24(6):777-81.
7. Hoppe C, Molgaard C, Thomsen BL, Juul A, Michaelsen KF. Protein intake at 9 mo of age is associated with body size but not with body fat in 10-y-old Danish children. Am J Clin Nutr. 2004;79(3):494-501.
8. Skinner JD, Bounds W, Carruth BR, Morris M, Ziegler P. Predictors of children's body mass index: a longitudinal study of diet and growth in children aged 2-8 y. Int J Obes Relat Metab Disord. 2004;28(4): 476-82.
9. Koletzko B, von Kries R, Closa R, Escribano J, Scaglioni S, Giovannini M, et al. Lower protein in infant formula is associated with lower weight up to age 2 y: a randomized clinical trial. Am J Clin Nutr. 2009; 89(6):1836-45.
10. Weber M, Grote V, Closa-Monasterolo R, Escribano J, Langhendries JP, Dain E, et al. Lower protein content in infant formula reduces BMI and obesity risk at school age: follow-up of a randomized trial. Am J Clin Nutr. 2014;99(5):1041-51.
11. Thorisdottir B, Gunnarsdottir I, Palsson GI, Halldorsson TI, Thorsdottir I. Animal protein intake at 12 months is associated with growth factors at the age of six. Acta Paediatr. 2014;103(5):512-7.
12. Gunther AL, Remer T, Kroke A, Buyken AE. Early protein intake and later obesity risk: which protein sources at which time points throughout infancy and childhood are important for body mass index and body fat percentage at 7 y of age? Am J Clin Nutr. 2007;86(6):1765-72.
13. Knopf RF, Conn JW, Fajans SS, Floyd JC, Guntsche E, Rull JA. Plasma growth hormone response to intravenous administration of amino acids. The Journal of Clinical Endocrinology & Metabolism. 1965; 25(8):1140-4.
14. van Vught AJAH, Dagnelie PC, Arts ICW, Froberg K, Andersen LB, El-Naaman B, et al. Dietary arginine and linear growth: the Copenhagen School Child Intervention Study. British Journal of Nutrition. 2013;109(06):1031-9.
15. Alles MS, Eussen SRBM, van der Beek EM. Nutritional challenges and opportunities during the weaning period and in young childhood. Annals of Nutrition and Metabolism. 2014;64(3-4):284-93.
16. Kruithof CJ, Kooijman MN, van Duijn CM, Franco OH, de Jongste JC, Klaver CCW, et al. The generation R study: Biobank update 2015. European journal of epidemiology. 2014;29(12):911-27.
17. Kiefte-de Jong JC, de Vries JH, Bleeker SE, Jaddoe VWV, Hofman A, Raat H, et al. Socio-demographic and lifestyle determinants of 'Western-like'and 'Health conscious' dietary patterns in toddlers. British Journal of Nutrition. 2013;109(01):137-47.

18. Voortman T, Kiefte-de Jong JC, Geelen A, Villamor E, Moll HA, de Jongste JC, et al. The Development of a Diet Quality Score for Preschool Children and Its Validation and Determinants in the Generation R Study. The Journal of nutrition. 2015:jn. 114.199349.
19. Hulshof K, Breedveld B. Results of the study on nutrient intake in young toddlers 2002. TNO Nutrition: Zeist, The Netherlands. 2002.
20. Center NN. Nevo: Dutch food composition database 2006. The Hague, The Netherlands: Netherlands Nutrition Centre. 2006.
21. Paul AA, Southgate DAT, Russell J. First supplement to McCance and Widdowson's The Composition of Foods1980.
22. Cole TJ, Bellizzi MC, Flegal KM, Dietz WH. Establishing a standard definition for child overweight and obesity worldwide: international survey. Bmj. 2000;320(7244):1240.
23. Niklasson A, Ericson A, Fryer JG, Karlberg J, Lawrence C, Karlberg P. An Update of the Swedish Reference Standards for Weight, Length and Head Circumference at Birth for Given Gestational Age (1977-1981). Acta Paediatrica. 1991;80(8-9):756-62.
24. Raza JA, Babb JD, Movahed A. Optimal management of hyperlipidemia in primary prevention of cardiovascular disease. International journal of cardiology. 2004;97(3):355-66.
25. Voortman T, Leermakers ET, Franco OH, Jaddoe VW, Moll HA, Hofman A, et al. A priori and a posteriori dietary patterns at the age of 1 year and body composition at the age of 6 years: the Generation R Study. Eur J Epidemiol. 2016.
26. Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. The American journal of clinical nutrition. 1997;65(4):1220S-8S.
27. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ: British Medical Journal. 2009;338.
28. Leermakers ETM, Felix JF, Erler NS, Ćerimagić A, Wijtzes AI, Hofman A, et al. Sugar-containing beverage intake in toddlers and body composition up to age 6 years: The Generation R Study. European journal of clinical nutrition. 2015.
29. Closa-Monasterolo R, Ferré N, Luque V, Zaragoza-Jordana M, Grote V, Weber M, et al. Sex differences in the endocrine system in response to protein intake early in life. The American journal of clinical nutrition. 2011;94(6 Suppl):1920S-7S.
30. Trumbo P, Schlicker S, Yates AA, Poos M. Dietary reference intakes for energy, carbohydrate, fiber, fat, fatty acids, cholesterol, protein and amino acids. Journal of the American Dietetic Association. 2002; 102(11):1621-30.
31. Thorisdottir B, Gunnarsdottir I, Thorisdottir AV, Palsson GI, Halldorsson TI, Thorsdottir I. Nutrient intake in infancy and body mass index at six years in two population-based cohorts recruited before and after revision of infant dietary recommendations. Ann Nutr Metab. 2013;63(1-2):145-51.
32. Voortman T, Braun KV, Kiefte-de Jong JC, Jaddoe VW, Franco OH, van den Hooven EH. Protein intake in early childhood and body composition at the age of 6 years: The Generation R Study. Int J Obes (Lond). 2016;40(6):1018-25.
33. Wood AR, Esko T, Yang J, Vedantam S, Pers TH, Gustafsson S, et al. Defining the role of common variation in the genomic and biological architecture of adult human height. Nature genetics. 2014;46(11): 1173-86.
34. Weijs PJ, Kool LM, van Baar NM, van der Zee SC. High beverage sugar as well as high animal protein intake at infancy may increase overweight risk at 8 years: a prospective longitudinal pilot study. Nutr J. 2011;10:95.

35. Orr JB. Influence of Amount of Milk Consumption on the Rate of Growth of School Children. Br Med J. 1928;1(3499):140-1.
36. Axelsson I, Borulf S, Righard L, Räihä N. Protein and energy intake during weaning: I. Effects on growth. Acta Paediatrica. 1987;76(2):321-7.
37. van Vught AJ, Heitmann BL, Nieuwenhuizen AG, Veldhorst MA, Andersen LB, Hasselstrom H, et al. Association between intake of dietary protein and 3-year-change in body growth among normal and overweight 6-year-old boys and girls (CoSCIS). Public Health Nutr. 2010;13(5):647-53.
38. Braun KV, Voortman T, Kiefte-de Jong JC, Jaddoe VW, Hofman A, Franco OH, et al. Dietary Intakes of Folic Acid and Methionine in Early Childhood Are Associated with Body Composition at School Age. J Nutr. 2015;145(9):2123-9.

4.3

Protein intake in early childhood and body composition at the age of 6 years: The Generation R Study.

Based on:

Voortman T, **Braun KVE**, Kiefte-de Jong JC, Jaddoe VW, Franco OH, van den Hooven EH. Protein intake in early childhood and body composition at the age of 6 years: The Generation R Study. *International Journal of Obesity (Lond).* 2016;40(6):1018-25.

ABSTRACT

Background: Previous studies suggest that high protein intake in infancy leads to a higher body mass index in later childhood. We examined the associations of total, animal and vegetable protein intake in early childhood with detailed measures of body composition at the age of 6 years.

Methods: This study was performed in 2,911 children participating in a population-based cohort study. Protein intake at the age of 1 year was assessed with a validated food-frequency questionnaire and was adjusted for total energy intake. At the children's age of 6 years, we measured their anthropometrics and body fat (with dual-energy X-ray absorptiometry). We calculated age- and sex- specific SD scores for body mass index (BMI), fat mass index (FMI), and fat-free mass index (FFMI).

Results: After adjustment for confounders, a 10 g/d higher total protein intake at 1 year of age was associated with a 0.05 SD (95%CI 0.00, 0.09) higher BMI at age 6 years. This association was fully driven by a higher FMI (0.06 SD (95%CI 0.01, 0.11)) and not FFMI (-0.01 SD (95%CI -0.06, 0.05)). The associations of protein intake with FMI at 6 years remained significant after adjustment for BMI at the age of 1 year. Additional analyses showed that the associations of protein intake with FMI were stronger in girls than in boys (p for interaction=0.03), stronger among children who had catch-up growth in the first year of life (p for interaction<0.01), and stronger for intake of animal protein (both dairy and non-dairy protein) than protein from vegetable sources.

Conclusions: Our results suggest that high protein intake in early childhood is associated with higher body fat mass, but not fat-free mass. Future studies are needed to investigate whether these changes persist into adulthood and to examine the optimal range of protein intake for infants and young children.

INTRODUCTION

Protein is an important component of early life nutrition, as it provides essential amino acids required for growth. However, a high protein intake in early childhood has been linked to a higher risk of obesity.[1-3] Whereas in adults a higher total protein intake increases satiety and energy expenditure and is thereby beneficial for body weight maintenance,[4, 5] different mechanisms may play a role in infancy. The early protein hypothesis states that high protein in early childhood is linked to later obesity through hormonal responses,[1] including enhanced secretion of insulin-like growth factor-1 (IGF-1),[6] which may increase growth and adipogenesis.[7] This effect might have a specific risk window around the age of 1 year, when most children undergo a transition from milk-based and complementary feeding to table foods, often coinciding with a sharp increase in protein intake.[8] Furthermore, previous studies suggested that certain groups of children may be extra sensitive to potential adverse effect of high protein intake in early life, for example those who are genetically predisposed to obesity or children who experienced catch-up growth in early childhood.[2, 9]

Although several studies showed that high protein intake in early childhood is associated with a higher body mass index (BMI) in later childhood,[1, 2, 10] most of these studies did not distinguish between body fat and lean mass. Because variation in fat-free mass is an important source of variability in BMI in infants and children, BMI is considered to be of limited use to measure childhood adiposity.[11] Therefore, we studied the associations between protein intake at the age of 1 year and BMI, fat mass index, and fat-free mass index at the age of 6 years. Furthermore, we evaluated these associations in a subgroup of the children for protein intake at the age of 2 years; and we examined whether the associations differed by children's sex, ethnicity, genetic risk score for BMI, birth weight, or catch-up growth or by protein source.

SUBJECTS AND METHODS

Subjects and study design

This study was embedded in the Generation R Study, a population-based prospective cohort study from fetal life onward in Rotterdam, the Netherlands.[12] Pregnant women were enrolled between 2001 and 2005 and 7,893 children were available for follow-up studies in early childhood. A questionnaire on infant diet was introduced from 2003 onwards, and was sent to 5,088 mothers who provided consent for follow-up and had sufficient mastery of the Dutch language (**Figure 4.3.1**). Complete dietary data was available for 3,629 children,[13, 14] of whom 2,911 visited the research center at the age of 6 years for body composition measurements (**Figure 4.3.1**). The study was conducted according to the Declaration of Helsinki, approved by the Medical Ethics Committee of Erasmus Medical Center, Rotterdam, and parents of the children provided written informed consent.

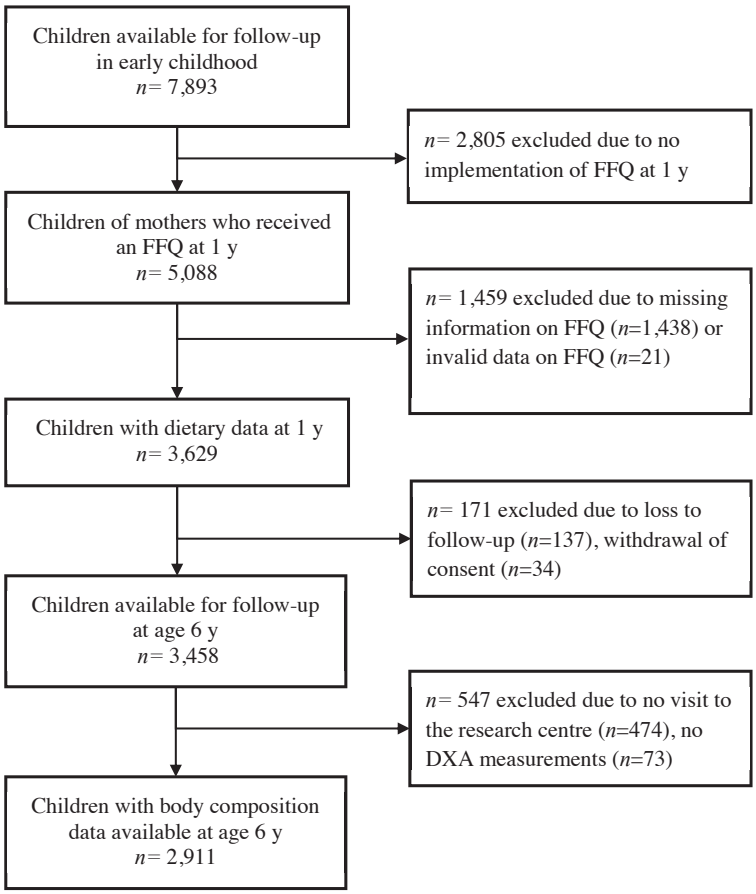


Figure 4.3.1. Flowchart: Population for analysis

Dietary assessment

Dietary intake was assessed at a median age of 12.9 months (95% range 12.2-18.9), with a semi-quantitative food-frequency questionnaire (FFQ) consisting of 211 food items and covering the past month.[13] On the basis of standardized portion sizes and the Dutch Food Composition Table 2006,[15] food frequencies were converted into energy and macronutrient intakes. Protein intake was further divided into protein from animal food sources (e.g., meat, fish, eggs, and dairy) and from plant food sources (e.g., grains, legumes, and soy). The FFQ was evaluated against three 24h-recalls in a representative sample of 32 Dutch children with a median age of 14 months (95% range 6 to 20) living in Rotterdam.[13, 14] To evaluate the agreement of the FFQ relative to these 24h-recalls, an intraclass correlation coefficient was computed, which was 0.74 for total protein intake.[13, 14] An additional FFQ around the age of 2 years was implemented in a subgroup of 899 Dutch children.[14] This questionnaire was completed for 844 children (94%) at a median age of 24.9 months (95% range 24.3-27.6),[14] of

whom 698 had body composition measurements taken at the age of 6 years. In total, 649 children had dietary data at both 1 and 2 years and body composition at 6 years available.

Body composition assessments

Children's anthropometrics and body composition were measured at a median age of 5.9 years (95% range 5.6 to 6.6) by well-trained staff in our research center.[12] Height was determined with a Harpenden stadiometer (Holtain Limited, Dyfed, U.K.) and weight was measured using a mechanical personal scale (SECA, Almere, the Netherlands). Total and regional body fat, lean, and bone mass were measured with a dual-energy X-ray absorptiometry (DXA) scanner (iDXA, GE-Lunar, 2008, Madison, WI, USA), using enCORE software version 13.6.

We calculated body mass index (BMI) [weight (kg) /height $(m)^2$], fat mass index (FMI) [fat mass (kg) /height $(m)^2$], and fat-free mass index (FFMI) [fat-free mass (kg) /height $(m)^2$].[16] As secondary outcomes we calculated body fat percentage (BF%) [fat mass (kg) /weight (kg)], and android/gynoid fat ratio (A/G ratio) [android fat mass (kg) /gynoid fat mass (kg)]. For all outcomes we calculated age- and sex-specific SD-scores (SDS) on the basis of the total Generation R study population with body composition data at 6 years (n=6,491).[12]

Covariates

We collected information on household income, maternal age, maternal educational level, and folic acid supplement use (as a proxy for maternal health-conscious behavior) with questionnaires at enrollment in the study.[12] Maternal smoking and alcohol consumption during pregnancy were assessed using questionnaires in each trimester and both variables were categorized into never; until pregnancy was known; or continued during pregnancy. Maternal anthropometrics were measured at enrollment in the study at the research center and BMI was calculated.[12]

Information on child's sex, gestational age, and birth weight was available from medical records. For birth weight, we calculated sex- and gestational age-specific SD-scores.[17] Child's ethnicity (Dutch or non-Dutch) was defined as Dutch if both parents were born in the Netherlands.[18] For a subgroup of 1,909 children, we assessed their genetic predisposition to obesity because this may be an effect modifier in the association of protein intake with obesity.[19] As described in detail previously,[20] cord blood samples were genotyped using Illumina Infinium II HumanHap610 Quad Arrays and we calculated a weighted genetic risk score for BMI using 29 independent variants.[20, 21]

We obtained information on breastfeeding in the first 4 months (never, partial, or exclusive) from delivery reports and from questionnaires. Timing of introduction of solids in the first year of life was obtained from a questionnaire administered at the child's age of 1 year. Children's carbohydrate and fat intake were measured using the previously mentioned FFQs. A diet score for preschool children was used as measure of overall diet quality and was calculated using data obtained with the FFQs.[14] Information on doctor-diagnosed food allergies was obtained with questionnaires at the child's ages of 6 and 12 months. Children's height and weight were measured at several time points between birth and the age of 4 years during routine visits at community health centres,[12] and sex- and age-specific height, weight, and BMI SD-scores were calculated.[22] For the current analyses, we used BMI at the age of 12 months and catch-up growth in the first year of life, defined as a weight change in SD scores greater than 0.67 between the age of 1 and 12 months.[23] Serum insulin levels at age 6 years, which we

examined as possible mediator,[24] were available for 1,966 children. Non-blood samples were obtained at the research center and serum concentrations of insulin were measured with enzymatic methods (Cobas 8000, Roche, Almere, the Netherlands).[25] Screen time (time spent watching television or using a computer), as a proxy for sedentary behavior, and participation in sports (yes/no), as a proxy for physical activity, were assessed with a questionnaire at 6 years.

Statistical analysis

Because we were interested in the effect of protein independent of its energy content, we adjusted dietary intake of protein and the other macronutrients for total energy intake using the nutrient residual method.[26] For interpretability, predicted protein intake at mean energy intake of our study population (1310 kcal/d) was added to the residuals as a constant.

Using natural cubic splines with three knots, we found no indication for non-linear associations between protein intake at 1 year and body composition outcomes at 6 years ($p>0.1$). Therefore, we used linear regression models to examine the associations of total, animal, and vegetable protein intake with body composition outcomes (BMI, FMI, and FFMI). Because a previous study suggested specific associations for dairy protein,[8] we performed *post hoc* analyses in which we further separated animal protein into energy-adjusted protein from dairy versus non-dairy sources. Basic models included energy-adjusted total protein intake, child's sex, total energy intake, and age at outcome measurement (model 1). For analyses with protein from different sources, total protein intake was replaced by energy-adjusted animal protein intake and energy-adjusted vegetable protein, i.e., animal and vegetable protein were adjusted for each other. Multivariable models were further adjusted for household income; maternal age, education, BMI, and smoking during pregnancy; and for child's ethnicity, birth weight SD-score, breastfeeding, total fat intake, diet quality score, screen time, and participation in sports (model 2). The covariates in model 2 were selected on the basis of theory or previous literature and were included if they resulted in a change of ≥10% in effect estimates when entered individually in model 1. The following covariates were considered, but were not included as they did not fulfill this criterion: maternal parity, folic acid supplement use during pregnancy, alcohol consumption during pregnancy, child's timing of introduction of solids, and child's food allergies. The final model was additionally adjusted for child BMI-SDS at 1 year of age to evaluate whether protein intake at age 1 year predicted body composition at age 6 years independent of baseline BMI (model 3).

On the basis of previously published literature, we assessed whether the associations differed by sex,[24, 27] ethnicity,[28] birth weight,[29] genetic risk score for BMI,[19] catch-up growth,[29] or age at dietary assessment.[8] To examine this, we examined the statistical interaction by adding the product term of total protein intake with the potential effect modifier to model 2. Interactions with age and sex were evaluated in models with body composition unadjusted for age and sex, respectively.

We performed several sensitivity analyses. Firstly, because the FFQ was developed for Dutch children,[13] we repeated our analyses restricted to children with a Dutch ethnic background. Secondly, to test whether it mattered if protein was consumed at the expense of fat or carbohydrate, we performed macronutrient substitution analyses in which we replaced total fat intake by carbohydrate intake, or by saturated, monounsaturated, and polyunsaturated fat intake. Thirdly, to examine whether associations were mediated by differences in insulin levels,[24] we additionally adjusted the multivariable

models for blood insulin levels at 6 years. Fourthly, we examined the direct effect and confounding effect of catch-up growth in the first year of life by adding this variable to model 2. Finally, we examined associations of protein intake at 2 years with body composition at 6 years in the subgroup of children (*n*=698) who had dietary data at this age, using the same models as for protein at 1 year.

To reduce potential bias associated with missing data, we imputed missing values of covariates (n=10 imputations),[30] according to the Fully Conditional Specification method (predictive mean matching) and assuming no monotone missing pattern (Supplementary Table 4.3.1). Analyses were performed in each of ten imputed datasets separately and results were pooled. Because we observed similar effect estimates in the original and imputed datasets, we present pooled estimates from the imputed datasets. Statistical analyses were performed using SPSS version 21.0 (IBM Corp., Armonk, NY, USA) and R version 3.1.2 (R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

Subject characteristics

Characteristics of the children and their mothers are presented in **Table 4.3.1**, for the whole group as well as for boys and girls separately. At the age of 1 year, mean (± SD) daily protein intake of the children was 41.2 gram (± 12.9), which corresponds to 12.9% of their total energy intake. Although this amount is higher than recommended for this age group,[31] it is similar to protein intake observed in other Western pediatric populations.[32, 33] Mean animal protein intake was 8.1 E% (± 2.4) and mean vegetable protein intake 4.7 E% (± 1.4). Boys had a higher absolute protein and energy intake than girls, but relative protein intake was the same for boys and girls (12.9 E%). At the age of 6 years, boys were taller and heavier than girls. Mean BMI was similar, but girls had a higher mean FMI and a lower mean FFMI than boys (**Table 4.3.1**). Subject characteristics before and after multiple imputation are presented in **Supplementary Table 4.3.2**.

Table 4.3.1. Subject characteristics[1]

	All (*n*=2,911)	**Boys (*n*=1,422)**	**Girls (*n*=1,489)**	***p*-value[2]**
Parental characteristics				
Maternal age (y)	31.9 (21.7-39.7)	31.8 (21.7-39.9)	31.8 (21.7-39.7)	*0.77*
Maternal BMI at enrollment (kg/m^2)	23.7 (18.9-35.2)	23.4 (18.7-35.2)	23.6 (19.0-36.0)	*0.34*
Higher maternal education (%)	58.4	60.0	57.3	*0.30*
High household income (%)	66.9	66.4	67.7	*0.40*
Smoking during pregnancy (%)				*0.06*
Never	77.9	78.5	78.0	
Until pregnancy was known	10.1	8.7	11.2	
Continued	12.0	12.8	10.8	
Infant characteristics				
Girls (%)	51.2	-	-	-
Dutch ethnicity (%)	68.9	69.4	68.4	*0.54*
Gestational age at birth (wk)	40.0 (1.7)	40.0 (1.8)	40.1 (1.7)	*0.40*

Table 4.3.1. Subject characteristics[1] (continued)

	All (*n*=2,911)	Boys (*n*=1,422)	Girls (*n*=1,489)	*p*-value[2]
Birth weight (g)	3452 (568)	3521 (577)	3387 (553)	*<0.001*
Breastfeeding in the first 4 months (%)				*0.92*
Exclusive	30.4	30.7	30.1	
Partial	60.1	59.9	60.2	
Never	9.5	9.4	9.7	
Child characteristics at 1 year				
Body mass index (kg/m^2)	17.4 (1.6)	17.6 (1.4)	17.2 (1.3)	*<0.001*
Catch-up growth first year (%)	21.4	22.9	20.0	*0.009*
Age at dietary measurement (mo)	12.9 (12.2-19.0)	12.9 (12.2-19.1)	12.9 (12.2-18.8)	*0.87*
Total energy intake (kcal/d)	1265 (676-2207)	1315 (691-2205)	1222 (651-2232)	*<0.001*
Diet quality score[3]	4.2 (1.3)	4.2 (1.3)	4.1 (1.3)	*0.04*
Protein intake (g/d)				
Total protein	41.8 (12.7)	42.4 (13.3)	40.0 (12.4)	*<0.001*
Animal protein	26.3 (10.2)	26.3 (10.6)	25.1 (10.0)	*0.006*
Dairy protein[4]	18.5 (8.3)	18.7 (8.1)	17.6 (8.4)	*0.22*
Non-dairy animal protein	7.7 (4.7)	7.5 (6.1)	7.3 (5.3)	*0.008*
Vegetable protein	15.0 (5.6)	15.5 (5.5)	14.4 (5.7)	*0.004*
Macronutrient composition of the diet				
Total protein intake (E%)	12.9 (2.4)	12.9 (2.4)	12.9 (2.4)	*0.94*
Total fat intake (E%)	28.6 (5.6)	28.6 (5.6)	28.6 (5.6)	*0.87*
Total carbohydrate intake (E%)	58.4 (6.0)	58.4 (5.9)	58.4 (6.0)	*0.95*
Child characteristics at 6 year visit				
Age (y)	5.9 (5.7-6.6)	5.9 (5.7-6.6)	5.9 (5.7-6.6)	*0.60*
Screen time (h/d)	1.2 (0.3-4.4)	1.3 (0.2-4.5)	1.2 (0.2-4.3)	*<0.001*
Participation in sports (%)	44.3	43.0	45.6	*0.15*
Height (cm)	118.2 (5.2)	118.5 (5.1)	117.9 (5.2)	*0.002*
Weight (kg)	22.4 (3.4)	22.5 (3.4)	21.7 (3.4)	*0.15*
Body mass index (kg/m^2)	16.0 (1.6)	16.0 (1.6)	16.0 (1.7)	*0.75*
Fat mass index (kg/m^2)	3.8 (1.2)	3.5 (1.1)	4.2 (1.2)	*<0.001*
Fat-free mass index (kg/m^2)	11.9 (0.9)	12.2 (0.8)	11.6 (0.8)	*<0.001*
Body fat percentage (%)	24.2 (16.2-36.4)	21.1 (15.7-33.5)	25.6 (18.8-37.5)	*<0.001*
Android/gynoid fat ratio	0.24 (0.06)	0.24 (0.05)	0.25 (0.06)	*0.09*
Overweight or obese (%)[5]	13.5	10.9	16.1	*<0.001*

[1] Values are percentages for categorical variables, means (SD) for continuous variables with a normal distribution, or medians (95% range) for continuous variables with a skewed distribution.
[2] p-values for differences in means between boys and girls, assessed using independent t-tests for continuous variables with a normal distribution, Mann Whitney U tests for continuous variables with a skewed distribution, and chi-square tests for categorical variables.
[3] Diet quality score for preschool children with a theoretical range of 0 to 10 (Voortman et al., 2015)
[4] Protein from dairy products including dairy-based infant formulas (mean 6.9 g/d) and breast milk (mean 0.6 g/d).
[5] According to international age- and sex-specific cut-offs for BMI (Cole et al., 2000)

Protein intake and body composition

Associations between protein intake at age 1 year and body composition at 6 years are presented in **Table 4.3.2**. In confounder-adjusted models (model 2), a 10 g/d higher total protein intake at 1 year was associated with a 0.05 SD (95%CI 0.00, 0.09) higher BMI and a 0.06 SD (95%CI 0.01, 0.11) higher FMI at 6 years, but not with FFMI (-0.01 SD (95%CI -0.06, 0.05)). The positive association with FMI slightly attenuated but remained statistically significant after additional adjustment for baseline BMI (model 3). The association of protein intake with BF% was similar to that observed for FMI, but protein intake was not associated with A/G fat ratio (**Supplementary Table 4.3.3**). Associations between protein intake and FMI were stronger for animal protein than for vegetable protein intake (**Table 4.3.2**). *Post hoc* analyses in which we further separated animal protein into protein from dairy versus non-dairy sources showed that associations with FMI were similar for protein from dairy (0.04 SD (95%CI 0.00, 0.08) versus protein from meat, fish, and eggs (0.05 SD (95%CI 0.00, 0.10).

Table 4.3.2. Associations of protein intake at the age of 1 year with body composition at the age of 6 years (*n*=2,911).

	BMI (SDS)	FMI (SDS)	FFMI (SDS)
Total protein intake (10 g/d)			
Model 1 (crude)	**0.06 (0.01, 0.10)**	**0.06 (0.02, 0.10)**	0.02 (-0.03, 0.06)
Model 2 (confounders)	**0.05 (0.01, 0.09)**	**0.06 (0.02, 0.09)**	0.02 (-0.03, 0.06)
Model 3 (baseline BMI)	0.03 (-0.01, 0.06)	**0.05 (0.00, 0.09)**	-0.01 (-0.05, 0.03)
Animal protein intake (10 g/d)			
Model 1 (crude)	**0.05 (0.01, 0.09)**	**0.06 (0.01, 0.10)**	0.01 (-0.03, 0.06)
Model 2 (confounders)	**0.06 (0.01, 0.10)**	**0.05 (0.01, 0.09)**	0.02 (-0.03, 0.07)
Model 3 (baseline BMI)	0.02 (-0.02, 0.06)	**0.05 (0.00, 0.10)**	-0.02 (-0.06, 0.03)
Vegetable protein intake (10 g/d)			
Model 1 (crude)	0.04 (-0.04, 0.11)	0.02 (-0.06, 0.09)	0.04 (-0.04, 0.12)
Model 2 (confounders)	0.01 (-0.07, 0.08)	-0.01 (-0.08, 0.07)	0.00 (-0.09, 0.08)
Model 3 (baseline BMI)	0.02 (-0.05, 0.09)	0.00 (-0.07, 0.07)	0.01 (-0.07, 0.09)

Values are based on multivariable linear regression models and reflect differences and 95% confidence intervals in body composition outcomes (age- and sex-specific SD scores) per 10 g/d higher energy-adjusted protein intake. In our population, a 0.05 SD higher BMI corresponds to an approximately 0.1 kg/m^2 higher BMI (Table 1). ***Bold*** *values indicate statistically significant effect estimates.*
Model 1 is adjusted for child sex, total energy intake at 1 y, and age at body composition measurement. Models with animal protein intake (energy-adjusted) are additionally adjusted for vegetable protein (energy-adjusted) intake and vice versa.
Model 2 is additionally adjusted for maternal age, pre-pregnancy BMI, educational level, and smoking during pregnancy; household income; and child's ethnicity, birth weight SD-score, breastfeeding in the first 4 months of life, age at dietary assessment, total fat intake at 1 y, diet quality score at 1 y, screen time at 6 y, and participation in sports at 6 y.
Model 3 is additionally adjusted for BMI at the age of 1 year
Abbreviations: BMI, body mass index; FMI, fat mass index; FFMI, fat-free mass index.

Stratified analyses

There were significant interactions between protein intake and sex for FMI (p=0.03), but not for BMI (p=0.15) or FFMI (p=0.24). After stratifying our analyses for sex, we observed that the association of protein intake with FMI was similar in direction, but slightly stronger in girls than in boys (**Supplementary Table 4.3.4**). We also observed a significant interaction of protein intake with catch-up growth in the first year on BMI and FMI at 6 years (p=0.002 and p<0.001, respectively). Results of stratified analyses showed that the positive associations between dietary protein and body fat were stronger among children who had experienced catch-up growth in the first year of life than among those without catch-up growth (**Table 4.3.4 and Supplementary Table 4.3.3**), whereas protein intake was similar in both groups (12.9 E% vs. 13.1 E%). Catch-up growth itself was associated with both fat and fat-free mass: after adjustment for confounders (model 2), children who experienced catch-up growth had a 0.22 (95%CI 0.13, 0.31) higher FFMI and a 0.26 (95% 0.19, 0.34) higher FMI than children who had not had catch-up growth. We observed no interactions of protein intake with ethnicity, age at dietary measurement, birth weight, or genetic risk score for BMI on any of the outcomes (p for interaction>0.05).

Table 4.3.3. Associations of protein intake at the age of 1 year with body composition at the age of 6 years, in children with or without catch-up growth in the first year of life (n=2,911).

Total protein intake (10 g/d)	BMI (SDS)	FMI (SDS)	FFMI (SDS)
No catch-up growth (n=2,287)			
Model 1 (crude)	0.02 (-0.02, 0.07)	0.02 (-0.02, 0.07)	0.00 (-0.06, 0.05)
Model 2 (confounders)	0.02 (-0.02, 0.06)	0.02 (-0.02, 0.06)	-0.01 (-0.06, 0.05)
Model 3 (baseline BMI)	0.00 (-0.04, 0.04)	0.01 (-0.03, 0.05)	-0.02 (-0.07, 0.02)
Catch-up growth (n=624)			
Model 1 (crude)	**0.18 (0.06, 0.29)**	**0.18 (0.07, 0.30)**	0.08 (-0.03, 0.20)
Model 2 (confounders)	**0.18 (0.08, 0.29)**	**0.19 (0.08, 0.30)**	0.09 (-0.02, 0.20)
Model 3 (baseline BMI)	**0.14 (0.04, 0.23)**	**0.15 (0.05, 0.26)**	0.05 (-0.06, 0.15)

Values are based on multivariable linear regression models and reflect differences and 95% confidence intervals in body composition outcomes (age- and sex-specific SD scores) per 10 g/d higher energy-adjusted protein intake. In our population, a 0.18 SD higher BMI corresponds to an approximately 0.3 kg/m^2 higher BMI (Table 1). ***Bold*** *values indicate statistically significant effect estimates. P-values for interaction between protein intake and catch-up growth were p=0.002 for BMI, p<0.001 for FMI, and p=0.17 for FFMI.*

Model 1 is adjusted for child sex, total energy intake at 1 y and age at body composition measurement; models with animal protein intake are additionally adjusted for vegetable protein intake and vice versa.

Model 2 is additionally adjusted for maternal age, pre-pregnancy BMI, educational level, and smoking during pregnancy; household income; and child's ethnicity, birth weight SD-score, breastfeeding in the first 4 months of life, age at dietary assessment, total fat intake at 1 y, diet quality score at 1 y, screen time at 6 y, and participation in sports at 6 y.

Model 3 is additionally adjusted for BMI at the age of 1 y.

Abbreviations: BMI, body mass index; FMI, fat mass index; FFMI, fat-free mass index, SDS, standard deviation score.

Sensitivity analyses

We performed several sensitivity analyses, but results remained similar as obtained in the original analyses. Analyses restricted to children with a Dutch background revealed similar effect estimates as for the whole group (**Supplementary Table 4.3.5**). Macronutrient substitution models in which we replaced total fat intake by either carbohydrate intake or by saturated, monounsaturated and polyunsaturated fat intake revealed similar effect estimates for protein intake (data not shown). Additional adjustment for insulin concentrations at 6 years as potential mediator, or for catch-up growth in the first year, did not affect the effect estimates for protein intake on BMI and FMI (data not shown).

Finally, protein intake at age 2 years was slightly higher (13.9 ± 1.9 E%) than at age 1 year,[34] but among the subgroup with dietary data at both the age of 1 and 2 years (*n*=649) associations with body composition were similar for protein intake at both ages (**Supplementary Table 4.3.6**).

DISCUSSION

Results from this large population-based cohort study suggest that protein intake in early childhood is associated with a higher fat mass index, but not with fat-free mass index at school age. These associations were stronger for protein from animal than from vegetable sources, but did not differ by food source of animal protein (i.e., dairy versus non-dairy). Furthermore, associations were slightly larger in girls than in boys, and were stronger among children who had experienced catch-up growth in the first year of life.

Comparison with previous studies

A number of observational studies showed that a higher protein intake in the complementary feeding period is associated with a higher BMI in later childhood.[1, 35-37] The causal effect of a higher protein intake in infancy on a higher BMI in later childhood was also confirmed in a large randomized trial in which children received high-protein or lower-protein infant formulas in their first year of life.[2] In our study, we also found that a higher protein intake at 1 year of age is associated with a higher BMI, but we additionally show that it is specifically associated with a higher fat mass, but not with fat-free mass index.

To our knowledge, only four previous studies examined protein intake in early life in relation to later measures of body fat.[1, 3, 38, 39] Mean protein intake in these populations ranged from 13 to 15 E%, which is close to the 12.9 E% in our study population. In the first study, Rolland-Cachera *et al.* reported that, among 112 French children, protein intake at the age of 2 years was associated with a higher BMI and higher subscapular skinfold thickness at the age of 8 years.[1] In the second study, performed by Hoppe *et al.*, protein intake at 9 months of age was not associated with BF% (assessed with DXA) at the age of 10 years in 142 Danish children.[38] The absence of an association in this study might be explained by the earlier age at which protein intake was measured, as results from the third study, performed by Günther *et al.*, suggests that the effect of protein intake on obesity risk may have a specific risk window, i.e., around the age of 12 months.[8] In this study in 203 German children, higher protein intake at the age of 12 months, but not at 6 or 18-24 months, was associated with a higher BF% (calculated from skinfold thicknesses) at the age of 7 years.[3] However, this relation was no longer

present when the children reached young adulthood (18-25 years).[40] In the fourth study, neither protein intake at 6 months nor at 12 months was associated with fat mass assessed with DXA at the age of 4 years in 556 British children.[39] In a smaller subgroup of our study population, we observed no differences in associations for protein intake at 1 year or 2 years and no interaction between age of dietary assessment and protein intake on child body composition. Thus, our results do not confirm that the period around the age of 1 year poses a specific risk window for high protein intake.

We observed no associations between protein intake in early childhood and FFMI at age 6 years. This is in contrast to findings in adults or older children, in whom several studies reported associations between high protein intake and higher FFMI or sparing of FFMI loss during weight loss.[41-43] To our knowledge, only one previous study has examined the association of protein intake at preschool age with measures of later fat-free mass. In this follow-up of 159 children participating in the previously mentioned study by Günther *et al.*, Assmann *et al.* reported that protein intake in early childhood was no longer associated with measures of body fat and was not associated with FFMI in young adulthood (18-25 years).[40] However, higher protein intake between the ages of 9 and 15 years was associated with a higher FFMI in young adulthood.[40] This association between protein intake and FFMI is in line with studies in adults.[44-46] Our study further supports the existing literature that suggests that higher protein intake in early childhood may have very different effects on body composition than protein intake in later life.[1]

Potential mechanisms

In adults and older children, a higher protein intake seems to increase weight loss while sparing fat-free mass.[43] A potential underlying mechanism is a higher satiety from high-protein diets, via increased release of satiety hormones and consequently a lower energy intake.[4, 5] However, a high protein intake in adults has also been associated with weight loss independent of energy intake, possibly because of the high thermic effect of protein resulting in a slightly increased energy expenditure.[4]

In early childhood, different mechanisms may play a role. The early protein hypothesis postulates that high protein in infancy triggers hormonal responses that may cause rapid weight gain in early life and persistent changes in later obesity risk.[1, 47, 48] Protein intake stimulates the secretion of insulin and insulin-like growth factor-1 (IGF-1),[6, 24] which may in turn increase growth and adipogenesis[7, 49] and have a long-term effect on body fat.[50] Unfortunately, in our study we did not measure IGF-1 or other growth hormone levels at the age of 1 year. We previously observed that protein intake was positively associated with insulin concentrations at age 6 years,[24] but adjustment for insulin levels had no effect on the association between protein intake and FMI in our population, suggesting that insulin at 6 years does not explain this association.

Associations between protein intake and FMI were stronger for animal than for vegetable protein intake. In line with this, a previous study in 2.5-year-old children found that dietary protein from animal, but not from vegetable sources, was associated with higher IGF-1 concentrations.[7] To our knowledge, only one previous study examined protein from different sources in early childhood in relation body composition. In line with our results, they found that animal but not vegetable protein intake was associated with a higher body fat.[8] More specifically, they observed that intake of dairy protein, but not meat protein, was associated with a higher BMI and body fat at the age of 7 years.[8]

Contrary to these results, additional analyses in our study showed that associations were similar for protein intake from dairy sources and from non-dairy animal sources.

It has been previously suggested that certain subgroups may be extra sensitive to the adverse effect of high protein intake in early life.[2] In our population we observed slightly stronger associations in girls than in boys. This may be explained by differences in the IGF-1 axis response to high protein intake, which a previous study suggested to be stronger in girls than in boys.[27] In line with this, we have previously shown associations of protein intake with insulin levels in girls, but not in boys.[24] The observed sex differences may also be related to a difference in timing of adiposity rebound (AR) between boys and girls, or to a difference in peak BMI at AR. This period, during which a child's BMI curve starts to rise again after a period of decrease during early childhood, usually occurs between 5 and 7 years of age,[51] corresponding to the age of outcome measurements in our study (6 years). An early AR and a higher BMI at AR are associated with an increased risk of obesity.[51] A higher intake of protein in early childhood has been linked to an earlier AR,[1] although not consistently;[52] and with a higher BMI at AR in girls, but not in boys.[52]

Our results suggest that growth patterns in infancy may also influence the relation between protein intake and later body composition, as we observed an interaction between catch-up growth in the first year of life –but not with birth weight– and protein intake at 1 year on later BMI and FMI. Catch-up growth in early childhood is an important risk factor for later obesity,[29, 53, 54] and our results suggest that especially children who experienced catch-up growth in the first year of life may have a higher risk of obesity with higher protein consumption. These children may already have a 'thrifty' energy metabolism leading to increased fat accumulation,[9] which may be further exacerbated by high protein intake. Previous studies also suggested that the effect of protein intake on obesity may be stronger in children who are at genetic risk to become obese,[19] but we did not find evidence for this interaction effect in our population. However, our study may have been underpowered to identify such an interaction. Future studies should further examine which children are at increased risk for adverse effects of high protein intake in early life.

Strengths and limitations

Main strengths of this study are its prospective population-based cohort design, the availability of information on many child and parental covariates, and the detailed measurements of body composition. We not only measured height and weight to calculate BMI, but we also measured body fat and fat-free mass with DXA. Because a given BMI can encompass a wide range of fat mass in childhood, BMI is considered to be of limited use to measure adiposity in children.[11] Although the effect estimates were relatively small and probably without direct consequences for individual children, the effects may be relevant on a population level in predicting later obesity.[55]

Unfortunately, not all children participating in the study were included in the current analyses, mainly because of missing dietary data. On average, mothers who filled out the FFQ had a higher educational level, a higher household income, and a healthier lifestyle than mothers who did not.[14] For missing values of covariates, we used multiple imputation procedures.[30]

Strengths of the dietary assessment in our study are that we used an extensive FFQ which was developed specifically for our population of young children, and that we calculated not only total dietary protein, but also protein from different food sources. A limitation is that there was no dietary

data available around the age of 6 years to examine possible confounding or effect modification by diet later in childhood. Another limitation of our dietary assessment is that food intake reported using an FFQ is subject to substantial measurement error.[56] By adjusting for total energy intake, we aimed to reduce this measurement error.[56, 57] However, when keeping energy intake constant, a higher protein intake involves a lower intake of fat and/or carbohydrate. Adjusting our models for either total fat, fat subtypes, or carbohydrate intake in macronutrient substitution models resulted in similar effect estimates for protein intake, suggesting that for body composition it makes no difference if dietary protein is replaced by carbohydrate or by fat in early childhood.

Implications

Previously, a large randomized controlled trial showed that a higher protein intake in the first year of life leads to a higher BMI in later childhood.[2] In our study, we observed that this higher BMI in relation to protein intake is specifically explained by a higher fat mass, and not a higher fat-free mass. Therefore, results from our study and previous studies combined,[1-3] suggest that high protein intake in early childhood increases the risk of adiposity. A reduction in protein intake in infancy and early childhood may thus be advised to prevent obesity, for example by lowering the amounts of protein in infant formula and toddler foods. However, a too low protein intake may restrict growth. Therefore future studies should explore what the optimal range of protein intake is for infants and young children for optimal growth[58] and other aspects of later health.[59] Furthermore, future studies should explore whether intake of different types of protein or certain amino acids are specifically associated with increased adiposity, for example by studying the endocrinal response to different protein subtypes. Results from these studies can be used to further optimize dietary recommendations regarding protein intake in early childhood.

Conclusions

In this prospective cohort study, high protein intake in early childhood is associated with higher fat mass, but not fat-free mass index in later childhood. This association was stronger for animal protein than for vegetable protein, but did not differ for dairy versus non-dairy animal protein. Further research should explore the underlying mechanisms and study the optimal amount of protein intake needed in infancy and early childhood for optimal growth and body composition.

SUPPLEMENTARY MATERIAL

Supplementary Table 4.3.1. Details of the multiple imputation procedure

Software	SPSS version 21 for Windows
Imputation method and key settings	Fully conditional specification (Markov chain Monte Carlo method); maximum iterations: 20
Number of imputed data sets created	10
Variables included as both predictor variable as a variable to be imputed	Maternal age, parity, educational level, and marital status at enrollment, gestational age at fatty acid measurement, ethnicity, pre-pregnancy BMI, alcohol consumption during pregnancy, smoking during pregnancy, folic acid supplement use during early pregnancy, energy intake, maternal diet quality score, gestational weight gain, maternal BMI at 6y follow-up, maternal educational level and employment status at 6 y follow-up; paternal education at enrollment and follow-up; household income at enrollment and at follow-up; and child's gestational age at birth, birthweight, birthweight Z-score, sex, breastfeeding, food allergies, timing of introduction of solid foods, age at follow-up, participation in sports, screen time
Variables included as predictors of missing data only	Maternal height and weight pre-pregnancy, in early, mid, and late pregnancy, and 6 years after delivery, maternal smoking and alcohol use in early, mid, and late pregnancy, serum folate levels in early pregnancy, gestational age at intake; paternal age, smoking, employment status, and ethnicity; child's diet quality score, dietary intake of calories, total protein, animal protein, vegetable protein, total fat, saturated fat, monounsaturated fat, polyunsaturated fat, carbohydrates; height, weight and BMI at the ages of 5, 12, 15, 20, 25, 32, 40, and 50 months; children's height, weight, BMI, fat mass, fat-free mass, insulin levels, C-peptide levels, total, HDL and LDL cholesterol levels, triacylglycerol, systolic and diastolic blood pressure, TV watching, playing outside, and computer use at the age of 6 years.
Treatment of not normally distributed variables	Predictive mean matching
Treatment of binary or categorical variables	Logistic regression and multinomial models
Population	All children with complete information on dietary intake at 1 year and body composition at 6 years (n=2,911)

Supplementary Table 4.3.2. Subject characteristics before and after imputation

	Unimputed	Imputed
Parental characteristics		
Maternal age (y)	31.9 (21.7-39.7)	31.9 (21.7-39.7)
Missing	*0.0%*	-
Maternal BMI at enrollment (kg/m^2)	23.5 (18.9-35.6)	23.7 (18.9-35.2)
Missing	*8.4%*	-
Higher maternal education (%)	59.3	58.4
Missing	*3.1%*	-
High household income (%)	66.1	66.9
Missing	*15.1%*	-
Smoking during pregnancy (valid %)		
Never	78.1	77.9
Until pregnancy was known	10.2	10.1
Continued	11.7	12.0
Missing	*9.8%*	-
Infant characteristics		
Girls (%)	51.2	51.2
Missing	*0.0%*	-
Dutch ethnicity (%)	68.9	68.9
Missing	*0.4%*	-
Gestational age at birth (wk)	39.9 (1.8)	40.0 (1.7)
Missing	*<0.1%*	-
Birth weight (g)	3452 (567)	3452 (568)
Missing	*<0.1%*	-
Breastfeeding in the first 4 months (valid %)		
Exclusive	29.2	30.4
Partial	61.5	60.1
Never	9.2	9.5
Missing	*10.1%*	-
Child characteristics at 1 year		
Body mass index (kg/m^2)	17.3 (1.5)	17.4 (1.6)
Missing	*16.1%*	-
Catch-up growth first year (%)	20.9	21.4
Missing	*17.8%*	-
Age at dietary measurement (mo)	12.9 (12.2-19.0)	not imputed
Total energy intake (kcal/d)	1265 (676-2207)	not imputed
Diet quality score[3]	4.2 (1.3)	not imputed
Protein intake (g/d)		
Total protein	41.8 (12.7)	not imputed
Animal protein	26.3 (10.2)	not imputed

Supplementary Table 4.3.2. Subject characteristics before and after imputation (continued)

	Unimputed	Imputed
Dairy protein[4]	18.5 (8.3)	not imputed
Non-dairy animal protein	7.7 (4.7)	not imputed
Vegetable protein	15.0 (5.6)	not imputed
Macronutrient composition of the diet		
Total protein intake (E%)	12.9 (2.4)	not imputed
Total fat intake (E%)	28.6 (5.6)	not imputed
Total carbohydrate intake (E%)	58.4 (6.0)	not imputed
Child characteristics at 6 year visit		
Age (y)	5.9 (5.7-6.6)	not imputed
Screen time (h/d)	1.2 (0.3-4.4)	1.2 (0.3-4.4)
Missing	*18.3%*	-
Participation in sports (%)	44.9	44.1
Missing	*7.3%*	-
Height (cm)	118.2 (5.2)	not imputed
Weight (kg)	22.4 (3.4)	not imputed
Body mass index (kg/m^2)	16.0 (1.6)	not imputed
Fat mass index (kg/m^2)	3.8 (1.2)	not imputed
Fat-free mass index (kg/m^2)	11.9 (0.9)	not imputed
Body fat percentage (%)	24.2 (16.2-36.4)	not imputed
Android/gynoid fat ratio	0.24 (0.06)	not imputed
Overweight or obese (%)[5]	13.5	not imputed

[1] Values are percentages for categorical variables, means (SD) for continuous variables with a normal distribution, or medians (95% range) for continuous variables with a skewed distribution.
[2] p-values for differences in means between boys and girls, assessed using independent t-tests for continuous variables with a normal distribution, Mann Whitney U tests for continuous variables with a skewed distribution, and chi-square tests for categorical variables.
[3] Diet quality score for preschool children with a theoretical range of 0 to 10 (Voortman et al., 2015)
[4] Protein from dairy products including dairy-based infant formulas (mean 6.9 g/d) and breast milk (mean 0.6 g/d).
[5] According to international age- and sex-specific cut-offs for BMI (Cole et al., 2000)

Supplementary Table 4.3.3. Associations of protein intake at the age of 1 year with body fat percentage and android/gynoid ratio at the age of 6 years (*n*=2,911).

Total protein intake (10 g/d)	BF% (SDS)	A/G ratio (SDS)
All children (*n*=2,911)		
Model 1 (crude)	**0.06 (0.02, 0.11)**	0.03 (-0.01, 0.07)
Model 2 (confounders)	**0.06 (0.02, 0.10)**	0.04 (-0.01, 0.09)
Model 3 (baseline BMI)	**0.05 (0.00, 0.10)**	0.04 (-0.02, 0.09)
Girls (*n*=1,489)		
Model 1 (crude)	**0.10 (0.03, 0.16)**	0.05 (-0.01, 0.11)
Model 2 (confounders)	**0.09 (0.02, 0.16)**	0.06 (-0.02, 0.14)
Model 3 (baseline BMI)	**0.08 (0.01, 0.15)**	0.06 (-0.02, 0.14)
Boys (*n*=1,422)		
Model 1 (crude)	0.04 (-0.03, 0.10)	0.01 (-0.05, 0.08)
Model 2 (confounders)	0.05 (-0.02, 0.12)	0.02 (-0.06, 0.09)
Model 3 (baseline BMI)	0.03 (-0.04, 0.09)	0.01 (-0.06, 0.09)
No catch-up growth (*n*=2,287)		
Model 1 (crude)	0.03 (-0.02, 0.08)	0.00 (-0.05, 0.04)
Model 2 (confounders)	0.02 (-0.02, 0.07)	-0.01 (-0.06, 0.04)
Model 3 (baseline BMI)	0.01 (-0.03, 0.06)	-0.01 (-0.06, 0.03)
Catch-up growth (*n*=624)		
Model 1 (crude)	**0.17 (0.06, 0.29)**	**0.16 (0.04, 0.27)**
Model 2 (confounders)	**0.17 (0.06, 0.28)**	**0.16 (0.04, 0.27)**
Model 3 (baseline BMI)	**0.15 (0.04, 0.26)**	**0.14 (0.02, 0.26)**

Values are based on multivariable linear regression models and reflect differences and 95% confidence intervals in body composition outcomes (age- and sex-specific SD scores) per 10 g/d higher energy-adjusted protein intake. ***Bold*** *values indicate statistically significant effect estimates.*
Model 1 is adjusted for total energy intake at 1 y and age at body composition measurement; models with animal protein intake are additionally adjusted for vegetable protein intake and vice versa.
Model 2 is additionally adjusted for maternal age, pre-pregnancy BMI, educational level, and smoking during pregnancy; household income; and child's ethnicity, birth weight SD-score, breastfeeding in the first 4 months of life, age at dietary assessment, total fat intake at 1 y, diet quality score at 1 y, screen time at 6 y, and participation in sports at 6 y.
Model 3 is additionally adjusted for BMI at the age of 1 y.
Abbreviations: BF%, body fat percentage; A/G ratio, android/gynoid fat ratio

Supplementary Table 4.3.4. Associations of protein intake at the age of 1 year with body composition at the age of 6 years, stratified by child sex (*n*=2,911).

Total protein intake (10 g/d)	BMI (SDS)	FMI (SDS)	FFMI (SDS)
Girls (*n*=1,489)			
Model 1 (crude)	**0.07 (0.01, 0.13)**	**0.09 (0.03, 0.15)**	0.00 (-0.07, 0.07)
Model 2 (confounders)	**0.06 (0.00, 0.12)**	**0.08 (0.02, 0.13)**	-0.01 (-0.08, 0.05)
Model 3 (baseline BMI)	0.05 (0.00, 0.11)	**0.07 (0.02, 0.12)**	-0.02 (-0.08, 0.04)
Boys (*n*=1,422)			
Model 1 (crude)	0.05 (-0.01, 0.11)	0.04 (-0.02, 0.10)	0.03 (-0.03, 0.20)
Model 2 (confounders)	0.04 (-0.02, 0.11)	0.05 (-0.02, 0.12)	0.02 (-0.06, 0.10)
Model 3 (baseline BMI)	0.01 (-0.06, 0.07)	0.03 (-0.04, 0.09)	-0.02 (-0.09, 0.06)

Values are based on multivariable linear regression models and reflect differences and 95% confidence intervals in body composition outcomes (age- and sex-specific SD scores) per 10 g/d higher energy-adjusted protein intake. ***Bold*** *values indicate statistically significant effect estimates. P-values for interaction between protein intake and sex were p=0.15 for BMI, p=0.03 for FMI, and p=0.24 for FFMI.*
Model 1 is adjusted for child sex, total energy intake at 1 y and age at body composition measurement; models with animal protein intake are additionally adjusted for vegetable protein intake and vice versa.
Model 2 is additionally adjusted for maternal age, pre-pregnancy BMI, educational level, and smoking during pregnancy; household income; and child's ethnicity, birth weight SD-score, breastfeeding in the first 4 months of life, age at dietary assessment, total fat intake at 1 y, diet quality score at 1 y, screen time at 6 y, and participation in sports at 6 y.
Model 3 is additionally adjusted for BMI at the age of 1 y.
Abbreviations: BMI, body mass index; FMI, fat mass index; FFMI, fat-free mass index.

Supplementary Table 4.3.5. Associations of protein intake at the age of 1 year with childhood body composition at the age of 6 years, in Dutch children only (*n*=2,006).

Total protein intake (10 g/d)	BMI (SDS)	FMI (SDS)	FFMI (SDS)
All children (*n*=2,006)			
Model 1 (crude)	**0.08 (0.04, 0.13)**	**0.06 (0.02, 0.10)**	**0.08 (0.02, 0.13)**
Model 2 (confounders)	**0.08 (0.03, 0.14)**	**0.07 (0.02, 0.12)**	0.06 (-0.01, 0.13)
Model 3 (baseline BMI)	0.04 (-0.01, 0.09)	0.04 (-0.01, 0.09)	0.02 (-0.04, 0.08)
Girls (*n*=1,018)			
Model 1 (crude)	**0.10 (0.03, 0.16)**	**0.10 (0.03, 0.16)**	0.04 (-0.04, 0.12)
Model 2 (confounders)	**0.12 (0.04, 0.20)**	**0.12 (0.05, 0.20)**	0.04 (-0.06, 0.13)
Model 3 (baseline BMI)	**0.10 (0.03, 0.17)**	**0.11 (0.04, 0.18)**	0.02 (-0.07, 0.10)
Boys (*n*=988)			
Model 1 (crude)	**0.07 (0.01, 0.14)**	0.03 (-0.03, 0.08)	**0.11 (0.04, 0.19)**
Model 2 (confounders)	0.05 (-0.03, 0.12)	0.02 (-0.05, 0.09)	0.08 (-0.01, 0.18)
Model 3 (baseline BMI)	-0.01 (-0.08, 0.06)	-0.01 (-0.08, 0.05)	0.03 (-0.06, 0.15)
No catch-up growth (*n*=2,287)			
Model 1 (crude)	0.02 (-0.02, 0.07)	0.02 (-0.02, 0.07)	0.00 (-0.05, 0.05)
Model 2 (confounders)	0.02 (-0.04, 0.07)	0.03 (-0.02, 0.08)	-0.03 (-0.09, 0.04)
Model 3 (baseline BMI)	0.00 (-0.05, 0.05)	0.02 (-0.03, 0.07)	-0.05 (-0.11, 0.01)
Catch-up growth (*n*=624)			
Model 1 (crude)	**0.18 (0.06, 0.29)**	**0.18 (0.06, 0.30)**	0.08 (-0.03, 0.20)
Model 2 (confounders)	**0.18 (0.05, 0.31)**	**0.19 (0.06, 0.33)**	0.09 (-0.05, 0.22)
Model 3 (baseline BMI)	**0.15 (0.03, 0.27)**	**0.17 (0.04, 0.30)**	0.05 (-0.07, 0.17)

Values are based on multivariable linear regression models and reflect differences and 95% confidence intervals in body composition outcomes (age- and sex-specific SD scores) per 10 g/d higher energy-adjusted protein intake. ***Bold*** *values indicate statistically significant effect estimates.*
Model 1 is adjusted for child sex, total energy intake at 1 y and age at body composition measurement; models with animal protein intake are additionally adjusted for vegetable protein intake and vice versa.
Model 2 is additionally adjusted for maternal age, pre-pregnancy BMI, educational level, and smoking during pregnancy; household income; and child's ethnicity, birth weight SD-score, breastfeeding in the first 4 months of life, age at dietary assessment, total fat intake at 1 y, diet quality score at 1 y, screen time at 6 y, and participation in sports at 6 y.
Model 3 is additionally adjusted for BMI at the age of 1 y.
Abbreviations: BMI, body mass index; FMI, fat mass index; FFMI, fat-free mass index.

Supplementary Table 4.3.6. Associations of protein intake at the age of 1 and 2 years with childhood body composition at the age of 6 years (*n*=649).

Total protein intake (10 g/d)	BMI (SDS)	FMI (SDS)	FFMI (SDS)
Protein intake at 1 year (*n*=649)			
Model 1 (crude)	0.05 (-0.03, 0.12)	-0.01 (-0.08, 0.07)	0.08 (-0.02, 0.18)
Model 2 (confounders)	0.03 (-0.06, 0.13)	-0.02 (-0.11, 0.08)	0.07 (-0.04, 0.19)
Model 3 (baseline BMI)	0.00 (-0.09, 0.08)	-0.04 (-0.13, 0.04)	0.03 (-0.07, 0.14)
Protein intake at 2 years (*n*=649)			
Model 1 (crude)	0.03 (-0.06, 0.11)	-0.03 (-0.11, 0.05)	0.09 (-0.02, 0.20)
Model 2 (confounders)	0.06 (-0.05, 0.16)	-0.01 (-0.11, 0.09)	0.10 (-0.02, 0.23)
Model 3 (baseline BMI)	-0.02 (-0.11, 0.07)	-0.07 (-0.17, 0.02)	0.04 (-0.08, 0.16)

Values are based on multivariable linear regression models and reflect differences and 95% confidence intervals in body composition outcomes (age- and sex-specific SD scores) per 10 g/d higher energy-adjusted protein intake. ***Bold*** *values indicate statistically significant effect estimates.*
Model 1 is adjusted for total energy intake at age of dietary assessment and for age at body composition measurement; models with animal protein intake are additionally adjusted for vegetable protein intake and vice versa.
Model 2 is additionally adjusted for maternal age, pre-pregnancy BMI, educational level, and smoking during pregnancy; household income; and child's ethnicity, birth weight SD-score, breastfeeding in the first 4 months of life, age at dietary assessment, total fat intake at age of dietary assessment, diet quality score at age of dietary assessment, screen time at 6 y, and participation in sports at 6 y.
Model 3 is additionally adjusted for BMI at the age of dietary assessment
Abbreviations: BMI, body mass index; FMI, fat mass index; FFMI, fat-free mass index.

REFERENCES

1. Rolland-Cachera MF, Deheeger M, Akrout M, Bellisle F. Influence of macronutrients on adiposity development: A follow up study of nutrition and growth from 10 months to 8 years of age. Int J Obes. 1995;19(8):573-8.
2. Weber M, Grote V, Closa-Monasterolo R, Escribano J, Langhendries JP, Dain E, et al. Lower protein content in infant formula reduces BMI and obesity risk at school age: follow-up of a randomized trial. Am J Clin Nutr. 2014;99(5):1041-51.
3. Gunther ALB, Buyken AE, Kroke A. Protein intake during the period of complementary feeding and early childhood and the association with body mass index and percentage body fat at 7 y of age. Am J Clin Nutr. 2007;85(6):1626-33.
4. Westerterp-Plantenga MS, Lemmens SG, Westerterp KR. Dietary protein - its role in satiety, energetics, weight loss and health. Br J Nutr. 2012;108 Suppl 2:S105-12.
5. Halton TL, Hu FB. The effects of high protein diets on thermogenesis, satiety and weight loss: a critical review. J Am Coll Nutr. 2004;23(5):373-85.
6. Socha P, Grote V, Gruszfeld D, Janas R, Demmelmair H, Closa-Monasterolo R, et al. Milk protein intake, the metabolic-endocrine response, and growth in infancy: Data from a randomized clinical trial. Am J Clin Nutr. 2011;94(6 SUPPL):1776S-84S.
7. Hoppe C, Udam TR, Lauritzen L, Molgaard C, Juul A, Michaelsen KF. Animal protein intake, serum insulin-like growth factor I, and growth in healthy 2.5-y-old Danish children. Am J Clin Nutr. 2004; 80(2):447-52.
8. Gunther ALB, Remer T, Kroke A, Buyken AE. Early protein intake and later obesity risk: Which protein sources at which time points throughout infancy and childhood are important for body mass index and body fat percentage at 7 y of age? Am J Clin Nutr. 2007;86(6):1765-72.
9. Dulloo AG, Jacquet J, Seydoux J, Montani JP. The thrifty 'catch-up fat' phenotype: its impact on insulin sensitivity during growth trajectories to obesity and metabolic syndrome. Int J Obes (Lond). 2006;30 Suppl 4:S23-35.
10. Thorisdottir B, Gunnarsdottir I, Palsson GI, Halldorsson TI, Thorsdottir I. Animal protein intake at 12 months is associated with growth factors at the age of six. Acta Paediatr. 2014;103(5):512-7.
11. Wells JC. A Hattori chart analysis of body mass index in infants and children. Int J Obes Relat Metab Disord. 2000;24(3):325-9.
12. Jaddoe VW, van Duijn CM, Franco OH, van der Heijden AJ, van Iizendoorn MH, de Jongste JC, et al. The Generation R Study: design and cohort update 2012. Eur J Epidemiol. 2012;27(9):739-56.
13. Kiefte-de Jong JC, de Vries JH, Bleeker SE, Jaddoe VW, Hofman A, Raat H, et al. Socio-demographic and lifestyle determinants of 'Western-like' and 'Health conscious' dietary patterns in toddlers. Br J Nutr. 2013;109(1):137-47.
14. Voortman T, Kiefte-de Jong JC, Geelen A, Villamor E, Moll HA, de Jongste JC, et al. The development of a diet quality score for preschool children and its validation and determinants in the Generation R Study. J Nutr. 2015;145(2):306-14.
15. Netherlands Nutrition Centre. Dutch Food Composition Database 2006 (NEVO-tabel 2006). Den Haag: Netherlands Nutrition Centre (Voedingscentrum), 2006.
16. Wells JC, Cole TJ. Adjustment of fat-free mass and fat mass for height in children aged 8 y. Int J Obes Relat Metab Disord. 2002;26(7):947-52.
17. Niklasson A, Ericson A, Fryer JG, Karlberg J, Lawrence C, Karlberg P. An update of the Swedish reference standards for weight, length and head circumference at birth for given gestational age (1977-1981). Acta Paediatr Scand. 1991;80(8-9):756-62.

18. Statistics Netherlands. Immigrants in the Netherlands 2004 (Allochtonen in Nederland 2004). Den Haag/Heerlen: Statistics Netherlands (Centraal Bureau voor de Statistiek), 2004.
19. Qi Q, Downer MK, Kilpelainen TO, Taal HR, Barton SJ, Ntalla I, et al. Dietary intake, FTO genetic variants and adiposity: a combined analysis of over 16,000 children and adolescents. Diabetes. 2015.
20. Vogelezang S, Monnereau C, Gaillard R, Renders CM, Hofman A, Jaddoe VW, et al. Adult adiposity susceptibility loci, early growth and general and abdominal fatness in childhood: the Generation R Study. Int J Obes (Lond). 2015.
21. Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, Jackson AU, et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. Nature genetics. 2010;42(11): 937-48.
22. Fredriks AM, van Buuren S, Burgmeijer RJ, Meulmeester JF, Beuker RJ, Brugman E, et al. Continuing positive secular growth change in The Netherlands 1955-1997. Pediatr Res. 2000;47(3):316-23.
23. Ong KK, Ahmed ML, Emmett PM, Preece MA, Dunger DB. Association between postnatal catch-up growth and obesity in childhood: prospective cohort study. BMJ. 2000;320(7240):967-71.
24. Voortman T, van den Hooven EH, Tielemans MJ, Hofman A, Kiefte-de Jong JC, Jaddoe VW, et al. Protein intake in early childhood and cardiometabolic health at school age: The Generation R Study. Eur J Nutr. 2015:in press.
25. Kruithof CJ, Kooijman MN, van Duijn CM, Franco OH, de Jongste JC, Klaver CC, et al. The Generation R Study: Biobank update 2015. Eur J Epidemiol. 2014;29(12):911-27.
26. Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. Am J Clin Nutr. 1997;65(4 Suppl):1220S-8S; discussion 9S-31S.
27. Closa-Monasterolo R, Ferre N, Luque V, Zaragoza-Jordana M, Grote V, Weber M, et al. Sex differences in the endocrine system in response to protein intake early in life. Am J Clin Nutr. 2011;94(6 Suppl): 1920S-7S.
28. van Rossem L, Hafkamp-de Groen E, Jaddoe VW, Hofman A, Mackenbach JP, Raat H. The role of early life factors in the development of ethnic differences in growth and overweight in preschool children: a prospective birth cohort. BMC public health. 2014;14:722.
29. Ong KK, Loos RJ. Rapid infancy weight gain and subsequent obesity: systematic reviews and hopeful suggestions. Acta Paediatr. 2006;95(8):904-8.
30. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ. 2009;338:b2393.
31. Institute of Medicine. Dietary reference intakes for energy, carbohydrate, fiber, fat, fatty acids, cholesterol, protein, and amino acids. Washington D.C.: National Academy of Sciences, 2005.
32. Hornell A, Lagstrom H, Lande B, Thorsdottir I. Protein intake from 0 to 18 years of age and its relation to health: a systematic literature review for the 5th Nordic Nutrition Recommendations. Food Nutr Res. 2013;57.
33. Ocké MC, Rossum CTMv, Fransen HP, Buurma EJM, Boer EJd, Brants HAM, et al. Dutch National Food Consumption Survey-Young Children 2005/2006. National Institute for Public Health and the Environment (RIVM), 2008.
34. Voortman T, Bakker H, Sedaghat S, Kiefte-de Jong JC, Hofman A, Jaddoe VW, et al. Protein intake in infancy and kidney size and function at the age of 6 years: The Generation R Study. Pediatr Nephrol. 2015.
35. Ohlund I, Hernell O, Hornell A, Stenlund H, Lind T. BMI at 4 years of age is associated with previous and current protein intake and with paternal BMI. Eur J Clin Nutr. 2010;64(2):138-45.
36. Scaglioni S, Agostoni C, De Notaris R, Radaelli G, Radice N, Valenti M, et al. Early macronutrient intake and overweight at five years of age. Int J Obes. 2000;24(6):777-81.

37. Gunnarsdottir I, Thorsdottir I. Relationship between growth and feeding in infancy and body mass index at the age of 6 years. Int J Obes. 2003;27(12):1523-7.
38. Hoppe C, Molgaard C, Thomsen BL, Juul A, Michaelsen KF. Protein intake at 9 mo of age is associated with body size but not with body fat in 10-y-old Danish children. Am J Clin Nutr. 2004;79(3):494-501.
39. Robinson S, Crozier S, Marriott L, Harvey N, Inskip H, Godfrey K, et al. Protein intake in infancy is not associated with adiposity at 4 years of age: findings from the Southampton Women's Survey. Proceedings of the Nutrition Society2008. p. E347.
40. Assmann KE, Joslowski G, Buyken AE, Cheng G, Remer T, Kroke A, et al. Prospective association of protein intake during puberty with body composition in young adulthood. Obesity (Silver Spring). 2013;21(12):E782-9.
41. Westerterp-Plantenga MS, Nieuwenhuizen A, Tome D, Soenen S, Westerterp KR. Dietary protein, weight loss, and weight maintenance. Annual review of nutrition. 2009;29:21-41.
42. Leidy HJ, Clifton PM, Astrup A, Wycherley TP, Westerterp-Plantenga MS, Luscombe-Marsh ND, et al. The role of protein in weight loss and maintenance. Am J Clin Nutr. 2015.
43. Krieger JW, Sitren HS, Daniels MJ, Langkamp-Henken B. Effects of variation in protein and carbohydrate intake on body mass and composition during energy restriction: a meta-regression 1. AM J CLIN NUTR. 2006;83(2):260-74.
44. Lang PO, Trivalle C, Vogel T, Proust J, Papazian JP. Markers of metabolic and cardiovascular health in adults: Comparative analysis of DEXA-based body composition components and BMI categories. J Cardiol. 2015;65(1):42-9.
45. Atlantis E, Martin SA, Haren MT, Taylor AW, Wittert GA, Members of the Florey Adelaide Male Ageing S. Inverse associations between muscle mass, strength, and the metabolic syndrome. Metabolism. 2009; 58(7):1013-22.
46. Srikanthan P, Karlamangla AS. Relative muscle mass is inversely associated with insulin resistance and prediabetes. Findings from the third National Health and Nutrition Examination Survey. J Clin Endocrinol Metab. 2011;96(9):2898-903.
47. Koletzko B, von Kries R, Monasterolo RC, Subias JE, Scaglioni S, Giovannini M, et al. Infant feeding and later obesity risk. Adv Exp Med Biol. 2009;646:15-29.
48. Grote V, von Kries R, Closa-Monasterolo R, Scaglioni S, Gruszfeld D, Sengier A, et al. Protein intake and growth in the first 24 months of life. J Pediatr Gastroenterol Nutr. 2010;51 Suppl 3:S117-8.
49. Ong K, Kratzsch J, Kiess W, Dunger D. Circulating IGF-I levels in childhood are related to both current body composition and early postnatal growth rate. J Clin Endocrinol Metab. 2002;87(3):1041-4.
50. Michaelsen KF, Larnkjaer A, Molgaard C. Early diet, insulin-like growth factor-1, growth and later obesity. World review of nutrition and dietetics. 2013;106:113-8.
51. Rolland-Cachera MF, Deheeger M, Maillot M, Bellisle F. Early adiposity rebound: causes and consequences for obesity in children and adults. Int J Obes (Lond). 2006;30 Suppl 4:S11-7.
52. Gunther ALB, Buyken AE, Kroke A. The influence of habitual protein intake in early childhood on BMI and age at adiposity rebound: Results from the DONALD Study. Int J Obes. 2006;30(7):1072-9.
53. Druet C, Stettler N, Sharp S, Simmons RK, Cooper C, Smith GD, et al. Prediction of childhood obesity by infancy weight gain: an individual-level meta-analysis. Paediatric and perinatal epidemiology. 2012; 26(1):19-26.
54. Monteiro PO, Victora CG. Rapid growth in infancy and childhood and obesity in later life--a systematic review. Obes Rev. 2005;6(2):143-54.
55. Singh AS, Mulder C, Twisk JW, van Mechelen W, Chinapaw MJ. Tracking of childhood overweight into adulthood: a systematic review of the literature. Obes Rev. 2008;9(5):474-88.

56. Kipnis V, Subar AF, Midthune D, Freedman LS, Ballard-Barbash R, Troiano RP, et al. Structure of dietary measurement error: results of the OPEN biomarker study. Am J Epidemiol. 2003;158(1):14-21; discussion 2-6.
57. Freedman LS, Schatzkin A, Midthune D, Kipnis V. Dealing with dietary measurement error in nutritional cohort studies. J Natl Cancer Inst. 2011;103(14):1086-92.
58. Abrams SA, Hawthorne KM, Pammi M. A systematic review of controlled trials of lower-protein or energy-containing infant formulas for use by healthy full-term infants. Adv Nutr. 2015;6(2):178-88.
59. Voortman T, Vitezova A, Bramer W, Ars C, Bautista P, Buitrago-Lopez A, et al. Effects of protein intake on blood pressure, insulin sensitivity and blood lipids in children: a systematic review. Br J Nutr. 2015; 113(3):383-402.

4.4

Intake of different types of fatty acids in infancy and growth, adiposity, and cardiometabolic health up to 6 years of age: The Generation R Study

Based on:

Stroobant W*, **Braun KVE***, Kiefte-de Jong JC, Moll HA, Jaddoe VWV, Brouwer IA, Franco OH, Voortman T. Intake of Different Types of Fatty Acids in Infancy Is Not Associated with Growth, Adiposity, or Cardiometabolic Health up to 6 Years of Age. *Journal of Nutrition. 2017;147(3):413-20.*

** Denotes equal contribution*

ABSTRACT

Background: Studies in adults indicate that a lower saturated and higher unsaturated fat intake is associated with a lower risk for metabolic syndrome and cardiovascular diseases. However, studies on fat intake in relation to cardiometabolic health during childhood are scarce. We examined associations between dietary intake of fatty acids at age 1 year and measures of growth, adiposity and cardiometabolic health up to age 6 years.

Methods: This study was performed among 2,927 children participating in the Generation R Study, a multiethnic prospective population-based cohort in the Netherlands. We measured children's total fat intake, and intake of saturated (SFA), monounsaturated (MUFA), and polyunsaturated fatty acids (PUFA) at their median age of 12.9 months (95%range 12.2-18.9) using a food-frequency questionnaire. We repeatedly measured their height and weight up to age 6 years. At age 6, we measured their body fat percentage, diastolic and systolic blood pressure, and serum insulin, triacylglycerol and HDL-cholesterol. These outcomes were combined into a cardiometabolic risk factor score. We examined associations of fatty acid intake with repeatedly measured height, weight and BMI using linear mixed models and with cardiometabolic outcomes using linear regression models, adjusting for sociodemographic and lifestyle factors and taking into account macronutrient substitution effects.

Results: In multivariable models, we observed no association of a higher intake of total fat or SFA, MUFA, or PUFA with growth, adiposity, or cardiometabolic health when fat was consumed at the expense of carbohydrates. In subsequent models, also no associations were observed for a higher MUFA or PUFA intake at the expense of SFA with any of outcomes. Results did not differ by sex, ethnicity, age, or birth weight.

Conclusions: The results of this study did not support our hypothesis that intake of different types of fatty acids was associated with adiposity or cardiometabolic health among children up to 6 years of age.

INTRODUCTION

Cardiometabolic diseases, such as cardiovascular disease and type 2 diabetes, are highly prevalent. Risk factors for these diseases include obesity, dyslipidemia, elevated blood pressure, and insulin resistance and a clustering of these risk factors is often referred to as the metabolic syndrome.[1] Several studies have suggested that the development of cardiometabolic risk factors already begins in early life and that these risk factors track during further life.[2-4] Therefore, gaining knowledge about factors that may influence cardiometabolic health among children is very relevant for early prevention of later cardiovascular disease and type 2 diabetes.

Several studies among adults have reported associations for different types of dietary fat intakes with cardiometabolic risk factors. Overall, these studies suggest that a lower saturated fatty acid (SFA) intake, when replaced by a higher polyunsaturated fatty acid (PUFA) or monounsaturated fatty acid (MUFA) intake, may reduce coronary heart disease risk and provide cardiometabolic benefits.[5-7] Suggested underlying mechanisms include for example differences in inflammatory responses, changes in endothelial function, or changes in lipid metabolism in response to different fatty acids.[8] Based on these studies, dietary guidelines generally advise to lower SFA intake and increase PUFA intake.[9]

In contrast to these dietary recommendations for adults, no clear guidelines are available regarding fat intake for young children. Nutritional requirements in this period may be different, however, evidence on the cardiometabolic health effects of fatty acid intake in early life is scarce.[10] A few studies among school-aged children have indicated that the intake of different types of fatty acids may be associated with certain cardiometabolic risk factors,[11-13] but studies on fat intake in early childhood are lacking,[14] whereas early life may be an important period for programming of long-term health.[15] Considering the divergent effects of different types of fatty acids on cardiometabolic outcomes in previous studies, we hypothesized that intake of SFA is negatively associated with cardiometabolic health, and that intake of MUFA and PUFA is positively associated with cardiometabolic health. Therefore, the objective of this study was to examine the association between the intake of different types of fatty acids in early childhood and body composition and cardiometabolic health at 6 years of age, in a prospective cohort study.

METHODS

Study design and subjects

This study was embedded in the Generation R Study, a prospective cohort from fetal life onward in Rotterdam, the Netherlands.[16] All pregnant women living in the urban area of Rotterdam and with an expected delivery date between April 2002 and January 2006 were approached to participate. Response at baseline was 61% (n=9,778 women). After birth, 7,893 children were available for follow-up studies. The study was approved by the Medical Ethical Committee of Erasmus Medical Center in Rotterdam and parents of all children provided written informed consent. To determine children's dietary intake, mothers of 5,088 children received a food-frequency questionnaire (FFQ). After excluding children whose parents did not return the FFQ and children whose FFQ was not completely filled out, complete and valid food intake data was available for 3,629 (72%) children.[17, 18] Of these children, 2,967 visited the research center around the age of 6 years and had data available about at least one cardiometabolic outcome (**Figure 4.4.1**).

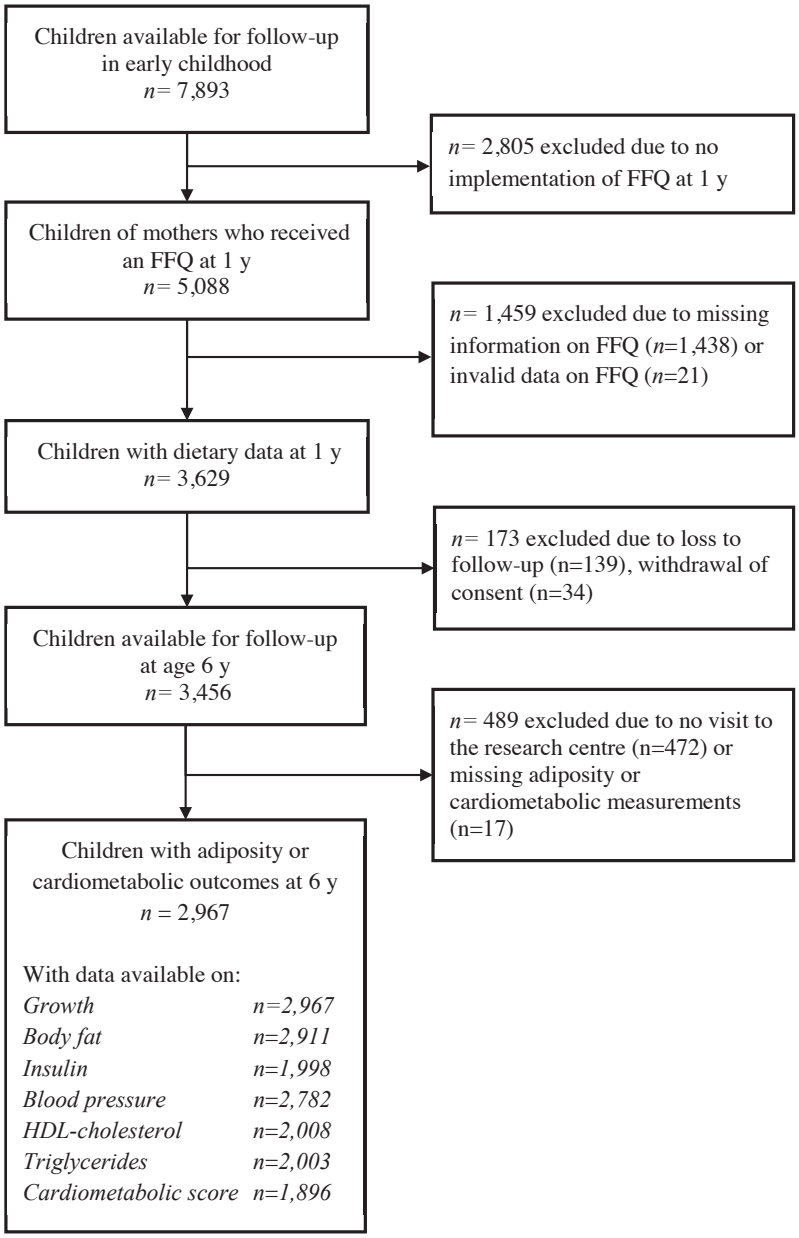


Figure 4.4.1. Flowchart of participants included for analysis

Dietary intake assessment

Food intake was assessed between 2003 and 2006 when the children had a median age of 12.9 months (95% range: 12.2-18.9), using a semi-quantitative FFQ. This questionnaire was completed by the parents or caregivers of the children. The FFQ was specifically designed for this age group and contained

211 food items.[17] The FFQ was validated against three 24-hour recalls in 32 children aged 14 months living in Rotterdam and interclass correlation coefficients for nutrient intakes ranged from 0.4 to 0.7.[17,18] With the Dutch Food Composition Database (2006), we calculated intake of total energy, total fat, SFA, MUFA, PUFA, *n*-3 PUFA, and *n*-6 PUFA.[19] Intakes of different types of fat were calculated as a percentage of total energy intake. Information on content of trans fat was unfortunately not available, but was assumed to be low.[20,21] To be able to apply macronutrient substitution models, we also calculated intake of other protein and carbohydrates, and we calculated energy from total fat minus that from SFAs, MUFAs and PUFAs (i.e., from glycerol, trans fatty acids, sterols, and phospholipids[22]), which amounted to approximately 4% of total energy intake in our study population.

Assessment of anthropometrics, adiposity and cardiometabolic health

Data on height, weight, and BMI were collected at eight different time points between the ages of 1 and 6 years. Children's height and weight up to the age of 4 years were measured during routine visits to Child Health Centers at median (95% range) ages of 11 (10-13), 14 (13-16), 19 (17-21), 24 (23-28), 30 (29-34), 36 (35-40), and 45 (44-48) months. At their median age of 6.0 years (95% range: 5.7-6.6), we measured the children's height, weight, body composition, and cardiometabolic health factors in our research center at the Erasmus Medical Center.[16] Height was determined in standing position to the nearest millimeter with a Harpenden stadiometer (Holtain Limited, Dyfed, U.K.). Weight was measured using a mechanical personal scale (SECA, Almere, the Netherlands) to the nearest 0.1 kg. For each time point, BMI was calculated as total body weight (kg) divided by height $(m)^2$. Overweight was determined according to the Cole criteria.[23]

Body fat mass was measured with a dual-energy x-ray absorptiometry (DXA) scanner (iDXA, General Electrics –Lunar, 2008, Madison, WI, USA) and was analyzed with Encore software (version 13.6).[24] Body fat percentage (BF%) was calculated by expressing total fat mass as percentage of total body weight; fat mass index (FMI) was calculated as fat mass (kg) divided by height $(m)^2$ and fat-free mass index (FFMI) as fat-free mass (kg) divided by height $(m)^2$. Blood pressure was measured four times at the right brachial artery using a validated automatic sphygmomanometer (Datascope Accutorr Plus™, Paramus, USA). The mean of the last three measurements was calculated and used for analyses. Non-fasting blood samples were obtained and concentrations of insulin, triacylglycerol (TAG), total cholesterol, HDL-cholesterol, LDL-cholesterol, and C-peptide were determined with enzymatic methods (using a Cobas 8000 analyzer, Roche, Almere, the Netherlands).[16] Quality control samples demonstrated intra-assay and inter-assay coefficients of variation ranging from 0.77% to 1.69%. For all body composition and cardiometabolic outcomes, age- and sex- specific SD-scores were calculated on the basis of our study population.

We applied a cardiometabolic risk factor score, modelled following metabolic syndrome-like definitions used in adults and adapted for use in pediatric population.[25,26] As described in more detail elsewhere, this score consisted of the sum of SD-scores from BF%, SBP, DBP, insulin concentrations, TAG concentrations, and the inverse of HDL-cholesterol concentrations,[26] with a higher cardiometabolic risk factor score reflecting a higher cardiometabolic risk. For our analyses, we chose the cardiometabolic risk factor score and its components as primary outcomes of interest, but we additionally examined the other cardiometabolic outcomes that were available in our study (i.e., total cholesterol, LDL-cholesterol, and C-peptide concentrations).

Covariates

Information on maternal age, parental educational level, net household income, folic acid supplement use during pregnancy, and maternal parity was collected via self-administered questionnaires at enrollment in the study. Parental highest finished educational levels were classified into no or primary education, middle school or less than 4 years of high school, or higher education. Net household income was categorized into less than 1400, 1400 to 2200, or more than 2200 euros per month. Parity was categorized into 0, 1, or 2 or more. Information on smoking and alcohol consumption during pregnancy was obtained through questionnaires in each trimester and both variables were categorized into never, till pregnancy was known, or continued during pregnancy. Maternal food intake during first trimester was assessed with an FFQ and overall diet quality was assessed with a predefined quality score.[27]

Information on children's sex, birth weight, and gestational age was available from obstetric records.[16] Sex- and gestational age-specific z-scores for birth weight were determined according to reference data.[28] Child's ethnic background was defined as Dutch or non-Dutch on the basis of the countries of birth of the parents.[18] Information on breastfeeding was available from postnatal questionnaires and was categorized into never, partially, or exclusively for at least 4 months. Information on timing of introduction of solid foods was available from a questionnaire. A previously developed diet score was applied to assess overall diet quality of the children at the age of 1 year, based on adherence to dietary guidelines for preschool children. Briefly, the ratio of reported and recommended intake for ten food groups was calculated, each truncated at 1. The scores of the 10 diet score components were summed, resulting in an overall score ranging from 0 to 10 on a continuous scale with a higher score representing a better diet quality.[18] Screen time (watching television or using a computer, hours/day), and participation in sports (yes or no) were assessed with questionnaires when the children were 6 years old.

Statistical analysis

Associations between the intake of fatty acids and repeatedly measured height, weight and BMI were analyzed with linear mixed models. Associations of dietary fatty acid intakes with adiposity and cardiometabolic outcomes at the age of 6 years were analyzed with linear regression analyses. To test whether exposure variables had a linear relation with the outcomes we applied natural cubic splines with 3 degrees of freedom.[29] We found indications for a non-linear association only for the relation between PUFA intake and the cardiometabolic risk score ($P<0.05$). Therefore, we also examined the associations of quartiles of PUFA intake with cardiometabolic health.

To account for the effect of energy intake and substitution effects of macronutrients in the diet, two different multivariable nutrient density substitution models were used.[30] We chose to enter all nutrients in the model per 5 energy percent (E%) in line with previous studies.[6, 31] Firstly, we examined the association of fat intake at the expense of carbohydrates with cardiometabolic outcomes. In order to study this, we included total energy, protein and total fat intake; or total energy, protein, SFA, MUFA, and PUFA intake in the same model. Because total carbohydrate intake is the only macronutrient left out of this model, coefficients for fat intake can be interpreted as a 5 E% higher intake of fat the expense of carbohydrate. Secondly, we examined the association of MUFA and PUFA intake at the expense of SFA with cardiometabolic health. In these models, total energy intake and

energy from carbohydrate, protein, total fat minus fatty acids, MUFA, and PUFAs were included in one model, i.e., all energy sources except for SFA.

In addition to the nutrients, we included child's sex, ethnicity, and age at filling out the FFQ in our basic models. In this model sex, ethnicity, age at filing out the FFQ, and birth weight were also tested as possible effect modifiers, but none of the interaction terms were statistically significant (all $P>0.05$). Covariates were selected based on theory and previous studies and were included in the multivariable model when adding them to the basic model resulted in changes in effect estimates larger than 10%. Based on this criterion, multivariable models were additionally adjusted for the following variables: maternal BMI, age, alcohol consumption during pregnancy, smoking during pregnancy, education, and parity, household income, birth weight Z-score, breastfeeding, and children's screen time and playing sports. The following covariates were considered but not included because they did not meet the 10%-change criterion: folic acid supplement use during pregnancy, maternal diet quality score, paternal educational level, timing of introduction of solid foods, and child's diet quality score. As sensitivity analysis, we repeated our models among children with a Dutch ethnic background only, because the FFQ was originally designed for Dutch children. Furthermore, we performed sensitivity analyses in which we replaced protein intake by carbohydrate intake in the models, in order to examine fat intake at the expense of protein. Finally, to assess if potential associations with the cardiometabolic risk factor score were driven by a single cardiometabolic risk factor, we performed sensitivity analyses in which we excluded each cardiometabolic outcome from the cardiometabolic score one by one.

To increase statistical power and to reduce bias due to missing data, missing values of covariates (ranging from 0 to 19%) were multiple imputed (10 imputations) according to the Fully Conditional Specification method using predictive mean matching.[32] We report the pooled regression coefficients after the multiple imputation procedure. Statistical analyses were performed using SPSS version 21.0 (IBM Corp., Armonk, NY, USA) and R version 3.1.2 (R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

Subject characteristics

Maternal characteristics, child characteristics, dietary intake and cardiometabolic outcomes are presented in **Table 4.4.1**. Most children had a Dutch ethnic background (68.8 %) and had mothers with a high educational level (59.4%). Children's mean (± SD) total energy intake was 1,310 (±388) kcal per day and 28.5 (±5.6) E% was derived from fat. The mean intake of SFAs in our population was 11.5 (±3.6) E%, of MUFAs 9.8 (±3.1) E%, and of PUFAs 5.5 (±1.5) E%. The main food sources contributing to SFA intake in our study population were sauces, solid cooking fats, and cakes and pastry. Main food sources of MUFA intake were sauces, and cooking fats and oils; and main sources of PUFA intake were margarines, liquid cooking fats and oils, and mayonnaise. At the age of 6 years, 404 (13.6%) of the children in this study population were overweight or obese. In nonresponse analyses, we observed that children without FFQ were slightly taller and heavier at follow-up than children with FFQ data.[33]

Table 4.4.1. Characteristics of the mothers and children included in the analyses

	Population characteristics[1] (*n*=2,967)
Maternal characteristics	
Maternal age (years)	31.5 ± 4.5
Maternal BMI at intake (kg/m^2)	23.5 (18.8-35.3)
Nulliparous	60.5%
Education level	
- No or primary education	5.5%
- Middle school or < 4 years of high school	36.0%
- Higher education	58.6%
Monthly household income	
- < € 1400	13.9%
- € 1400 – € 2200	19.4%
- > € 2200	66.6%
Smoking during pregnancy	
- Never	77.3%
- Until pregnancy was known	10.7%
- Continued	11.9%
Alcohol consumption during pregnancy	
- Never	37.7%
- Until pregnancy was known	16.4%
- Continued	45.9%
Infant characteristics	
Sex (% boys)	48.7%
Dutch ethnic background	68.6%
Gestational age at birth (weeks)	40.1 (35.6-42.3)
Birth weight (g)	3,452 ± 569
Breastfeeding	
- Exclusive the first 4 months	29.8%
- Partial the first 4 months	59.9%
- Never	10.3%
Child characteristics at dietary assessment	
Age at FFQ (months)	12.9 (12.2-18.9)
Total energy intake (kcal)	1,310 ± 388
Total protein intake (E%)	12.9 ± 2.4
Total carbohydrate intake (E%)	58.4 ± 5.9
Total fat intake (E%)	28.5 ± 5.6
- SFA intake (E%)	11.5 ± 3.6
- MUFA intake (E%)	9.8 ± 3.1
- PUFA intake (E%)	5.5 ± 1.5

Table 4.4.1. Characteristics of the mothers and children included in the analyses (continued)

	Population characteristics[1] (*n*=2,967)
Child characteristics at follow-up visit	
Age (years)	6.0 (5.7-6.6)
Height (cm)	118 ± 5
Weight (kg)	21.8 (17.4-30.4)
BMI (kg/m^2)	16.0 ± 1.6
Screen time (h/d)	1.25 (0.25-4.71)
Participation in sports	44.9%
Body fat percentage (*n*=2911)	24.2 ± 5.0
Systolic blood pressure (mmHg) (*n*=2843)	102 ± 8
Diastolic blood pressure (mmHg) (*n*=2843)	60 ± 7
HDL-cholesterol (mmol/L) (*n*=2008)	1.34 ± 0.31
Insulin (pmol/L) (*n*=1998)	115 (18-397)
Triacylglycerol (mmol/L) (*n*=2003)	0.97 (0.40-2.36)

[1]Values are given as mean ± SD when continuous variables were normally distributed, as median (95% range) for continuous variables with a skewed distribution, and as percentages for categorical variables.

Table 4.4.2. Associations between fat intake at the expense of carbohydrates at age 1 year and growth between the ages of 1 and 6 years.[1,2]

	Height (SDS) *n*=2967	Weight (SDS) *n*=2967	BMI (SDS) *n*=2967
Total fat intake (5E%)	-0.023 (-0.056;0.010)	-0.027 (-0.059;0.007)	-0.003 (-0.030;0.024)
SFA intake (5E%)	0.088 (-0.049;0.225)	0.074 (-0.062;0.211)	0.049 (-0.062;0.159)
MUFA intake (5E%)	-0.139 (-0.314;0.035)	-0.102 (-0.275;0.072)	-0.016 (-0.157;0.125)
PUFA intake (5E%)	-0.006 (-0.163;0.151)	-0.082 (-0.238;0.074)	-0.094 (-0.221;0.033)

[1]Values are based on multivariable linear mixed models and reflect differences (95% CI) of the whole trajectory of the anthropometric outcome between the ages of 1 and 6 years (age and sex adjusted SD-scores) per 5 E% higher intake of a specific fat intake, at the expense of carbohydrates.
[2]Models were adjusted for intake of total energy and energy from protein and all types of fats (i.e., all energy sources except for carbohydrates) to examine macronutrient substitution effects. Models were additionally adjusted for age at FFQ, sex, ethnicity, maternal education, household income, parity, maternal BMI, maternal age, smoking during pregnancy, drinking alcohol during pregnancy, breastfeeding, birth weight, screen time and playing sports at 6 years of age.
Abbreviations: SDS, Standard deviation score; SFA, saturated fatty acids; MUFA, monounsaturated fatty acids; PUFA, polyunsaturated fatty acids; E, energy.

Table 4.4.3. Associations between fat intake at the expense of carbohydrates at ae 1 year and cardiometabolic health at age 6 years.[1,2]

	BF% (SDS) *n*=2911	SBP (SDS) *n*=2843	DBP (SDS) *n*=2843	HDL (SDS) *n*=2008	Insulin (SDS) *n*=1998	Triacylglycerol (SDS) *n*=2003	Cardiometabolic risk factor score *n*=1896
Total fat intake (5E%)	-0.014 (-0.043;0.014)	0.020 (-0.014;0.055)	-0.006 (-0.040;0.028)	0.025 (-0.017;0.066)	-0.009 (-0.051;0.032)	-0.024 (-0.066;0.018)	-0.076 (-0.181;0.030)
SFA intake (5E%)	-0.028 (-0.142;0.086)	0.074 (-0.065;0.214)	0.073 (-0.063;0.209)	0.043 (-0.119;0.205)	-0.158 (-0.321;0.005)	-0.073 (-0.236;0.089)	-0.213 (-0.624;0.199)
MUFA intake (5E%)	0.020 (-0.125;0.164)	-0.065 (-0.241;0.110)	-0.132 (-0.303;0.039)	-0.010 (-0.217;0.197)	0.163 (-0.045;0.371)	0.028 (-0.179;0.235)	0.040 (-0.483;0.564)
PUFA intake (5E%)	-0.079 (-0.213;0.054)	0.109 (-0.053;0.272)	0.121 (-0.037;0.272)	0.053 (-0.138;0.245)	-0.065 (-0.258;0.127)	-0.038 (-0.230;0.154)	-0.017 (-0.503;0.470)

[1]Values are based on multivariable linear regression models and reflect differences (95% CI) in cardiometabolic outcomes (age and sex adjusted SD-scores) and cardiometabolic risk factor score per 5 E% higher intake of a specific fat, at the expense of carbohydrates.

[2]Models were adjusted for intake of total energy, energy from protein, and energy from all types of fats (i.e., all energy sources except for carbohydrates) to examine macronutrient substitution effects. Models were additionally adjusted for age at FFQ, sex, ethnicity, maternal education, household income, parity, maternal BMI, maternal age, smoking during pregnancy, drinking alcohol during pregnancy, breastfeeding, birth weight, screen time and playing sports at 6 years of age.

Abbreviations: SDS, Standard deviation score; SFA, saturated fatty acids; MUFA, monounsaturated fatty acids; PUFA, polyunsaturated fatty acids; BF, body fat; SBP, systolic blood pressure; DBP, diastolic blood pressure; HDL, high-density-lipoprotein; E%, energy percent.

Associations of fat intake with growth, adiposity and cardiometabolic health

The covariate-adjusted associations for the carbohydrate substitution models are presented in **Table 4.4.2** and **Table 4.4.3**. A higher total fat intake at the expense of a lower carbohydrate intake in early childhood was not associated with any of the anthropometric or cardiometabolic outcomes at the ages of 6 years. Similarly, higher intakes of either SFAs, MUFAs, or PUFAs at the expense of carbohydrates were not associated with growth (**Table 4.4.2**) or cardiometabolic health (**Table 4.4.3**). For analyses with PUFA intake in quartiles also no significant associations or clear non-linear trends were observed (data not shown).

Substitution models in which we examined the substitution of SFAs by unsaturated fat showed similar results as observed for carbohydrate substitution models: neither MUFA nor PUFA intake were associated with any of the anthropometric (data not shown) or cardiometabolic outcomes (**Table 4.4.4**) when consumed at the expense of SFAs.

Associations with cardiometabolic health unadjusted for confounders are shown in **Supplementary Table 4.4.1**. In these models, a higher SFA intake was associated with a higher SBP and lower insulin concentrations; and a higher MUFA intake was associated with higher insulin levels when consumed that the expense of carbohydrates. Other results were similar to those of the adjusted model.

Additional analysis

Sensitivity analyses restricted to children with a Dutch ethnic background showed similar associations between fat intakes, either at the expense of carbohydrates or of SFAs, with cardiometabolic health as observed for the whole study population (**Supplementary Table 4.4.2**). Intake of fatty acids was also not associated with other measures of adiposity (FMI or FFMI) or other cardiometabolic health outcomes that were not included in the cardiometabolic score (i.e., total cholesterol, LDL-cholesterol, or C-peptide concentrations) (**Supplementary Table 4.4.3**).

When we explored substitution of protein instead of carbohydrates we found that a higher total fat, PUFA, or MUFA intake at the expense of protein was associated with a lower BF%, but not with any of the other outcomes (data not shown). This is in line with results for protein intake, either at the expense of fat or carbohydrates, which we described previously.[26, 34] Finally, in line with the main analysis, in sensitivity analyses in which we excluded each component from the cardiometabolic score one by one no significant associations were observed.

DISCUSSION

The aim of this study was to examine the associations between intake of total, saturated, monounsaturated and polyunsaturated fat in children in early childhood and their growth and cardiometabolic health up to the age of 6 years. Overall, our results show no consistent association between intake of total fat or different types of fatty acids, at the expense of either carbohydrates or saturated fat with growth or cardiometabolic outcomes.

One large intervention study in children focused on similar outcome variables as our study. This Special Turku coronary Risk factor Intervention Project (STRIP) study investigated the influence

Table 4.4.4. Associations for unsaturated fat intake at the expense of saturated fat at age 1 year and cardiometabolic health at age 6 years.[1,2]

	BF% (SDS) *n*=2911	SBP (SDS) *n*=2843	DBP (SDS) *n*=2843	HDL (SDS) *n*=2008	Insulin (SDS) *n*=1998	Triacylglycerol (SDS) *n*=2003	Cardiometabolic risk factor score *n*=1896
MUFA intake (5E%)	0.048 (-0.202;0.298)	-0.139 (-0.443;0.164)	-0.205 (-0.501;0.091)	-0.053 (-0.408; 0.302)	0.321 (-0.036; 0.678)	0.102 (-0.253;0.457)	0.253 (-0.647; 1.153)
PUFA intake (5E%)	-0.051 (-0.182;0.079)	0.035 (-0.124;0.194)	0.048 (-0.107;0.203)	0.010 (-0.178;0.197)	0.093 (-0.096;0.281)	0.035 (-0.152;0.223)	0.196 (-0.283;0.675)

*[1]Values are based on multivariable linear regression models and reflect differences (95% CI) in cardiometabolic outcomes (age and sex adjusted SD-scores) and cardiometabolic risk factor score per 5 E% or per 1 E% increase of monounsaturated and polyunsaturated fatty acid intake at the expense of saturated fatty acids. * P<0.05*

[2]Models were adjusted for intake of total energy and energy from protein, energy carbohydrates and energy from all fats except for SFA to examine macronutrient substitution effects. Models were additionally adjusted for age, sex, ethnicity, maternal education, household income, parity, maternal BMI, maternal age, smoking during pregnancy, drinking alcohol during pregnancy, breastfeeding, birth weight, screen time and playing sports at 6 years of age.

Abbreviations: SDS, Standard deviation score; MUFA, monounsaturated fatty acids; PUFA, polyunsaturated fatty acids; BF, body fat; SBP, systolic blood pressure; DBP, diastolic blood pressure; HDL, high-density-lipoprotein; E%, energy percent.

of low-saturated fat counseling versus no dietary counseling on cardiometabolic health in over one thousand children from 7 months of age onwards in Finland. Outcomes that were examined included metabolic syndrome, blood pressure, insulin sensitivity and cholesterol levels, and in several follow-up studies in later childhood or adolescence a beneficial effect of the intervention was observed for many of these cardiometabolic outcomes.[13, 35-37] In contrast to the findings of the STRIP study, in our study we did not find associations between lower SFA intake and a better cardiometabolic health. However, because the intervention in the STRIP study consisted of dietary counseling it is not certain whether the effects were caused by a low-saturated fat diet or other effects of the long-term lifestyle advice. For example, in one of the analyses in the STRIP study it was shown that the intervention group had a lower HOMA-IR, but that actual SFA intake was not significantly associated with HOMA-IR in multivariable analyses.[37] Additionally, dietary counseling in the STRIP study remained until 20 years of age and observed effects may therefore also be caused by dietary changes in later childhood rather than in early childhood.

A few observational studies focused on dietary fat intake in young children in relation to their weight or BMI. In previous analyses in our study population, Heppe et al. (2013) observed that a higher PUFA intake at the age of 1 year was associated with a lower BMI at 4 years of age – an earlier follow-up measurement in our cohort.[38] Considering that we did not find this association with BMI up to the age of 6 years might suggest that the potential effects of fat intake fade away after a longer follow-up period. Another possible explanation for our null-findings with BF% may be that the adiposity rebound, that occurs around the age of 6 years,[39] have obscured a possible inverse association between PUFA intake and body fat at this age specifically. In line with our results, Williams et al. (2008), who performed a prospective study in 519 children, did not find any significant associations between SFA or MUFA intake at 3-4 years of age and BMI at 7-10 years.[40] Also Agostoni et al. (2000), who measured dietary intake at 1 and 5 years and BMI at 5 years in 147 children, observed that total fat, SFA, MUFA or PUFA intake at neither 1 year nor at 5 years were associated with BMI at 5 years of age.[41] Also in line with our findings, PUFA supplementation in infancy has been shown not to influence growth.[42] No previous studies were identified that examined fat intake in relation to measures of body fat in young children.

Two previous studies investigated the association between fatty acid intake with cholesterol and TAG concentrations in children. Contrary to our findings, Cowin et al. (2000) observed that a higher intake of PUFAs or SFAs at the age of 18 months was associated with a lower HDL-cholesterol concentration at the age of 31 months.[43] However, this association was only present among girls and not among boys. Moreover, among boys, but not girls, a higher total fat intake or SFA intake was associated with a higher total cholesterol concentration.[43] In our population, we found no significant interaction between child sex and fatty acid intake on cardiometabolic health. In line with our results, the study of Williams et al. (2008) did not find any association between SFA or MUFA intake at 3-4 years of age and total cholesterol, HDL-cholesterol and TAG concentrations at 7-10 years.[40] However, cross-sectional data from the same study population showed that intake of MUFAs, but not SFAs, in later childhood were inversely associated with total cholesterol and TAG concentrations.[40]

In contrast to our hypothesis, in our study no associations were found for lower SFA intake and higher MUFA and PUFA intake and better cardiometabolic health. This is in line with a review and meta-analysis of studies among adults by Chowdhury et al., who also reported no consistent effects

of dietary or circulating SFAs, MUFAs, or PUFAs on cardiovascular disease.[44] A possible explanation for the fact that we did not find associations might be that different individual fatty acids within the subgroups of SFA, MUFA and PUFA, may have different effects on cardiometabolic health. Unfortunately, we could not reliably calculate intake of specific fatty acids from our FFQ data. We also did not have information available on trans fatty acids, but intake was assumed to be less than 1% of total energy intake based on national food consumption data in the years of our dietary assessment.[20, 21] Although many food items were included in the FFQ, including many detailed subtypes of oils and margarines, the intraclass correlation coefficient for total fat intake in a validation study was only 0.4. This indicates measurement errors in our assessment of fat intake in early childhood. However, dietary fat intake remains difficult to estimate and our reference method with 24-hour recalls is also not an optimal way to measure this.[45] Therefore, it is difficult to judge the quality of the estimates of fat intake in our study and whether our exposure measurement was specific enough to detect a potential association of fatty acid intake with growth or cardiometabolic health. Future studies should use more detailed measurements of fat intake, such as weighed food records or repeated 24-hour recalls, or should combine dietary intake data with fatty acid concentrations in for example blood or adipose tissue. Especially serum or erythrocyte concentrations of certain *n*-3 or *n*-6 PUFAs may provide good estimates for dietary intake of fatty acids.[46, 47]

An important strength compared with other studies is that we accounted for the effect of energy intake and macronutrient substitution by using a multivariable nutrient density substitution model. By using different substitution models, we could examine the effects of replacing different macronutrients by fatty acids. Finally, a strength of our study is the large number of variables measured in this cohort, which made it possible to examine many potential confounders and effect modifiers. However, a limitation was that no information was available about the children's cardiometabolic health earlier in childhood or their dietary intake at the age of 6 years. Therefore, it was not possible to perform longitudinal analyses or to examine whether fatty acid intake at the age of 6 years was associated with cardiometabolic health or diluted a potential effect of fat intake in early childhood.

CONCLUSION

Results from this prospective cohort study show no consistent associations between intakes of total, saturated, monounsaturated, or polyunsaturated fat in early childhood with growth or cardiometabolic health outcomes at school age. Because the number of studies examining these associations is scarce, future studies are needed to further investigate the associations between intake of different types of fat in early childhood and adiposity and cardiometabolic health, preferably with the use of more detailed dietary assessment methods and combined with fatty acid biomarkers.

SUPPLEMENTARY MATERIAL

Supplementary Table 4.4.1. Minimally adjusted associations between fat intake, at the expense of carbohydrates, at age 1 year and cardiometabolic health at 6 years of age.

Fat intake, per 5E%	BF% (SDS) *n=2911*	SBP (SDS) *n=2843*	DBP (SDS) *n=2843*	HDL (SDS) *n=2008*	Insulin (SDS) *n=1998*	Triacylglycerol (SDS) *n=2003*	Cardiometabolic risk factor score *n=1896*
Total fat intake	-0.007 (-0.036;0.022)	0.019 (-0.015;0.053)	-0.002 (-0.035;0.030)	0.026 (-0.014;0.066)	-0.015 (-0.055;0.026)	-0.024 (-0.064;0.016)	-0.073 (-0.176;0.030)
SFA intake	-0.073 (-0.196;0.050)	0.074 (0.000;0.148)*	0.079 (-0.049;0.207)	0.048 (-0.035;0.131)	-0.165 (-0.249;-0.081)*	-0.035 (-0.119;0.049)	-0.281 (-0.497;0.066)
MUFA intake	0.053 (-0.099;0.206)	-0.081 (-0.256;0.093)	-0.138 (-0.308;0.032)	-0.027 (-0.130;0.077)	0.180 (0.076;0.285)*	0.011 (-0.093;0.115)	0.132 (-0.406;0.670)
PUFA intake	-0.050 (-0.189;0.090)	0.137 (-0.023;0.298)	0.149 (-0.006;0.305)	0.041 (-0.120;0.202)	-0.071 (-0.1760;0.035)	-0.043 (-0.140;0.053)	-0.008 (-0.489; 0.473)

*Values are based on linear regression models and reflect differences (95% CI) in cardiometabolic outcomes (age and sex adjusted SD-scores) and cardiometabolic risk factor score per 5 E% increase of a specific fat intake, at the expense of carbohydrates. * p <0.05*

Models were adjusted for intake of total energy and energy from protein and all types of fats (i.e., all energy sources except for carbohydrates) to examine macronutrient substitution effects. Models were additionally adjusted for child's age at FFQ, sex and ethnicity.

Abbreviations: SDS, Standard deviation score; SFA, saturated fatty acids; MUFA, monounsaturated fatty acids; PUFA, polyunsaturated fatty acids; BF, body fat; SBP, systolic blood pressure; DBP, diastolic blood pressure; HDL, high-density-lipoprotein; E, energy.

Supplementary Table 4.4.2. Adjusted associations between fat intake, at the expense of carbohydrates, at age 1 and cardiometabolic health at age 6 years in Dutch children only.

Fat intake, per 5E%	BF% (SDS) *n=2911*	SBP (SDS) *n=2843*	DBP (SDS) *n=2843*	HDL (SDS) *n=2008*	Insulin (SDS) *n=1998*	Triacylglycerol (SDS) *n=2003*	Cardiometabolic risk factor score *n=1896*
Total fat intake	-0.024 (-0.058;0.009)	0.019 (-0.024;0.063)	-0.020 (-0.063;0.022)	0.011 (-0.042;0.064)	0.017 (-0.035;0.069)	-0.041 (-0.094;0.012)	-0.064 (-0.193;0.064)
SFA intake	-0.073 (-0.211;0.064)	0.067 (-0.115;0.248)	0.036 (-0.140;0.211)	0.077 (-0.138;0.292)	-0.125 (-0.337;0.086)	-0.031 (-0.246;0.184)	-0.273 (-0.800;0.254)
MUFA intake	0.019 (-0.150;0.189)	-0.065 (-0.290;0.159)	-0.136 (-0.353;0.080)	-0.093 (-0.362;0.176)	0.241 (-0.024;0.506)	-0.054 (-0.324;0.216)	0.182 (-0.475;0.840)
PUFA intake	-0.072 (-0.231;0.088)	0.096 (-0.112;0.304)	0.116 (-0.084;0.317)	0.131 (-0.122;0.384)	-0.171 (-0.420;0.078)	-0.055 (-0.309;0.199)	-0.321 (-0.945;0.303)

Values are based on multivariable linear regression models and reflect differences (95% CI) in cardiometabolic outcomes (age and sex adjusted SD-scores) and cardiometabolic risk factor score per 5 E% higher intake of a specific fat intake, at the expense of carbohydrates.
Models were adjusted for intake of total energy and energy from protein and all types of fats (i.e., all energy sources except for carbohydrates) to examine macronutrient substitution effects. Models were additionally adjusted for age at FFQ, sex, ethnicity, maternal education, household income, parity, maternal BMI, maternal age, smoking during pregnancy, drinking alcohol during pregnancy, breastfeeding, birth weight, screen time and playing sports at 6 years of age.
Abbreviations: SDS, Standard deviation score; SFA, saturated fatty acids; MUFA, monounsaturated fatty acids; PUFA, polyunsaturated fatty acids; BF, body fat; SBP, systolic blood pressure; DBP, diastolic blood pressure; HDL, high-density-lipoprotein; E, energy.

Supplementary Table 4.4.3. Adjusted associations between fat intake, at the expense of carbohydrates, at the age of 1 year with fat mass index, fat-free mass index, LDL cholesterol, total cholesterol, and C-peptide concentrations at 6 years of age.

Fat intake, per 5E%	Total-C (SDS) *n=1947*	LDL-C (SDS) *n=1948*	C-peptide (SDS) *n=1939*	FMI (SDS) *n=2911*	FFMI (SDS) *n=2911*
Total fat intake	-0.007 (-0.048;0.034)	-0.006 (-0.047;0.036)	0.006 (-0.035;0.048)	-0.014 (-0.041;0.013)	-0.016 (-0.047;0.015)
SFA intake	-0.092 (-0.258;0.075)	-0.085 (-0.252;0.082)	-0.145 (-0.312;0.021)	-0.023 (-0.032;0.065)	0.017 (-0.108;0.143)
MUFA intake	0.058 (-0.149;0.265)	0.074 (-0.134;0.282)	0.198 (-0.008;0.405)	0.025 (-0.131;0.085)	-0.010 (-0.169;0.148)
PUFA intake	-0.005 (-0.198;0.187)	-0.039 (-0.232;0.154)	-0.033 (-0.226;0.159)	-0.105 (-0.112;0.162)	-0.114 (-0.260;0.033)

Values are based on multivariable linear regression models and reflect differences (95% CI) in cardiometabolic outcomes (age and sex adjusted SD-scores) per 5 E% higher intake of a specific fat intake, at the expense of carbohydrates.

Models were adjusted for intake of total energy and energy from protein and all types of fats (i.e., all energy sources except for carbohydrates) to examine macronutrient substitution effects. Models were additionally adjusted for age at FFQ, sex, ethnicity, maternal education, household income, parity, maternal BMI, maternal age, smoking during pregnancy, drinking alcohol during pregnancy, breastfeeding, birth weight, screen time and playing sports at 6 years of age.

Abbreviations: E, energy; SDS, standard deviation score; LDL, low density lipoprotein; FMI, fat mass index; fat-free mass index; SFA, saturated fatty acids; MUFA, monounsaturated fatty acids; PUFA, polyunsaturated fatty acids.

REFERENCES

1. Grundy SM, Brewer HB, Jr., Cleeman JI, Smith SC, Jr., Lenfant C, National Heart L, et al. Definition of metabolic syndrome: report of the National Heart, Lung, and Blood Institute/American Heart Association conference on scientific issues related to definition. Arteriosclerosis, thrombosis, and vascular biology. 2004;24(2):e13-8.
2. Morrison JA, Glueck CJ, Woo JG, Wang P. Risk factors for cardiovascular disease and type 2 diabetes retained from childhood to adulthood predict adult outcomes: the Princeton LRC Follow-up Study. International journal of pediatric endocrinology. 2012;2012(1):6.
3. Katzmarzyk PT, Perusse L, Malina RM, Bergeron J, Despres JP, Bouchard C. Stability of indicators of the metabolic syndrome from childhood and adolescence to young adulthood: the Quebec Family Study. Journal of clinical epidemiology. 2001;54(2):190-5.
4. Morrison JA, Friedman LA, Gray-McGuire C. Metabolic syndrome in childhood predicts adult cardiovascular disease 25 years later: the Princeton Lipid Research Clinics Follow-up Study. Pediatrics. 2007; 120(2):340-5.
5. Micha R, Mozaffarian D. Saturated fat and cardiometabolic risk factors, coronary heart disease, stroke, and diabetes: a fresh look at the evidence. Lipids. 2010;45(10):893-905.
6. Li Y, Hruby A, Bernstein AM, Ley SH, Wang DD, Chiuve SE, et al. Saturated Fats Compared With Unsaturated Fats and Sources of Carbohydrates in Relation to Risk of Coronary Heart Disease: A Prospective Cohort Study. J Am Coll Cardiol. 2015;66(14):1538-48.
7. Jakobsen MU, O'Reilly EJ, Heitmann BL, Pereira MA, Balter K, Fraser GE, et al. Major types of dietary fat and risk of coronary heart disease: a pooled analysis of 11 cohort studies. Am J Clin Nutr. 2009;89(5): 1425-32.
8. Endo J, Arita M. Cardioprotective mechanism of omega-3 polyunsaturated fatty acids. Journal of cardiology. 2016;67(1):22-7.
9. Food and Agriculture Organization of the United Nations (FAO). Fats and fatty acids in human nutrition. Report of an expert consultation. Geneva: FAO, 2010 Contract No.: FAO Food and Nutrition paper 91.
10. Aranceta J, Perez-Rodrigo C. Recommended dietary reference intakes, nutritional goals and dietary guidelines for fat and fatty acids: a systematic review. The British journal of nutrition. 2012;107 Suppl 2: S8-22.
11. Simons-Morton DG, Hunsberger SA, Van Horn L, Barton BA, Robson AM, McMahon RP, et al. Nutrient intake and blood pressure in the Dietary Intervention Study in Children. Hypertension. 1997;29(4): 930-6.
12. Carlson SE, DeVoe PW, Barness LA. Effect of infant diets with different polyunsaturated to saturated fat ratios on circulating high-density lipoproteins. Journal of pediatric gastroenterology and nutrition. 1982;1(3):303-9.
13. Niinikoski H, Lagstrom H, Jokinen E, Siltala M, Ronnemaa T, Viikari J, et al. Impact of repeated dietary counseling between infancy and 14 years of age on dietary intakes and serum lipids and lipoproteins: the STRIP study. Circulation. 2007;116(9):1032-40.
14. Braun KV, Voortman T, Kiefte-de Jong JC, Jaddoe VW, Hofman A, Franco OH, et al. Dietary Intakes of Folic Acid and Methionine in Early Childhood Are Associated with Body Composition at School Age. J Nutr. 2015;145(9):2123-9.
15. Innis SM. Metabolic programming of long-term outcomes due to fatty acid nutrition in early life. Matern Child Nutr. 2011;7 Suppl 2:112-23.

16. Kruithof CJ, Kooijman MN, van Duijn CM, Franco OH, de Jongste JC, Klaver CC, et al. The Generation R Study: Biobank update 2015. Eur J Epidemiol. 2014;29(12):911-27.
17. Kiefte-de Jong JC, de Vries JH, Bleeker SE, Jaddoe VW, Hofman A, Raat H, et al. Socio-demographic and lifestyle determinants of 'Western-like' and 'Health conscious' dietary patterns in toddlers. The British journal of nutrition. 2013;109(1):137-47.
18. Voortman T, Kiefte-de Jong JC, Geelen A, Villamor E, Moll HA, de Jongste JC, et al. The development of a diet quality score for preschool children and its validation and determinants in the Generation R Study. The Journal of nutrition. 2015;145(2):306-14.
19. Netherlands Nutrition Centre. Dutch Food Composition Database 2006 (NEVO 2006). The Hague2006.
20. Ocke M, van Rossum C, Fransen H, Buurma E, de Boer E, Brants H, et al. Dutch National Food Consumption Survey Young Children 2005/2006. National Institute for Public Health and the Environment (RIVM), 2008 Contract No.: RIVM Report 350070001.
21. Hulshof K, Ocke M, van Rossum C, Buurma-Rethans E, Brants H, Drijvers J, et al. Results of the national food consumption survey 2003. National Institute for Public Health and the Environment (RIVM), 2004 Contract No.: RIVM Report 350030002.
22. National Institute for Public Health and the Environment (RIVM). NEVO-online 2013: achtergrondinformatie Nederlands Voedingsstoffenbestand. Bilthoven, the Netherlands2013.
23. Cole TJ, Bellizzi MC, Flegal KM, Dietz WH. Establishing a standard definition for child overweight and obesity worldwide: international survey. Bmj. 2000;320(7244):1240.
24. Gishti O, Gaillard R, Manniesing R, Abrahamse-Berkeveld M, van der Beek EM, Heppe DH, et al. Fetal and infant growth patterns associated with total and abdominal fat distribution in school-age children. The Journal of clinical endocrinology and metabolism. 2014;99(7):2557-66.
25. Eisenmann JC. On the use of a continuous metabolic syndrome score in pediatric research. Cardiovascular Diabetology. 2008;7(17).
26. Voortman T, van den Hooven EH, Tielemans MJ, Hofman A, Kiefte-de Jong JC, Jaddoe VW, et al. Protein intake in early childhood and cardiometabolic health at school age: The Generation R Study. Eur J Nutr. 2015;55(6):2117-27.
27. van den Broek M, Leermakers ET, Jaddoe VW, Steegers EA, Rivadeneira F, Raat H, et al. Maternal dietary patterns during pregnancy and body composition of the child at age 6 y: the Generation R Study. The American journal of clinical nutrition. 2015;102(4):873-80.
28. Niklasson A, Ericson A, Fryer JG, Karlberg J, Lawrence C, Karlberg P. An update of the Swedish reference standards for weight, length and head circumference at birth for given gestational age (1977-1981). Acta paediatrica Scandinavica. 1991;80(8-9):756-62.
29. Hastie TJ. Generalized additive models. Chapter 7 of Statistical Models in S. In: Chambers JM, editor.: Wadsworth & Brooks/Cole; 1992.
30. Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. The American journal of clinical nutrition. 1997;65(4 Suppl):1220S-8S; discussion 9S-31S.
31. Wang DD, Li Y, Chiuve SE, Stampfer MJ, Manson JE, Rimm EB, et al. Association of Specific Dietary Fats With Total and Cause-Specific Mortality. JAMA internal medicine. 2016;176(8):1134-45.
32. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. Bmj. 2009;338:b2393.
33. Braun K, Erler N, Kiefte-de Jong JC, Jaddoe V, van den Hooven E, Franco O, et al. Dietary intake of protein in early childhood is associated with growth trajectories between 1 and 9 years of age. The Journal of nutrition. 2016: doi: 10.3945/jn.116.237164.

34. Voortman T, Braun KV, Kiefte-de Jong JC, Jaddoe VW, Franco OH, van den Hooven EH. Protein intake in early childhood and body composition at the age of 6 years: The Generation R Study. Int J Obes (Lond). 2016;40(6):1018-25.
35. Nupponen M, Pahkala K, Juonala M, Magnussen CG, Niinikoski H, Ronnemaa T, et al. Metabolic syndrome from adolescence to early adulthood: effect of infancy-onset dietary counseling of low saturated fat: the Special Turku Coronary Risk Factor Intervention Project (STRIP). Circulation. 2015;131(7): 605-13.
36. Niinikoski H, Jula A, Viikari J, Ronnemaa T, Heino P, Lagstrom H, et al. Blood pressure is lower in children and adolescents with a low-saturated-fat diet since infancy: the special turku coronary risk factor intervention project. Hypertension. 2009;53(6):918-24.
37. Kaitosaari T, Ronnemaa T, Viikari J, Raitakari O, Arffman M, Marniemi J, et al. Low-saturated fat dietary counseling starting in infancy improves insulin sensitivity in 9-year-old healthy children: the Special Turku Coronary Risk Factor Intervention Project for Children (STRIP) study. Diabetes care. 2006;29(4):781-5.
38. Heppe DH, Kiefte-de Jong JC, Durmus B, Moll HA, Raat H, Hofman A, et al. Parental, fetal, and infant risk factors for preschool overweight: the Generation R Study. Pediatric research. 2013;73(1):120-7.
39. Rolland-Cachera MF, Deheeger M, Bellisle F, Sempe M, Guilloud-Bataille M, Patois E. Adiposity rebound in children: a simple indicator for predicting obesity. The American journal of clinical nutrition. 1984;39(1):129-35.
40. Williams CL, Strobino BA. Childhood diet, overweight, and CVD risk factors: the Healthy Start project. Preventive cardiology. 2008;11(1):11-20.
41. Agostoni C, Riva E, Scaglioni S, Marangoni F, Radaelli G, Giovannini M. Dietary fats and cholesterol in italian infants and children. The American journal of clinical nutrition. 2000;72(5 Suppl):1384S-91S.
42. Simmer K, Patole SK, Rao SC. Long-chain polyunsaturated fatty acid supplementation in infants born at term. Cochrane Database Syst Rev. 2011(12):CD000376.
43. Cowin IS, Emmett PM, Alspac Study Team. Associations between dietary intakes and blood cholesterol concentrations at 31 months. European journal of clinical nutrition. 2001;55(1):39-49.
44. Chowdhury R, Warnakula S, Kunutsor S, Crowe F, Ward HA, Johnson L, et al. Association of dietary, circulating, and supplement fatty acids with coronary risk: a systematic review and meta-analysis. Annals of internal medicine. 2014;160(6):398-406.
45. Block G. Human dietary assessment: methods and issues. Preventive medicine. 1989;18(5):653-60.
46. Uusitalo L, Nevalainen J, Salminen I, Ovaskainen ML, Kronberg-Kippila C, Ahonen S, et al. Fatty acids in serum and diet--a canonical correlation analysis among toddlers. Matern Child Nutr. 2013;9(3): 381-95.
47. Burrows T, Berthon B, Garg ML, Collins CE. A comparative validation of a child food frequency questionnaire using red blood cell membrane fatty acids. European journal of clinical nutrition. 2012;66(7): 825-9.

Chapter 5

Nutrition & cardiometabolic health in adults

5.1

Associations of serum folate and vitamin B12 with body composition in elderly: The B-PROOF study

Based on:

Oliai Araghi S*, Braun KVE*, van der Velde N, van Dijk SC, van Schoor NM, Zillikens C, de Groot L, Uitterlinden A, Stricker B, Voortman T*, Kiefte-de Jong JC*. Associations of serum folate and vitamin B12 with body composition in elderly: The B-PROOF study. Submitted for publication. *Submitted for publication.*

**Denotes equal contribution*

ABSTRACT

Background: Folate and vitamin B12 have been reported to be associated with body composition, however, evidence on this association in elderly individuals is scarce. We aimed to investigate the associations of serum folate and vitamin B12 levels, as well as dietary intake of folic acid and vitamin B12 with body composition in a large population of older persons.

Methods: We used baseline data from the B-PROOF study, a randomized controlled trial which included 2919 individuals aged 65 years and older with elevated homocysteine levels. Venous blood samples were obtained to measure serum folate, vitamin B12, HoloTC, and MMA. Dietary intake of folate and vitamin B12 was measured with use of a food frequency questionnaire. Height and weight were measured and fat mass and fat-free mass were measured with use of Dual Energy X-ray assessment DXA. Body mass index (BMI), fat mass index (FMI) and fat-free mass index (FFMI) were calculated.

Results: In linear regression models adjusted for several medical history and lifestyle variables, higher levels of serum folate (nmol/L) were associated with a 0.02 lower BMI (95% CI -0.04; -0.01). This association was stronger in participants with cardiometabolic diseases (p-interaction=0.03), and was mainly driven by FFMI. Higher levels of HoloTC (pmol/L) were associated with a 0.96 higher FMI (95%CI 0.26; 1.65) and higher MMA (μgmol/L) was associated with a 1.11 lower FMI (95%CI -1.90; -0.32). No associations were observed between dietary intake of folate or vitamin B12 with body composition.

Conclusions: In this population of elderly, higher folate levels were associated with a lower BMI. This association was most pronounced in participants with prevalent cardiometabolic diseases, and was mainly driven by FFMI. In contrast, higher levels of vitamin B12, measured by HoloTC and MMA, were associated with a higher FMI. Future studies should further investigate underlying mechanisms and long-term effects on health.

INTRODUCTION

The prevalence overweight is growing in all age categories.[1] Being overweight or obese can alter the absorption, distribution, metabolism and/or excretion of micronutrients, which may cause several vitamin deficiencies.[2-4] This may be particularly of importance in the elderly as this group is mainly at high risk of deficiency.[5] For instance, vitamin B12 levels have been shown to decline with age and therefore prevalence of vitamin B12 deficiency and depletion is high in elderly.[6] This increase in prevalence of vitamin B12 deficiency is caused by both inadequate dietary intake as well as malabsorption.[7]

Several studies showed that there is an association between overweight and obesity with serum vitamin B12 and folate.[8-13] However, it is not clear in whether vitamin B deficiency is a cause or a consequence of obesity. It is suggested that obesity leads to malabsorption of vitamin B12 and folate, which causes deficiencies. On the other hand, vitamin B12 and folate deficiency due to low dietary intake may also lead to obesity. For example, Guananti et al. reported that vitamin B12 and folate intake were inversely associated with adiposity in children.[14] In addition, we also observed an inverse association between folate intake and BMI in children from a Dutch birth cohort.[15] These findings could be explained by the role of folate and vitamin B12 in epigenetic mechanisms. Vitamin B12 and folate act as co-factors in one-carbon metabolism, which is important for the forming of methyl donors for DNA methylation.[16, 17] Deficiency of these nutrients could lead to dysregulation of DNA methylation and might generate metabolic disturbances, including disturbed energy and lipid metabolism, leading to adiposity.[18, 19]

To date, studies exploring the association between folate and vitamin B12 with body composition in older individuals have not been conducted. Furthermore, measuring overweight by BMI in elderly is under debate.[20]Aging is associated with loss of muscle mass, and therefore BMI could underestimate the prevalence of obesity in this population.[21] To address these relations in an elderly population with changing body composition, the use of serum and dietary intake data simultaneously can provide a further insight on the relation of folate and vitamin B12 with body composition. Therefore, we aimed to investigate the associations of serum folate and vitamin B12 levels as well as dietary intake of folic acid and vitamin B12 from both food and supplements, with detailed measures of body composition in a large population of older persons.

MATERIALS AND METHODS

Study participants

For the present cross-sectional analyses, baseline data of the B-PROOF study (B-vitamins for the Prevention Of Osteoporotic Fractures) were used. The B-PROOF study is a multi-center, randomized, placebo-controlled, double-blind intervention study, investigating the effect of a 2-year daily oral vitamin B12 (500 µg) and folic acid (400 µg) supplementation on fracture incidence. The study was conducted in three research centers in the Netherlands: VU University Medical Center (Amsterdam), Wageningen University (Wageningen), and Erasmus Medical Center (Rotterdam). This study included 2919 individuals, aged 65 years and older with elevated homocysteine levels (12 - 50 µmol/l).

Participants were excluded if they had a renal insufficiency (creatinine level > 150 µmol/l) or presence of a malignancy in the past 5 years. A detailed description of the trial has been reported elsewhere.[22] All participants gave written informed consent before the start of the study. The B-PROOF study has been registered in the Netherlands Trial Register (NTRNTR1333) and with ClinicalTrials.gov (NCT00696514). The WU Medical Ethics Committee approved the study protocol, and the Medical Ethics committees of Erasmus MC and VUmc gave approval for local feasibility.[22]

Anthropometrics measurements

Height was measured in duplicate to the nearest 0.1 cm with the participant standing erect and without wearing shoes, using a stadiometer.[22] Weight was measured to the nearest 0.5 kg using a calibrated weighing device (SECA 761) with the participant wearing light garments, empty pockets and without wearing shoes.[22] BMI was calculated as weight in kilograms divided by square of height in meters and expressed as kg/m^2. Participants were categorized in underweight (BMI < 20), normal weight (BMI 20 to <25), overweight (BMI 25 to <30) and obesity (BMI ≥ 30).[23]

Body composition

At baseline, a subsample of participants from the Amsterdam and Rotterdam region (n=1227) underwent Dual Energy X-ray assessment (DXA) using the GE Lunar Prodigy device (GE Healthcare, USA, CV = 0.08%), (Erasmus MC) and the Hologic QDR 4500 Delphi device (Hologic Inc., USA, CV – 0.45%), (VuMC). The two devices were cross-calibrated by measuring a European spine phantom (ESP) five times on both devices and all results were adjusted accordingly. Total fat mass and total fat-free mass was calculated from the DXA scan.[22] Fat Mass Index (FMI) and Fat-Free Mass Index (FFMI) was calculated as total fat mass or total fat-free mass in kilograms divided by square of height in meters and expressed as kg/m^2.

Laboratory measurements

Venous blood samples were obtained in the morning, when the participants were in a fasted state, or had taken a restricted breakfast. Plasma homocysteine was determined at baseline, using the Architect i2000 RS analyser (VUmc, intra assay CV – 2%, inter assay CV = 4%), HPLC method[24] (WU, intra assay CV = 3.1%, inter assay CV = 5.9%) and LCMS/MS (EMC, CV = 3.1%). According to a cross-calibration, outcomes of the three centers did not differ significantly. Serum folate was determined by imunoelectrochemiluminescence on a Roche Modular E170 (Roche, Almere, The Netherlands) (CV = 5.9% at 5.7 nmol/L and 2.8% at 23.4 nmol/L). Serum methylmalonic acid (MMA) was measured by LC-MS/MS (CV < 9%) and holo-transcobalamin (HoloTC), and was determined by the AxSYM analyser (Abbott Diagnostics, Hoofddorp, the Netherlands) (CV < 8%).[22, 25] HoloTC was used as measure of vitamin B12 status, because it has been shown to better reflect vitamin B12 status than serum total vitamin B12.[26] To isolate DNA, buffy coats were used. The MTHFR genotypes, 677CC, 677CT or 677 TT, were determined using the Illumina Omni-express array (Illumina Inc., San Diego, CA, USA).

Dietary assessment

Dietary intake was estimated in a subsample, i.e. all participants of the Wageningen region (n=603), by a Food Frequency Questionnaire (FFQ) with a focus on macronutrients, vitamin B12, folate, vitamin D, and calcium. The FFQ was developed by the dietetics group of Wageningen University and was derived from an FFQ which was validated for energy, fat, cholesterol, folate and vitamin B12 intake.[22, 27, 28] Folic acid and vitamin B12 supplement use was defined as users or non-users.

Covariates

Demographic characteristics and health status variables, which included age, sex, self-reported medical history (cardiovascular disease and diabetes), alcohol intake, smoking habits, physical activity (Kcal/day), education and vitamin supplement use, were determined using a structured questionnaire. Alcohol intake was categorized according to the Dutch method of Garretsen et al. into 'never', 'light', 'moderate' and 'excessive' drinkers, based on the number of days per week alcohol was consumed and the number of glasses per time.[29, 30] Smoking habits were defined as never smoked; former smoker or current smoker.

Statistical analyses

Normal distribution for all variables was examined by visual inspection of histograms. When necessary, data were transformed. Second, linear regression analysis was used to determine associations of serum folate, vitamin B12, HoloTC, MMA, folic acid intake, and vitamin B12 intake (from food and supplement) with BMI, FMI and FFMI. All the analyses were adjusted for age and sex, and for energy intake for the association between folic acid and vitamin B12 intake (from food and supplements) (model 1). Subsequently, based on literature, smoking, alcohol, PA, homocysteine, education, hypertension, and hypercholesterolemia were added as fixed confounders.[31, 32] In addition, the interaction of age, gender, prevalent cardiometabolic diseases (cardiac disease or diabetes), MTHFR with serum folate and serum vitamin B12 on body composition was tested in the model 2. If the P value for interaction was < 0.1, stratified analyses were performed. In addition, we tested whether there was an interaction of serum folate with serum vitamin B12, MMA and HoloTC on body composition. Statistical analyses were performed using the statistical software package of SPSS 21.0 (SPSS Inc., Chicago, Illinois, USA). P-values of < 0.05 were considered statistically significant for all the analyses other than the interaction analyses (<0.1).

RESULTS

Population characteristics

Population characteristics are presented in **Table 5.1.1**. Mean age was 74.0 years (6.5 SD) for the total population (n=2919), 73 years (5.7 SD) for the DXA population (n=1227) and 72.8 years (5.7 SD) for the population with FFQ data available (n=603). Mean BMI was 27.1 kg/m^2 (4.0 SD) for the total population, 27.0 kg/m^2 (3.8 SD) for participants who underwent a DXA measurement, and 26.9 kg/m^2 (SD 3.6) for participants with FFQ data. The participants who underwent a DXA scan were younger; different alcohol consumption patterns (more moderate and excessive alcohol intake and

Table 5.1.1. Population characteristics

	B-PROOF Participants (N =2919)	DXA-test Participants (N = 1227)	FFQ participants (N=603)
Age (years)[a]	74 (6.5)	72.9 (5.7)	72.8 (5.7)
Sex			
Female (%)	50	48.3	42.1
Body Mass Index (kg/m^2)[a]	27.1 (4.0)	27.0 (3.8)	26.9 (3.6)
Underweight (%)	0.4	0.2	0.3
Normal weight (%)	28.6	30.4	28.4
Overweight (%)	50.9	50.1	54.7
Obesity (%)	20.1	19.2	16.6
Fat	NA		NA
Total Fat Mass (kg)		25.5 (8.4)	
Total Fat Percentage (%)		32.4	
FMI (kg/m^2)		8.9 (3.2)	
FFMI (kg/m^2)		18.0 (2.2)	
Smoking (%)			
Current	56.5	57.6	58.8
Former	9.6	9.0	10.4
Never	33.9	33.3	31.0
Alcohol intake (%)			
Light	67.4	64.1	64.2
Moderate	28.8	31.4	32.8
Excessive	3.4	4.0	2.5
Very excessive	0.4	0.6	0.5
Self-reported medical history of			
Cardiac disease (%)	25.1	25	25.5
Diabetes (%)	10.3	10.8	7.1
Hypercholesterolemia (%)	24.7	28.5	21.2
Measured hypertension (%)	51.5	58.8	38.6
Homocysteine (mmol/L)[b]	14.4 [3.6]	14.3 [3.3]	14.0 [3.2]
Serum Folate (nmol/L)[a]	21.0 (11.62)	21.3 (9.3)	20.1 (17.2)
Serum Vitamin B12 (pmol/L)[a]	285.5 (115.95)	287.5 (115.2)	281.2 (107.9)
Holotranscobalamin (pmol/L)[b]	59.0 [34]	67 [40]	60.0 [34.0]
MMA (mcgmol/L)[b]	0.2 [0.1]	0.2 [0.1]	0.2 [0.1]
MTHFR (%)			
CC	44.9	46.4	44.3
CT	42.1	40.8	42.6
TT	13.0	12.8	13.1
Folic Acid supplement use (%)	16.0	16.3	11.1
Vitamin B12 supplement use (%)	16.2	16.8	11.1

Table 5.1.1. Population characteristics (continued)

	B-PROOF Participants (N =2919)	DXA-test Participants (N = 1227)	FFQ participants (N=603)
Folate intake from food (mcg/day)[a]	NA	NA	191.5 (53.9)
Vitamin B12 intake from food (mcg/day)[a]	NA	NA	4.1 (2.0)
Total activity (Kcal/day)[b]	550.8 [483.8]	591.2 [505.4]	586.4 [512.4]
Total energy intake (Kcal/day)[a]	NA	NA	2005.9 (473.4)
Education (%)			
Low	32.1	32.5	25.9
Middle	42.0	41.1	40.8
High	26.0	26.4	33.3
Region (%)			
Amsterdam	26.6	34.6	0
Rotterdam	29.4	65.4	0
Wageningen	44.0	0	100

[a]Presented as mean (SD); or [b] median [IQR]

less very excessive drinkers); had a higher serum HoloTC when compared to the total B-PROOF population. The participants with FFQ data were also younger; less likely to have diabetes mellitus; had lower serum folate and vitamin B12; had lower use of folic acid and vitamin B12 supplementation; had higher PA and were more often high educated when compared to the total B-PROOF population. HoloTC was log-transformed, as this variable was non-normally distributed.

Folate and vitamin B12 biomarkers and BMI

Results of the linear regression analyses of serum folate, serum vitamin B12, HoloTC, MMA, folic acid supplements and vitamin B12 supplements with BMI for the total population, are shown in **Table 5.1.2**. Serum folate was inversely significantly associated with BMI after adjustments for covariates, indicating that for each nmol/L increase in serum folate, there was a 0.02 kg/m^2 lower BMI (β -0.02; 95% CI -0.04; -0.01) in the total population. Serum vitamin B12 was also inversely significantly associated with BMI in the crude model, but this association was no longer significant after adjustment for covariates (β -0.001; 95% CI -0.003; 0.000 vs. β 0.000; 95% CI -0.002; 0.001) in the total population. HoloTC, MMA, folic acid supplements and vitamin B12 supplements were not associated with BMI (**Table 5.1.2**).

We found a significant association of HoloTC log transformed and BMI in the population with a DXA measurement after adjustment for covariates (β 1.37; 95% CI 0.43; 2.32). Additionally, MMA was significant inversely associated with BMI in this particular population (β -1.27; 95% CI -2.35; -0.20)(**Supplementary Table 5.1.1**). We found no significant associations between intake of folic acid and vitamin B12 (total and from food only) with BMI (**Table 5.1.4**).

Table 5.1.2. Associations between serum folate, vitamin B12, HoloTC, MMA, folic acid supplements and vitamin B12 supplements with BMI in Total Population (n=2919)

	BMI	
	Model 1 β *95% CI*	*Model 2* β *95% CI*
Serum Folate (nmol/L)	-0.02 [0.03; -0.01]*	-0.02 [-0.04; -0.01]*
Serum Vitamin B12 (pmol/L)	-0.001 [-0.003; 0.000]*	-0.001 [-0.002; 0.001]
HoloTC_log[a] (pmol/L)	0.53 [-0.09; 1.15]	0.61 [-0.10; 1.32]
MMA (μgmol/L)	-0.26 [-0.742; 0.225]	-0.68 [-1.47; 0.12]
Folic acid supplements	-0.27 [-0.67; 0.12]	-0.10 [-0.54; 0.34]
Vitamin B12 supplements	-0.37 [-0.76; 0.03]	-0.27 [-0.71; 0.17]

Values are regression coefficients and 95% CIs based on linear regression models and reflect differences in BMI per 1 unit increase of serum foliate, serum vitamin B12, HoloTC, MMA, folic acid supplements, vitamin B12 supplements;
Model 1 is adjusted for age, sex.
Model 2 is additionally adjusted for smoking, alcohol consumption, physical activity, education, hypertension and hypercholesterolemia.
[a]log transformed
**P-value <0.05.*

Folate and vitamin B12 biomarkers and FMI and FFMI in DXA population

Results of the linear regression analyses of serum folate, serum vitamin B12, HoloTC, MMA, folic acid supplements and vitamin B12 supplements with FMI and FFMI for the DXA population are shown in **Table 5.1.3**. There was a significant association between HoloTC and FMI (β 0.96; 95% CI 0.26; 1.65) in the adjusted model, indicating that for each pmol/L increase in log transformed HoloTC, there was a 0.96 kg/m^2 higher FMI. However, MMA was significantly inversely associated with FMI, indicating that for each μgmol/L increase in MMA, there was a 1.11 kg/m^2 lower FMI (β -1.11; 95% CI -1.90; -0.32). Moreover, serum folate, serum vitamin B12, HoloTC, MMA, folic acid supplements and vitamin B12 supplements were not associated with FFMI.

Additional analyses

Analyses were stratified in case of a significant interaction was observed with one of the potential effect modifiers. Results of these stratified analyses are presented in **Supplementary Table 5.1.3**. We found a significant interaction between serum folate and cardiometabolic diseases in association with BMI (p-interaction=0.03) as well as FFMI (p-interaction=0.06). Stratification showed that the association between serum folate and BMI was most pronounced in participants with cardiometabolic diseases compared to participants without cardiometabolic diseases. In participants with cardiometabolic diseases a significant inverse association was observed between serum folate and BMI (β -0.05; 95%CI -0.09; -0.02), whereas no significant association was observed for participants without cardiometabolic diseases (β -0.01; 95% CI 0.03; -0.01) (**Supplementary Table 5.1.3**). Similarly, there was an inverse association between serum folate and FFMI in participants with cardiometabolic diseases (β -0.03; 95% CI -0.05; -0.01), but in participants without cardiometabolic diseases, this association

Table 5.1.3. Associations between serum folate, vitamin B12, HoloTC, MMA, folic acid supplements and vitamin B12 supplements with FMI and FFMI in DXA population (n=1227)

	FMI		FFMI	
	Model 1 β *95% CI*	*Model 2* β *95% CI*	*Model 1* β *95% CI*	*Model 2* β *95% CI*
Serum Folate (nmol/L)	-0.004 [-0.02; 0.01]	-0.002 [-0.02; 0.02]	-0.009 [-0.020; 0.002]	-0.01 [-0.02; 0.000]
Serum Vitamin B12 (pmol/L)	0.000 [-0.002; 0.001]	-0.00 [-0.001; 0.001]	0.000 [-0.001; 0.001]	0.000 [0.001; 0.001]
HoloTC_log[a] (pmol/L)	0.75 [0.07; 1.42]*	0.96 [0.26; 1.65]*	0.23 [-0.19; 0.65]	0.40 [-0.03; 0.84]
MMA (μgmol/L)	-0.75 [-1.53; 0.04]	-1.11 [-1.90; -0.32]*	-0.14 [-0.63; 0.35]	-0.17 [-0.67; 0.33]
Folic acid supplements	-0.03 [-0.46; 0.40]	0.17 [-0.27; 0.62]	-0.18 [-0.44; 0.09]	-0.11 [-0.38; 0.17]
Vitamin B12 supplements	-0.18 [-0.61; 0.24]	-0.05 [-0.49; 0.39]	-0.19 [-0.46; 0.08]	-0.11 [-0.39;0.16]

Values are regression coefficients and 95% CIs based on linear regression models and reflect differences in BMI per 1 unit increase of serum foliate, serum vitamin B12, HoloTC, MMA, folic acid supplements, vitamin B12 supplements.
Model 1 is adjusted for age, sex.
Model 2 is additionally adjusted for smoking, alcohol consumption, physical activity, education, hypertension, hypercholesterolemia.
Model 3 is additionally adjusted for diabetes and homocysteine.
[a]log transformed
**P-value <0.05.*

Table 5.1.4. Associations between serum folate, vitamin B12, HoloTC, MMA and folic acid intake and supplements, vitamin B12 intake and supplements with BMI in FFQ population (n=603)

	BMI	
	Model 1 β *95% CI*	*Model 2* β *95% CI*
Serum Folate (nmol/L)	-0.004 [-0.02; 0.01]	-0.02 [-0.06; 0.03]
Serum Vitamin B12 (pmol/L)	-0.003 [-0.006; 0.000]*	-0.001 [-0.005; 0.002]
HoloTC_log[a] (pmol/L)	0.04 [-1.26; 1.35]	0.03 [-1.74; 1.81]
MMA (μgmol/L)	-0.15 [-0.76; 0.46]	-0.78 [-2.99; 1.43]
Folic acid total intake (FFQ)	0.000 [-0.002; 0.001]	0.00 [-0.003; 0.002]
Vitamin B12 total intake (FFQ)[a]	-0.84 [-2.13; 0.46]	-1.06 [-2.76; 0.63]
Folic acid intake from food (FFQ)	0.005 [-0.002; 0.011]	0.007 [-0.001; 0.02]
Vitamin B12 total intake from food (FFQ)	0.03 [-0.014; 0.079]	0.07 [-0.004; 0.15]
Folic acid supplements	-0.16 [-1.08; 0.75]	-0.45 [-1.53; 0.64]
Vitamin B12 intake supplements	-0.12 [-1.04; 0.80]	-0.51 [-1.58; 0.56]

Values are regression coefficients and 95% CIs based on linear regression models and reflect differences in FMI per 1 unit increase of serum foliate, serum vitamin B12, HoloTC, MMA, folic acid intake (from food and supplements), vitamin B12 intake (from food and supplements); Model 1 is adjusted for age, sex (and energy intake for intake form food). Model 2 is additionally adjusted for smoking, alcohol consumption, physical activity, education, hypertension and hypercholesterolemia.
*[a]log transformed *P-value <0.05.*

was not significant (β -0.01; 95%CI 0.05; -0.01). We also observed a significant interaction between vitamin B12 supplement use and cardiometabolic diseases with FFMI, but stratified analyses showed no significant associations.

Age was found to be an effect modifier in the association between MMA and FMI (p-interaction=0.03). MMA was inversely associated with FMI in the population older than 73 years (β -1.92; 95% CI -2.90; -0.94), but not in participants younger than 73 years (β -0.07; 95% CI -1.33; 1.19). In addition, we found a significant interaction between sex and MMA (p=0.04) with FFMI. In women, an inverse association was observed between (β -0.65; 95% CI -1.26; -0.03). In contrast, a positive association between MMA and FFMI was observed for men, although this was non-significant (β 0.68; 95% CI -0.20; 1.55). There was also a significant interaction between sex and vitamin B12 intake with BMI. A higher vitamin B12 intake was associated with a lower BMI in men (β -2.86; 95% CI -4.58; -1.13), but not in women.

DISCUSSION

In the current study, we investigated associations between serum levels as well as dietary intake of vitamin B12 and folate with body composition in elderly individuals. We observed that a higher level of serum folate was associated with a lower BMI. This association was most pronounced in participants with cardiometabolic diseases, and seemed to be driven by FFMI rather than FMI in this subgroup. Furthermore, HoloTC was positively associated with FMI and MMA was negatively associated with FMI, both suggesting that higher levels of vitamin B12 are associated with a higher FMI. However, we did not observe any associations between dietary intake of vitamin B12 or folate with body composition.

Our findings regarding folate are in line with results from several previous studies. In a cross sectional study among adults Kimmons et al. showed that compared with normal-weight adults, overweight and obese adults were more likely to have low folate levels.[8] In line with this, a study by Mahabir et al. showed that adiposity was associated with lower serum folate levels in postmenopausal women.[9] The inverse association we observed between folate levels and BMI was modified by the presence of cardiometabolic diseases. Stratified analyses showed that associations were stronger in participants with cardiometabolic diseases, and that this association was mainly driven by FFMI. However, no associations between folate levels with BMI or FFMI were observed in participants without cardiometabolic diseases. Possibly, our findings in participants with cardiometabolic diseases are a reflection of a different underlying mechanisms. Another explanation could be reverse causality. Changes in lifestyle that are made after disease incidence may result in an improved folate status, while a detrimental physical status due to cardiometabolic diseases, such as a lower FFMI may still be present.

In contrast to our expectations, we observed that several biomarkers of vitamin B12, including higher levels of HoloTC and lower levels of MMA, were associated with a higher FMI. In a previous study among children we observed similar detrimental associations between vitamin B12 intake and android-gynoid ratio.[15] However, in another study among adults from a primary care-based setting vitamin B12 was negatively correlated with BMI.[10] Considering the age of our study population, lack

of intrinsic factor could play a role in these findings. This could lead to malabsorption of vitamin B12, which could be related to other health outcomes affecting body composition. Therefore, participants with lower vitamin B12 levels could have lower FMI as a reflection of a detrimental physical status. This is supported by our results showing that the negative association for MMA with FMI is stronger in older participants (>73 years).

From these studies investigating levels of folate and vitamin B12 with body composition it is not possible to determine the temporal direction. Hypotheses has been proposed for both directions: folate and vitamin B12 deficiency may be a consequence of obesity due to malabsorption, but may also be a cause of obesity due to metabolic disturbances through insulin resistance, or epigenetic mechanism, such as DNA methylation.[18, 19] In addition to biomarkers of folate and vitamin B12, we examined associations between dietary intake of folate and vitamin B12 with body composition. However, our results do not provide evidence that dietary intake of vitamin B12 and folate are involved in the development of overweight or obesity. This is in line with a previous study in which the causal relationship between vitamin B12 and BMI was studied with use of a Mendelian randomization approach. The authors reported that vitamin B12 levels were associated with BMI, however, there was no evidence that a higher BMI was caused by lower vitamin B12 levels.[33] In line with this, in a study among Mexican children an inverse association was observed between folate vitamin B12 levels and body composition, but not with dietary intake of folate and vitamin B12.[14]

Strengths and limitations

The current study has several strengths and limitations. An important strength is the size of the population included and the large amount of information available on potentially confounding factors. However, residual confounding cannot be ruled out, as in any observational study. For instance, the positive association between biomarkers of vitamin B12 with FMI could be due to other dietary factors. Considering that vitamin B12 only naturally occurs in animal products, the positive association between vitamin B12 biomarkers and BMI could be explained by a higher animal-based diet accompanied with a higher caloric intake. However, we only had data available on energy intake in a subset, but not the total population. Therefore, we could not determine whether this association was explained by other dietary factors in this population. As dietary intake was only available in a subgroup, this may have resulted in insufficient statistical power, and therefore we may have been unable to detect an association between dietary intake and body composition. Furthermore, due to the cross-sectional nature of our data, we are not able to determine a temporal relationship of the associations we observed. We had several methods to measure vitamin B12 levels which provided better insight in potential deficiencies. For instance, HoloTC has been shown to be a more accurate to detect vitamin B12 deficiency compared to other methods.[34] Another strength of this study was that detailed body composition measurements were available, in addition to the use of BMI. Although BMI may be practical to use as outcome of interest, this measure is limited regarding prediction of several health outcomes because it does not take into account body composition.[35, 36] Therefore, we used more detailed body composition measured with DXA, including FMI and FFMI. However, we did not have data on distribution of body fat. This may also be of importance when predicting disease risk, as abdominal fat rather than gynoid fat may have different effects on health.[37]

Conclusion

In this population of elderly, higher folate serum levels were associated with a lower BMI. This association was most pronounced in participants with prevalent cardiometabolic diseases, and was mainly driven by FFMI rather than FMI among these participants. In contrast, higher levels of vitamin B12, measured by HoloTC and MMA, were associated with a higher FMI. Future large-scale studies should further investigate underlying mechanisms and long-term effects on health.

SUPPLEMENTARY MATERIAL

Supplementary Table 5.1.1. Associations between serum folate, vitamin B12, HoloTC, MMA, folic acid supplements and vitamin B12 supplements with BMI in DXA population (n=1227)

	BMI	
	Model 1 β *95% CI*	*Model 2* β *95% CI*
Serum Folate (nmol/L)	-0.01 [-0.03; 0.01]	-0.01 [-0.04; 0.01]
Serum Vitamin B12 (pmol/L)	0.000 [-0.002; 0.001]	0.000 [-0.002; 0.002]
HoloTC_log[a] (pmol/L)	0.98 [0.07; 1.90]*	1.37 [0.43; 2.32]*
MMA (µgmol/L)	-0.88 [-1.943; 0.185]	-1.27 [-2.35; -0.20]*
Folic acid supplements	-0.17 [-0.75; 0.41]	0.06 [-0.54; 0.67]
Vitamin B12 supplements	-0.34 [-0.91; 0.24]	-0.17 [-0.76; 0.43]

Values are regression coefficients and 95% CIs based on linear regression models and reflect differences in BMI per 1 unit increase of serum folate, serum vitamin B12, HoloTC, MMA, folic acid supplements, vitamin B12 supplements;
Model 1 is adjusted for age, sex.
Model 2 is additionally adjusted for smoking, alcohol consumption, physical activity, education, hypertension and hypercholesterolemia.
*[a]log transformed *P-value <0.05.*

Supplementary Table 5.1.2. Associations between intake of folate and vitamin B12 and serum folate, vitamin B12, holoTC, MMA

	Folic acid intake β *95% CI*[a]	Vitamin B12 intake β *95% CI*[a]
Serum Folate (nmol/L)	0.065 [0.060; 0.069]*	NA
Serum Vitamin B12 (pmol/L)	NA	1.469 [0.087; 2.851]*
HoloTC (pmol/L)	NA	-0.001 [-0.007; 0.005]
MMA (µgmol/L)	NA	0.970 [0.479; 1.462]*

[a]adjusted for age and sex

Supplementary Table 5.1.3. Associations between serum folate, MMA, vitamin B12 supplements and vitamin B12 intake with BMI, FMI, and FFMI stratified for effect modifiers (p-interaction < 0.10)

	BMI β *95% CI*[1] β *95% CI*[2]	FMI	FFMI
Serum folate	**Cardiometabolic diseases(y/n) (p-interaction 0.03)** No: -0.01 (-0.03; 0.01) Yes: -0.05 (-0.09; -0.02)*	-	**Cardiometabolic diseases (y/n) (p-interaction 0.06)** No: -0.01 (-0.02; 0.01) Yes: -0.03 (-0.05; -0.01)*
MMA	-	***Age* (p-interaction *0.03*)** Age <73 yrs.: -0.07 (-1.33; 1.19) Age >73 yrs.: -1.92 (-2.90; -0.94)*	***Sex* (p-interaction *0.04*)** Male: 0.68 (-0.20; 1.55) Female: -0.65 (-1.26; -0.03)*
Vitamin B12 supplements	-	-	**Cardiometabolic diseases (y/n) (p-interaction *0.07*)** No: -0.25 (-0.55; 0.06) Yes: 0.33 (-0.25; 0.92)
Vitamin B12 total_log	***Age* (p-interaction *0.07*)** Age <73 yrs.: -0.76 (-3.20; 1.68) Age >73 yrs.: -1.84 (-4.23; 0.55) ***Sex* (p-interaction *0.03*)** Male: -2.86 (-4.58; -1.13)* Female: 1.63 (-1.48; 4.74)	-	-

Values are regression coefficients and 95% CIs based on linear regression models and reflect differences in BMI, FMI, or FFMI per 1 unit increase of serum folate, serum vitamin B12, HoloTC, MMA, folic acid supplements, vitamin B12 supplements;
Models are adjusted for age, sex, smoking, alcohol consumption, physical activity, education, hypertension and hypercholesterolemia.
**P-value <0.05.*

REFERENCES

1. Mathus-Vliegen EM. Obesity and the elderly. J Clin Gastroenterol. 2012;46(7):533-44.
2. Patrini C, Griziotti A, Ricciardi L. Obese individuals as thiamin storers. Int J Obes Relat Metab Disord. 2004;28(7):920-4.
3. Mensink GB, Fletcher R, Gurinovic M, Huybrechts I, Lafay L, Serra-Majem L, et al. Mapping low intake of micronutrients across Europe. Br J Nutr. 2013;110(4):755-73.
4. Lips P. Vitamin D deficiency and secondary hyperparathyroidism in the elderly: consequences for bone loss and fractures and therapeutic implications. Endocr Rev. 2001;22(4):477-501.
5. Hughes CF, Ward M, Hoey L, McNulty H. Vitamin B12 and ageing: current issues and interaction with folate. Ann Clin Biochem. 2013;50(Pt 4):315-29.
6. Baik HW, Russell RM. Vitamin B12 deficiency in the elderly. Annu Rev Nutr. 1999;19:357-77.
7. Allen LH. How common is vitamin B-12 deficiency? Am J Clin Nutr. 2009;89(2):693S-6S.
8. Kimmons JE, Blanck HM, Tohill BC, Zhang J, Khan LK. Associations between body mass index and the prevalence of low micronutrient levels among US adults. MedGenMed. 2006;8(4):59.
9. Mahabir S, Ettinger S, Johnson L, Baer DJ, Clevidence BA, Hartman TJ, et al. Measures of adiposity and body fat distribution in relation to serum folate levels in postmenopausal women in a feeding study. Eur J Clin Nutr. 2008;62(5):644-50.
10. Baltaci D, Kutlucan A, Turker Y, Yilmaz A, Karacam S, Deler H, et al. Association of vitamin B12 with obesity, overweight, insulin resistance and metabolic syndrome, and body fat composition; primary care-based study. Med Glas (Zenica). 2013;10(2):203-10.
11. Abu-Samak M, Khuzaie R, Abu-Hasheesh M, Jaradeh M, Fawzi M. Relationship of vitamin B12 deficiency with overweight in male Jordanian youth. Journal of Applied Sciences. 2008;8(17):3060-3.
12. Pinhas-Hamiel O, Doron-Panush N, Reichman B, Nitzan-Kaluski D, Shalitin S, Geva-Lerner L. Obese children and adolescents: a risk group for low vitamin B12 concentration. Arch Pediatr Adolesc Med. 2006;160(9):933-6.
13. Reitman A, Friedrich I, Ben-Amotz A, Levy Y. Low plasma antioxidants and normal plasma B vitamins and homocysteine in patients with severe obesity. Isr Med Assoc J. 2002;4(8):590-3.
14. Gunanti IR, Marks GC, Al-Mamun A, Long KZ. Low serum vitamin B-12 and folate concentrations and low thiamin and riboflavin intakes are inversely associated with greater adiposity in Mexican American children. J Nutr. 2014;144(12):2027-33.
15. Braun KV, Voortman T, Kiefte-de Jong JC, Jaddoe VW, Hofman A, Franco OH, et al. Dietary Intakes of Folic Acid and Methionine in Early Childhood Are Associated with Body Composition at School Age. J Nutr. 2015;145(9):2123-9.
16. Niculescu MD, Zeisel SH. Diet, methyl donors and DNA methylation: interactions between dietary folate, methionine and choline. The Journal of nutrition. 2002;132(8):2333S-5S.
17. Friso S, Choi S-W. Gene-nutrient interactions and DNA methylation. The Journal of nutrition. 2002; 132(8):2382S-7S.
18. Milagro FI, Mansego ML, De Miguel C, Martinez JA. Dietary factors, epigenetic modifications and obesity outcomes: progresses and perspectives. Mol Aspects Med. 2013;34(4):782-812.
19. Glier MB, Green TJ, Devlin AM. Methyl nutrients, DNA methylation, and cardiovascular disease. Molecular Nutrition & Food Research. 2014;58(1):172-82.
20. Heim N, Snijder, M.B., Deeg, D.J.H., Seidell, J.C., Visser, M. Obesity in older adults is associated with an increased prevalence and incidence of pain. Obesity Journal. 2008;16(11):2510-7.

21. Gallagher D, Visser M, De Meersman RE, Sepulveda D, Baumgartner RN, Pierson RN, et al. Appendicular skeletal muscle mass: effects of age, gender, and ethnicity. J Appl Physiol (1985). 1997;83(1): 229-39.
22. van Wijngaarden JP, Dhonukshe-Rutten RA, van Schoor NM, van der Velde N, Swart KM, Enneman AW, et al. Rationale and design of the B-PROOF study, a randomized controlled trial on the effect of supplemental intake of vitamin B12 and folic acid on fracture incidence. BMC Geriatr. 2011;11:80.
23. Must A, Dallal GE, Dietz WH. Reference data for obesity: 85th and 95th percentiles of body mass index (wt/ht2) and triceps skinfold thickness. Am J Clin Nutr. 1991;53(4):839-46.
24. Ubbink JB, Hayward Vermaak WJ, Bissbort S. Rapid high-performance liquid chromatographic assay for total homocysteine levels in human serum. J Chromatogr. 1991;565(1-2):441-6.
25. van Wijngaarden JP, Swart KM, Enneman AW, Dhonukshe-Rutten RA, van Dijk SC, Ham AC, et al. Effect of daily vitamin B-12 and folic acid supplementation on fracture incidence in elderly individuals with an elevated plasma homocysteine concentration: B-PROOF, a randomized controlled trial. Am J Clin Nutr. 2014;100(6):1578-86.
26. Heil SG, de Jonge R, de Rotte MC, van Wijnen M, Heiner-Fokkema RM, Kobold AC, et al. Screening for metabolic vitamin B12 deficiency by holotranscobalamin in patients suspected of vitamin B12 deficiency: a multicentre study. Ann Clin Biochem. 2012;49(Pt 2):184-9.
27. Feunekes GI, Van Staveren WA, De Vries JH, Burema J, Hautvast JG. Relative and biomarker-based validity of a food-frequency questionnaire estimating intake of fats and cholesterol. Am J Clin Nutr. 1993;58(4):489-96.
28. Verkleij-Hagoort AC, de Vries JH, Stegers MP, Lindemans J, Ursem NT, Steegers-Theunissen RP. Validation of the assessment of folate and vitamin B12 intake in women of reproductive age: the method of triads. Eur J Clin Nutr. 2007;61(5):610-5.
29. Garretsen H. Probleemdrinken: Prevalentiebepaling, beinvloedende factoren en preventiemogelijkheden: Theoretische overwegingen en onderzoek in Rotterdam. (in Dutch). Lisse, Swets & Zeitlinger 1983.
30. Mulder PG, Garretsen HF. Are epidemiological and sociological surveys a proper instrument for detecting true problem drinkers? (The low sensitivity of an alcohol survey in Rotterdam). Int J Epidemiol. 1983;12(4):442-4.
31. Koolhaas CM, van Rooij FJA, Schoufour JD, Cepeda M, Tiemeier H, Brage S, et al. Objective Measures of Activity in the Elderly: Distribution and Associations With Demographic and Health Factors. J Am Med Dir Assoc. 2017;18(10):838-47.
32. Bradshaw M, Kent BV, Henderson WM, Setar AC. Subjective social status, life course SES, and BMI in young adulthood. Health Psychol. 2017;36(7):682-94.
33. Allin KH, Friedrich N, Pietzner M, Grarup N, Thuesen BH, Linneberg A, et al. Genetic determinants of serum vitamin B12 and their relation to body mass index. Eur J Epidemiol. 2017;32(2):125-34.
34. Clarke R, Sherliker P, Hin H, Nexo E, Hvas AM, Schneede J, et al. Detection of vitamin B12 deficiency in older people by measuring vitamin B12 or the active fraction of vitamin B12, holotranscobalamin. Clin Chem. 2007;53(5):963-70.
35. Nevill AM, Stewart AD, Olds T, Holder R. Relationship between adiposity and body size reveals limitations of BMI. Am J Phys Anthropol. 2006;129(1):151-6.
36. Freedman DS, Wang J, Ogden CL, Thornton JC, Mei Z, Pierson RN, et al. The prediction of body fatness by BMI and skinfold thicknesses among children and adolescents. Ann Hum Biol. 2007;34(2):183-94.
37. Klein S, Allison DB, Heymsfield SB, Kelley DE, Leibel RL, Nonas C, et al. Waist circumference and cardiometabolic risk: a consensus statement from shaping America's health: Association for Weight

Management and Obesity Prevention; NAASO, the Obesity Society; the American Society for Nutrition; and the American Diabetes Association. Diabetes Care. 2007;30(6):1647-52.

5.2

Methyl donor nutrient intake and incidence of type 2 diabetes mellitus: results from the Nurses' Health Study and Health Professionals Follow-Up Study.

Based on:

Braun KVE, Satija A, Voortman T, Franco OH, Sun Q, Bhupathiraju SN, Hu FB. Methyl donor nutrient intake and incidence of type 2 diabetes mellitus: results from the Nurses' Health Study and Health Professionals Follow-Up Study. *Manuscript in preparation*

ABSTRACT

Background There is increasing evidence that DNA methylation may play a role in the development of diabetes. DNA methylation can be influenced by environmental factors, such as diet. These dietary factors include methyl donor nutrients, including vitamin B2, vitamin B6, vitamin B12, folate, and methionine, which are important in one-carbon metabolism. Intake of these nutrients may influence risk of type 2 diabetes via epigenetic changes.

Methods We included 69,949 women from the Nurses' Health Study (1984-2012), 90,239 women from the Nurses' Health Study 2 (1991-2011), and 40,539 men from the Health Professionals Follow-Up Study (1986-2010). Dietary data were collected every 2-4 years using a semi-quantitative food frequency questionnaire, from which dietary intake of vitamin B2, vitamin B6, vitamin B12, folate, and methionine were calculated. Hazard ratios were estimated using Cox proportional hazard regression analyses with time-varying covariates.

Results We documented 8,141 cases of type 2 diabetes during 1,763,428 years of follow-up. In pooled multivariable-adjusted analysis, the highest quartiles of vitamin B2 and vitamin B6 were associated with a lower risk of type 2 diabetes compared to the lowest quartile (Q1) (HR [95% CI]: 0.90 [0.84, 0.97], and 0.91 [0.84, 0.97], respectively). Intake of total vitamin B12 was not associated with risk of diabetes. However, when analyses were stratified by food source, vitamin B12 intake from food in Q5 was associated with a higher risk of diabetes compared to Q1 (HR [95% CI]: 1.11 [1.02, 1.19], whereas a high B12 intake from supplements was associated with lower risk of diabetes (HR [95% CI]: 0.92 [0.85, 0.98]). There was no significant association between folate and methionine intake and risk of diabetes.

Conclusion Our study suggests that higher intakes of vitamin B2 and vitamin B6 are associated with a lower risk of diabetes. A higher vitamin B12 intake from food seems to be associated with a higher risk of diabetes, which may be due to animal products.

BACKGROUND

The prevalence of type 2 diabetes is growing worldwide and leads to a major increase in morbidity and mortality.[1] Therefore, it is important to identify risk factors in the development of type 2 diabetes in order to develop targeted prevention strategies.

There is increasing evidence that DNA methylation may play a role in the development of metabolic disorders, including diabetes and diabetes related complications.[2-4] DNA methylation can be influenced by environmental factors, such as diet.[5] These dietary factors include methyl donor nutrients, such as folate and vitamin B12, which are important in one-carbon metabolism. Intake of these nutrients may influence risk of type 2 diabetes via epigenetic changes.[6] Only a few studies have investigated the association between methyl donor nutrients and diabetes. Krishnaveni et al. reported that vitamin B12 deficiency in pregnant women was associated with risk of both gestational diabetes and diabetes after 5 years of follow up.[7] In contrast, in a randomized controlled trial among women at high risk of CVD, Song et al. reported no significant effect of supplementation of vitamin B6, vitamin B12, and folate on the risk of type 2 diabetes.[8] However, in the same study a trend towards a protective effect in participants with family history of type 2 diabetes was reported. These findings suggest that vitamin B6, vitamin B12 and folate may be beneficial for a specific subgroup which is predisposed to development of type 2 diabetes.

To date, no studies have been conducted on the association between the intake of methyl donor nutrients and risk of type 2 diabetes in adults in large prospective cohort studies. Therefore, the aim of this study was to investigate the association between intake of methyl donor nutrients and risk of type 2 diabetes, and whether these associations are modified by family history of type 2 diabetes.

METHODS

Study population

This study was embedded in three prospective cohort studies in the USA: the Nurses' Health Study (NHS), Nurses' Health Study II (NHS2), and the Health Professionals' Follow-Up Study (HPFS). The NHS started in 1976 with 121,701 female nurses aged 30-55 years, the NHS2 started in 1989 with 116,430 female nurses aged 25-42 years, and the HPFS 1986 with 51,529 male health professionals aged 40-75 years. In all three studies, information on lifestyle and medical history was collected by questionnaires at baseline and every 2 years during follow-up. The response rate per cycle was ~90% for all cohorts. In this study, the 1984, 1991, and 1986 cycles were baselines for NHS, NHS2, and HPFS, respectively, since most covariates of interest were comprehensively measured from this cycle onwards. Participants with type 1 diabetes or type 2 diabetes mellitus, history of gestational diabetes, cardiovascular disease, cancer, missing dietary data, or inadequate dietary data (adequate dietary information was defined as:>50 of 61 items completed, yielding 500-3500 kcal/d, for women; >60 of 131 items completed, yielding 800-4200 kcal/d, for men). The population for analyses consisted of 69,949 women from the Nurses' Health Study, 90,239 women from the Nurses' Health Study 2, and 40,539 men from the Health Professionals Follow-Up Study. Study protocols for all cohorts were approved by the institutional review board of Brigham and Women's Hospital and Harvard T.H. Chan School of Public Health and all participants provided informed consent.

Dietary assessment

Dietary data were collected every 2-4 years using a validated semi-quantitative food frequency questionnaire as described previously.[9, 10] Participants were asked how often they consumed a serving of ~130 food items over the previous year. The frequency of a food item could range from "never or less than once/month" to "≥ 6 times/day". Dietary intake of vitamin B2, vitamin B6, vitamin B12, folate, and methionine were calculated by multiplying the consumption frequency of each food by the nutrient content of the portion specified. Nutrient values of foods were obtained from the Harvard University Food composition Database, derived mainly from the US Department of Agriculture sources. All nutrients were adjusted for total energy intake with use of the residual method.[11]

Ascertainment of type 2 diabetes

Participants in all three cohorts were asked whether they were diagnosed with diabetes by a physician every two years. Participants who self-reported physician diagnosed diabetes were sent a supplementary questionnaire to confirm diagnosis.[12, 13] Participants were defined as type 2 diabetes cases if they met at least one of the following criteria (according to the National Diabetes Data Group)[14]: classic symptoms plus fasting blood glucose ≥7.8 mmol/L or random blood glucose ≥11.1mmol/L; no symptoms, but increased blood glucose levels on two different occasions (i.e. fasting blood glucose ≥7.8 mmol/L or random blood glucose ≥11.1mmol/L); treatment with hypoglycemic drugs. Starting June 1998 the threshold for fasting plasma glucose was changed to ≥7.0 mmol/L.[15] Starting 2010 HbA1c ≥ 6.5% was added to the diagnosis criteria.[16]

Covariates

At baseline height was ascertained and during follow-up information on weight, smoking, marital status, family history of type 2 diabetes, hypertension, hypercholesterolemia, level of physical activity, total energy intake, alcohol consumption, and ethnicity was ascertained with use of biennial questionnaires. For participants from NHS and NHS2, information on postmenopausal status and postmenopausal hormone use, and for participants from NHS2 only information on oral contraceptive use was obtained.

Statistical analysis

We calculated cumulative updated averages of intake of vitamin B2, vitamin B6, vitamin B12, folate and methionine for each cohort to better estimate long-term intakes and to reduce within-person variation. Missing data on dietary intake during follow-up were replaced by the cumulative average of previous assessments. Updating was stopped in case of a cancer or CVD event, as diet may be changed in response to these events. Quantiles of each nutrient were calculated to limit the influence of outliers and to detect potential non-linear associations. Person time was calculated for each participant from questionnaire return date until type 2 diabetes diagnosis, death, or end of follow-up. Hazard ratios were estimated using Cox proportional hazard regression analyses with time-varying covariates to assess the association between quintiles of vitamin B2, vitamin B6, vitamin B12, folate and methionine and risk of type 2 diabetes. Age was included as the timescale. All models were adjusted for smoking, marital status, family history of type 2 diabetes, hypertension, hypercholesterolemia, level of physical activity, total energy intake, alcohol consumption, ethnicity, and BMI. In additional

models adjustments were made for other nutritional factors, including cereal fiber, animal protein, and PUFA:SFA ratio. Effect modification by age, family history of diabetes, and alcohol consumption was tested by including cross-product terms in the models. Analysis were performed per cohort and then combined in meta-analysis using a fixed effects model: the Cochrane Q statistic, the I^2 statistic, and between-study coefficient of variation were used to assess heterogeneity. All analyses were performed using SAS version 9.4 for UNIX (SAS Institute).

RESULTS

The baseline characteristics of participants are presented in **Table 5.2.1**. 5.2Mean and median intakes of vitamin B2, vitamin B6, vitamin B12, and methionine remained similar between baseline and during follow-up (data not shown). In contrast, due to the folate fortification which was implemented in 1998 (**Figure 5.2.1**),[17] median [95% range] intake of folate changed from 301.0 [230.0-474] in 1984 to 372.0 [262.0-650.0] in 2010.

We documented 17,629 cases of type 2 diabetes during 4,454,307 years of follow-up. The associations between quintiles of vitamin B2, vitamin B6, vitamin B12, folate, and methionine with risk of type 2 diabetes are presented in **Tables 5.2.2-5.2.6**. In pooled multivariable-adjusted analysis, the highest quintile(Q5) of vitamin B2 was associated with a lower risk of type 2 diabetes compared to the lowest quintile(Q1) (HR [95% CI]: 0.92 [0.88, 0.97]). A higher vitamin B6 was also associated with a lower risk of type 2 diabetes (HR [95% CI]: Q5 vs. Q1 0.93 [0.89, 0.98]). The associations for vitamin B2 and vitamin B6 remained significant for after adjustment for several other nutrients. Intake of total vitamin B12 was not associated with risk of diabetes. However, when analyses were stratified by food source, vitamin B12 intake from food in Q5 was associated with a higher risk of diabetes compared to Q1, but this was no longer significant after adjustment of other nutrients (HR [95% CI]: 1.05 [0.99, 1.10]). In contrast, a high B12 intake from supplements was associated with lower risk of diabetes (HR [95% CI] Q5 vs. Q1: 0.94 [0.90, 0.99]). There was a tendency toward a protective association for a high folate intake on risk of diabetes (HR [95% CI] Q5 vs. Q1: 0.90 [0.85, 0.94]). However, this association was attenuated after adjustment of other nutrients (HR [95% CI] Q5 vs. Q1: 0.95 [0.90, 1.00]). A high methionine intake was associated with a higher risk of diabetes(HR [95% CI] Q5 vs. Q1: 1.06 [1.01, 1.11]), but this was explained by other nutrients (HR [95% CI] Q5 vs. Q1: 0.97 [0.86, 1.09]).

We tested interactions for sex, alcohol intake, and family history of diabetes, but none of the interaction terms were significant. Therefore, all results are presented for the total population. In order to determine whether the folic acid fortification implementation in 1998 affected the association between folate intake and risk of type 2 diabetes, we performed several sensitivity analyses. First, we continuously updated folate intake, instead of cumulatively updating, but this did not change the results (data not shown). Second, we separated analyses for time periods before and after 1998. From these two time periods, similar results were observed compared to the total follow-up period (data not shown).

Table 5.2.1. Age-standardized characteristics by quintiles of methyl donor nutrients.

	B2			B6			B12			Folate			Methionine		
	Q1	Q 3	Q5	Q1	Q3	Q 5	Q1	Q3	Q5	Q1	Q3	Q 5	Q 1	Q3	Q 5
NHS															
Age (years)	49.76 (7.07)	50.30 (7.20)	50.49 (7.14)	49.73 (7.03)	50.42 (7.15)	50.44 (7.15)	49.88 (7.10)	50.12 (7.14)	50.72 (7.14)	49.74 (7.06)	50.33 (7.15)	50.56 (7.19)	50.31 (7.17)	50.17 (7.11)	50.35 (7.13)
White, %	97	98	98	98	98	98	98	98	97	98	98	98	97	98	97
Current postmenopausal hormone use, %	24	24	26	23	24	26	23	25	26	23	24	27	25	24	25
Current smoker, %	26	24	23	27	23	23	25	23	24	27	23	23	26	24	23
Physical activity (MET-h/wk)	13.61 (19.94)	14.12 (20.00)	14.84 (22.07)	13.53 (19.45)	14.34 (21.05)	15.14 (23.29)	13.93 (20.34)	14.31 (21.82)	14.91 (23.51)	13.25 (20.35)	14.24 (20.24)	15.07 (23.36)	13.88 (20.27)	14.17 (20.86)	14.99 (23.15)
BMI (kg/m^2)	24.77 (4.50)	24.88 (4.54)	24.77 (4.61)	24.80 (4.56)	24.94 (4.60)	24.78 (4.61)	24.73 (4.51)	24.80 (4.59)	24.90 (4.63)	24.90 (4.67)	24.82 (4.57)	24.72 (4.53)	24.56 (4.45)	24.82 (4.59)	25.17 (4.75)
FH diabetes, %	28	28	27	28	29	27	28	28	28	28	28	27	28	28	28
Hypertension, %	8	8	8	7	8	7	8	8	8	8	8	7	8	8	8
Hypercholesterolemia, %	3	3	4	3	3	4	4	3	3	3	3	4	3	3	4
AHEI score	46.91 (10.86)	47.61 (10.49)	48.74 (11.05)	46.50 (10.68)	47.95 (10.63)	48.91 (11.14)	47.05 (10.81)	47.57 (10.91)	48.93 (10.74)	46.14 (10.74)	47.99 (10.63)	48.83 (11.04)	46.74 (10.68)	47.55 (10.79)	49.20 (10.89)
Energy intake (kcal/d)	1530 (493)	1887 (539)	1704 (534)	1561 (509)	1899 (529)	1691 (526)	1636 (541)	1869 (553)	1627 (514)	1638 (537)	1796 (529)	1659 (498)	1705 (562)	1774 (528)	1693 (522)
Alcohol intake (g/d)	9.064 (14.323)	5.645 (9.745)	6.615 (11.193)	7.475 (12.220)	6.395 (11.155)	6.739 (11.301)	7.724 (12.929)	7.137 (12.036)	6.058 (10.145)	7.388 (13.002)	6.767 (10.770)	6.462 (10.641)	9.812 (16.150)	6.714 (10.358)	4.588 (7.483)
Vitamin B2 intake (mg/d)	1.142 (0.150)	2.001 (0.166)	13.30 (11.62)	1.368 (0.481)	1.959 (0.635)	12.45 (12.10)	1.629 (3.230)	4.220 (4.386)	8.158 (12.122)	2.352 (3.408)	3.283 (4.111)	8.170 (11.696)	3.577 (6.145)	4.112 (6.622)	5.082 (7.912)
Vitamin B6 intake (mg/d)	3.662 (17.659)	4.139 (15.504)	27.57 (46.21)	1.298 (0.188)	2.126 (0.192)	38.63 (53.25)	4.521 (19.393)	9.097 (25.650)	17.35 (38.71)	4.348 (18.158)	6.656 (21.910)	21.23 (42.51)	7.857 (25.737)	8.965 (26.859)	12.13 (32.74)

Table 5.2.1. Age-standardized characteristics by quintiles of methyl donor nutrients. (continued)

	B2			B6			B12			Folate			Methionine		
	Q1	Q 3	Q5	Q1	Q3	Q 5	Q1	Q3	Q5	Q1	Q3	Q 5	Q 1	Q3	Q 5
Vitamin B12 intake (mcg/d)	4.646 (2.457)	10.64 (5.28)	20.71 (32.91)	6.414 (4.690)	9.779 (5.558)	20.25 (32.79)	3.432 (0.722)	8.846 (1.496)	25.69 (30.53)	6.839 (4.680)	9.843 (8.195)	20.41 (29.48)	8.362 (17.867)	11.37 (14.38)	15.23 (17.36)
Folate intake (mcg/d)	215.3 (72.1)	318.5 (93.2)	550.1 (329.6)	213.3 (69.6)	323.4 (82.7)	569.2 (326.6)	244.6 (91.9)	397.7 (182.8)	559.2 (309.1)	178.5 (28.1)	302.9 (21.4)	765.1 (214.0)	343.8 (227.0)	379.2 (220.0)	430.8 (253.6)
Methionine intake (g/d)	1.473 (0.334)	1.760 (0.332)	1.736 (0.396)	1.489 (0.302)	1.773 (0.357)	1.741 (0.397)	1.401 (0.283)	1.703 (0.365)	1.838 (0.396)	1.558 (0.347)	1.708 (0.351)	1.739 (0.391)	1.182 (0.135)	1.649 (0.050)	2.215 (0.257)
NHS2															
Age (years)	36.49 (4.65)	36.13 (4.61)	35.83 (4.73)	36.16 (4.67)	36.30 (4.58)	35.83 (4.74)	36.31 (4.61)	36.22 (4.61)	35.77 (4.77)	36.09 (4.69)	36.46 (4.56)	35.37 (4.76)	36.05 (4.75)	36.04 (4.63)	36.32 (4.60)
White, %	95	97	97	96	97	97	97	97	96	95	97	97	96	97	96
Current oc use, %	12	11	9	12	11	8	12	11	10	12	11	8	12	11	10
Current smoker, %	18	10	11	17	10	11	13	12	12	18	10	9	15	12	11
Physical activity (MET-h/wk)	17.68 (24.62)	20.78 (26.75)	23.78 (30.61)	16.22 (23.01)	22.02 (27.30)	23.95 (30.69)	19.90 (26.34)	20.11 (26.96)	22.90 (30.13)	15.77 (22.37)	22.27 (27.59)	23.99 (30.73)	20.98 (29.34)	20.04 (25.89)	22.68 (28.63)
BMI (kg/m^2)	24.56 (5.60)	24.58 (5.14)	24.28 (4.90)	24.61 (5.66)	24.54 (5.14)	24.35 (4.87)	24.14 (5.23)	24.78 (5.39)	24.45 (5.00)	25.06 (5.85)	24.29 (4.98)	24.25 (4.83)	23.65 (5.01)	24.38 (5.05)	25.52 (5.56)
FH diabetes, %	35	34	33	34	34	34	34	35	34	35	34	33	32	34	36
Hypertension, %	7	5	6	7	6	6	6	6	6	7	6	6	5	6	7
Hypercholesterolemia, %	15	13	14	14	14	15	14	14	14	16	13	14	13	14	16
Energy intake (kcal/d)	45.10 (11.01)	47.39 (10.55)	49.81 (11.04)	41.91 (9.76)	49.84 (10.55)	50.39 (11.23)	47.27 (11.52)	47.10 (10.67)	48.84 (10.66)	42.07 (9.61)	49.68 (10.64)	50.36 (10.97)	45.34 (11.98)	46.69 (10.28)	51.67 (10.12)
Alcohol intake (g/d)	1639 (538)	1903 (537)	1715 (535)	1649 (544)	1889 (545)	1740 (549)	1682 (545)	1873 (552)	1704 (539)	1704 (548)	1836 (547)	1684 (514)	1775 (584)	1818 (540)	1736 (526)
Vitamin B2 intake (mg/d)	3.673 (7.157)	2.822 (5.318)	2.844 (5.918)	3.002 (6.173)	3.205 (6.066)	2.986 (6.071)	3.396 (6.756)	3.128 (5.929)	2.755 (5.513)	3.119 (6.521)	3.312 (6.075)	2.616 (5.251)	3.816 (8.005)	3.147 (5.627)	2.414 (4.538)

Table 5.2.1. Age-standardized characteristics by quintiles of methyl donor nutrients. (continued)

	B2			B6			B12			Folate			Methionine		
	Q1	Q 3	Q5	Q1	Q3	Q 5	Q1	Q3	Q5	Q1	Q3	Q 5	Q 1	Q3	Q 5
Vitamin B6 intake (mg/d)	1.403 (0.176)	2.305 (0.172)	11.21 (16.38)	1.614 (0.439)	2.247 (0.643)	10.73 (16.49)	1.713 (1.286)	2.802 (2.301)	9.261 (16.818)	1.982 (2.344)	2.833 (3.299)	8.082 (15.632)	3.899 (9.392)	3.884 (7.999)	4.482 (9.086)
Vitamin B12 intake (mcg/d)	3.505 (15.929)	4.346 (15.846)	22.01 (41.37)	1.589 (0.200)	2.575 (0.225)	31.13 (50.36)	4.131 (17.737)	5.803 (19.642)	17.73 (37.83)	3.776 (16.011)	6.267 (21.415)	16.72 (37.05)	8.159 (26.111)	7.696 (23.903)	9.333 (26.791)
Folate intake (mcg/d)	4.653 (1.844)	7.952 (4.611)	19.62 (24.52)	5.436 (3.504)	7.499 (4.407)	18.69 (24.61)	3.495 (0.711)	7.427 (0.495)	22.72 (24.08)	5.714 (3.841)	7.689 (6.703)	17.92 (22.64)	8.366 (13.314)	9.752 (14.294)	11.55 (13.26)
Methionine intake (g/d)	258.3 (76.4)	371.1 (103.2)	828.9 (371.6)	249.8 (67.8)	384.1 (94.9)	825.3 (383.5)	294.6 (101.3)	394.5 (174.6)	810.1 (371.9)	217.7 (32.8)	374.8 (30.8)	951.3 (273.4)	460.9 (298.2)	475.3 (283.5)	502.9 (304.4)
Energy intake (kcal/d)	1.852 (0.421)	2.115 (0.402)	2.089 (0.450)	1.784 (0.343)	2.149 (0.433)	2.083 (0.462)	1.780 (0.417)	2.115 (0.405)	2.134 (0.451)	1.969 (0.441)	2.049 (0.413)	2.069 (0.444)	1.470 (0.195)	2.014 (0.057)	2.649 (0.286)
HPFS															
Age	51.40 (9.02)	53.09 (9.64)	53.88 (9.60)	50.99 (9.02)	53.72 (9.67)	54.12 (9.61)	51.81 (9.47)	52.85 (9.41)	53.86 (9.43)	51.70 (9.17)	53.23 (9.57)	54.11 (9.54)	52.93 (9.75)	52.69 (9.45)	53.73 (9.36)
White, %	93	96	95	95	95	95	94	96	95	94	95	95	94	96	95
Current smoker, %	11	9	9	14	6	9	8	9	11	14	7	8	11	9	8
Physical activity (MET-h/wk)	18.53 (25.84)	21.37 (29.79)	23.64 (31.12)	16.50 (22.81)	22.92 (33.03)	24.31 (32.34)	21.63 (28.98)	21.07 (27.48)	22.29 (30.98)	16.29 (25.46)	22.85 (32.87)	25.04 (33.36)	21.49 (32.41)	20.99 (26.29)	21.96 (30.83)
BMI (kg/m^2)	25.58 (3.24)	25.47 (3.27)	25.11 (3.14)	25.65 (3.32)	25.43 (3.28)	25.09 (3.13)	25.18 (3.14)	25.55 (3.40)	25.53 (3.48)	25.82 (3.51)	25.42 (3.21)	25.08 (3.24)	25.03 (3.16)	25.40 (3.11)	25.93 (3.58)
FH diabetes, %	19	21	21	19	20	20	19	20	20	19	20	20	19	21	21
Hypertension, %	21	17	20	18	19	20	20	18	19	20	18	20	18	19	21
Hypercholesterolemia, %	11	8	12	8	10	12	12	9	10	9	10	12	9	10	12
Energy intake (kcal/d)	50.87 (11.67)	51.55 (11.06)	55.05 (11.72)	47.07 (10.39)	54.71 (10.90)	56.11 (11.69)	53.10 (12.08)	51.40 (11.26)	53.86 (11.15)	46.61 (10.52)	53.96 (10.83)	56.18 (11.41)	48.99 (11.55)	52.51 (11.26)	56.04 (10.95)

Table 5.2.1. Age-standardized characteristics by quintiles of methyl donor nutrients. (continued)

	B2			B6			B12			Folate			Methionine		
	Q1	Q 3	Q5	Q1	Q3	Q 5	Q1	Q3	Q5	Q1	Q3	Q 5	Q 1	Q3	Q 5
Alcohol intake (g/d)	1770 (553)	2152 (639)	1964 (616)	1776 (572)	2138 (632)	1944 (605)	1921 (605)	2087 (622)	1908 (601)	1893 (615)	2062 (621)	1921 (584)	1971 (631)	2029 (613)	1962 (630)
Vitamin B2 intake (mg/d)	14.16 (17.88)	9.883 (13.976)	11.74 (15.87)	12.19 (16.31)	10.80 (14.49)	11.50 (15.46)	12.55 (17.22)	11.41 (15.17)	10.38 (14.11)	13.07 (17.95)	10.80 (14.26)	10.78 (14.24)	16.80 (21.65)	10.97 (13.84)	7.475 (9.870)
Vitamin B6 intake (mg/d)	1.417 (0.173)	2.389 (0.200)	16.62 (18.28)	1.649 (1.270)	2.355 (0.746)	15.93 (18.71)	2.260 (4.414)	3.786 (4.009)	11.61 (19.62)	2.993 (4.583)	3.704 (4.532)	10.95 (18.79)	4.619 (8.948)	5.028 (9.136)	6.045 (12.234)
Vitamin B12 intake (mcg/d)	2.316 (7.413)	3.412 (9.463)	28.05 (45.88)	1.583 (0.200)	2.674 (0.222)	32.74 (48.66)	3.915 (14.244)	5.964 (17.949)	19.96 (41.17)	3.462 (11.765)	5.466 (16.867)	22.21 (43.17)	7.639 (23.161)	8.139 (23.114)	11.19 (31.45)
Folate intake (mcg/d)	6.371 (2.675)	11.45 (6.34)	22.53 (36.02)	7.859 (4.772)	10.06 (5.62)	22.77 (35.77)	4.948 (1.137)	9.468 (0.499)	29.30 (35.58)	8.601 (5.515)	10.37 (6.85)	21.81 (33.22)	10.25 (12.82)	12.42 (12.50)	15.27 (25.59)
Methionine intake (g/d)	290.5 (92.1)	408.6 (114.1)	691.1 (410.6)	273.8 (81.9)	418.2 (98.5)	743.1 (403.8)	344.4 (119.8)	403.1 (163.1)	725.3 (401.3)	235.6 (35.6)	388.1 (24.1)	930.7 (275.9)	451.7 (269.3)	474.8 (262.5)	510.5 (304.1)
Energy intake (kcal/d)	2.096 (0.401)	2.317 (0.376)	2.291 (0.426)	2.028 (0.337)	2.350 (0.391)	2.305 (0.428)	2.066 (0.383)	2.282 (0.371)	2.369 (0.428)	2.169 (0.430)	2.277 (0.385)	2.282 (0.417)	1.710 (0.173)	2.229 (0.055)	2.822 (0.276)

Table 5.2.2. Hazard ratios (95% CIs) for type 2 diabetes according to quintiles of vitamin B2

	Q1	Q2	Q3	Q4	Q5
NHS					
Total vitamin B2(mcg/d)					
Median	1.5	2	2.7	4.5	10.6
Cases/person-years	1,728/347,616	1,727/357,685	1,662/352,678	1,569/352,563	1,455/352,887
Model 1	1	0.98 (0.91, 1.04)	0.94(0.87, 1.00)	**0.88 (0.83, 0.95)**	**0.81 (0.76, 0.87)**
Model 2	1	**0.93 (0.87, 0.99)**	**0.92 (0.84, 0.98)**	**0.90 (0.84, 0.97)**	**0.87 (0.81, 0.93)**
Model 3	1	0.95 (0.89, 1.02)	0.95 (0.88, 1.01)	0.94 (0.87, 1.01)	**0.90 (0.84, 0.97)**
Vitamin B2 from food (mcg/d)					
Median	1.2	1.4	1.6	1.8	2.2
Cases/person-years	1,749/347,085	1,680/355,479	1,593/346,420	1,581/363,583	1,538/350,861
Model 4	**1**	**0.91 (0.85,0.98)**	**0.91 (0.85, 0.98)**	**0.93 (0.87, 0.996)**	**0.85 (0.79, 0.92)**
Model 5	1	**0.93 (0.86, 0.99)**	**0.93 (0.86, 0.999)**	0.95 (0.88, 1.03)	**0.88 (0.81, 0.95)**
Vitamin B2 from supplements (mcg/d)					
Median	0	0.1	0.3	1.7	9.3
Cases/person-years	2,075/389,974	1,707/322,319	1,478/344,911	1,498/353,500	1,383/352,725
Model 6	1	0.94 (0.88, 1.01)	0.98 (0.92, 1.05)	0.96 (0.89, 1.03)	**0.92 (0.89, 0.99)**
Model 7	1	0.96 (0.89, 1.03)	0.995 (0.93, 1.07)	0.97 (0.90, 1.03)	**0.91 (0.85, 0.98)**
NHS II					
Total vitamin B2(mcg/d)					
Median	1.6	2.1	2.7	3.9	16
Cases/person-years	1,462/361,387	1,209/364,647	1,130/363,958	1,080/364,306	1,168/365,126
Model 1	1	**0.84 (0.78, 0.91)**	**0.79 (0.73, 0.86)**	**0.77 (0.71, 0.83)**	**0.78 (0.72, 0.84)**
Model 2	1	**0.90 (0.83, 0.96)**	**0.89 (0.82, 0.96)**	**0.90 (0.84, 0.98)**	**0.93 (0.85, 1.00)**
Model 3	1	**0.92 (0.85, 0.99)**	**0.92 (0.85, 0.99)**	0.93 (0.86, 1.01)	0.96 (0.88, 1.03)
Vitamin B2 from food (mcg/d)					
Median	1.4	1.7	1.9	2.1	2.5
Cases/person-years	1,515/360,779	1,283/365,758	1,229/364,291	1,112/364,686	910/363,910
Model 4	1	**0.90 (0.83, 0.97)**	**0.90 (0.84, 0.98)**	**0.85 (0.78, 0.92)**	**0.78 (0.72, 0.85)**
Model 5	1	**0.90 (0.84, 0.98)**	**0.91 (0.78, 0.93)**	**0.85 (0.78, 0.93)**	**0.78 (0.71, 0.86)**
Vitamin B2 from supplements (mcg/d)					
Median	0	0.2	0.7	1.6	10.1
Cases/person-years	1,389/407,950	1,127/317,401	1,200/364,806	1,144/364,202	1,189/365,064
Model 6	1	**0.87 (0.80, 0.94)**	0.93 (0.86, 1.01)	0.95 (0.88, 1.03)	0.98 (0.90, 1.06)
Model 7	1	**0.87 (0.80, 0.95)**	0.93 (0.86, 1.01)	0.95 (0.88, 1.03)	0.97 (0.89, 1.05)
HPFS					
Total vitamin B2 (mcg/d)					
Median	1.57	2.06	2.69	4.13	13.01
Cases/person-years	729/173,610	752/174,265	655/174,658	681/174,346	622/174,579
Model 1	1	1.01 (0.91, 1.12)	**0.87 (0.78, 0.97)**	**0.90 (0.81, 0.998)**	**0.82 (0.73, 0.91)**

Table 5.2.2. Hazard ratios (95% CIs) for type 2 diabetes according to quintiles of vitamin B2 (continued)

	Q1	Q2	Q3	Q4	Q5
Model 2	1	1.02 (0.92, 1.13)	**0.90 (0.80, 0.998)**	0.94 (0.85, 1.05)	**0.88 (0.79, 0.99)**
Model 3	1	1.05 (0.94, 1.16)	0.92 (0.82, 1.03)	0.97 (0.87, 1.08)	0.91 (0.82, 1.02)
Vitamin B2 from food (mcg/d)					
Median	1.47	1.73	1.94	2.18	2.59
Cases/person-years	717/173,562	656/174,986	707/174,544	709/174,255	650/174,109
Model 4	1	0.90 (0.81,0.999)	**0.96 (0.86, 1.06)**	0.94 (0.85, 1.05)	**0.84 (0.75, 0.94)**
Model 5	1	0.91 (0.81, 1.01)	0.97 (0.87, 1.09)	0.95 (0.85, 1.07)	**0.84 (0.75, 0.95)**
Vitamin B2 from supplements (mcg/d)					
Median	1.69	1.99	2.22	2.47	2.92
Cases/person-years	781/200,534	665/146,833	707/175,029	668/174,460	618/174,601
Model 6	1	1.11 (0.98, 1.25)	1.04 (0.93, 1.17)	1.00 (0.90, 1.12)	0.97 (0.87, 1.08)
Model 7	1	1.12 (0.999, 1.26)	1.07 (0.95, 1.19)	1.02 (0.91, 1.14)	0.97 (0.87, 1.08)
Pooled					
Total vitamin B2(mcg/d)					
Model 1	1	**0.93 (0.89, 0.98)***	**0.87 (0.83, 0.91)***	**0.84 (0.81, 0.88)***	**0.80 (0.77, 0.84)**
Model 2	1	**0.94 (0.89, 0.98)**	**0.90 (0.86, 0.95)**	**0.91 (0.87, 0.95)**	**0.89 (0.85, 0.94)**
Model 3	1	0.96 (0.91, 1.00)	**0.93 (0.89, 0.98)**	**0.94 (0.90, 0.99)**	**0.92 (0.88, 0.97)**
Vitamin B2 from food (mcg/d)					
Model 4	1	0.91 (0.86, 0.95)	**0.92 (0.87, 0.96)**	**0.90 (0.86, 0.94)**	**0.82 (0.79, 0.87)**
Model 5	1	**0.91 (0.87, 0.96)**	**0.93 (0.89, 0.98)**	**0.92 (0.87, 0.96)**	**0.84 (0.79, 0.88)**
Vitamin B2from supplements (mcg/d)					
Model 6	1	0.94 (0.90, 0.99)*	0.97 (0.93, 1.02)	0.97 (0.92, 1.01)	0.95 (0.91, 1.00)
Model 7	1	0.95 (0.91, 1.00)*	0.99 (0.94, 1.03)	0.97 (0.93, 1.02)	**0.94 (0.90, 0.99)**

Model 1: adjusted for age (continuously).
Model 2: model 1 + ethnicity (caucasian, black, or other), smoking (never, past, current: 1-14, 15-24, or ≥25 cigarettes/day), marital status (currently married, widowed, divorced or separated, or never married), family history of type 2 diabetes (yes/no), hypertension(yes/no), hypercholesterolemia (yes/no), postmenopausal status and hormone use(premenopausal, or if postmenopausal, never, current, or past postmenopausal hormone use), total energy intake (quintiles), level of physical activity (<3, 3-9, 9-18, 18-27, or ≥27 MET-h/week), alcohol consumption (0, 0.1-5, 5-10, 10-15, or ≥15 g/day), and BMI (<21, 21-23, 23-25, 25-27, 27-30, 30-33, 33-35, 35-40, or ≥40 kg/m2).
Model 3: model 2 + intake of cereal fiber (quintiles), animal protein (quintiles), and PUFA:SFA ratio (quintiles).
Model 4: model 2 + multivitamin use (yes/no).
Model 5: model 3 + multivitamin use (yes/no).
Model 6: model 2 + AHEI index: (quintiles).
Model 7: model 3 + AHEI index: (quintiles).
** p-Value for Q-statistic <0.05, indicating statistically significant heterogeneity among the three cohorts*

Table 5.2.3. Hazard ratios (95% CIs) for type 2 diabetes according to quintiles of vitamin B6

	Q1	Q2	Q3	Q4	Q5
NHS					
Total vitamin B6(mcg/d)					
Median	1.6	2.3	3.1	4.8	24.4
Cases/person-years	1,777/356,303	1,702/349,303	1,649/350,357	1,521/353,413	1,492/353,641
Model 1	1	0.94 (0.88, 1.01)	**0.91 (0.85, 0.97)**	**0.83 (0.78, 0.89)**	**0.81 (0.76, 0.87)**
Model 2	1	**0.93 (0.86, 0.99)**	**0.94 (0.88, 1.01)**	**0.90 (0.84, 0.96)**	**0.87 (0.81, 0.94)**
Model 3	1	0.94 (0.88, 1.01)	0.97 (0.90, 1.04)	**0.93 (0.87, 0.999)**	**0.91 (0.84, 0.97)**
Vitamin B6 from food (mcg/d)					
Median	1.3	1.6	1.7	1.9	2.3
Cases/person-years	1,749/347,085	1,680/355,479	1,593/346,420	1,581/363,583	1,538/350,861
Model 4	1	0.98 (0.91, 1.05)	0.98 (0.91, 1.05)	0.98 (0.92, 1.06)	**0.92 (0.86, 0.99)**
Model 5	1	0.99 (0.92, 1.06)	0.995 (0.92, 1.07)	1.02 (0.95, 1.10)	0.98 (0.90, 1.06)
Vitamin B6 from supplements (mcg/d)					
Median	0	0.1	0.4	2	16
Cases/person-years	2,016/379,056	1,703/323,397	1,557/354,942	1,456/353,018	1,409/353,015
Model 6	1	0.95 (0.88, 1.02)	1.01 (0.94, 1.08)	0.94 (0.88, 1.01)	0.94 (0.87, 1.01)
Model 7	1	0.96 (0.89, 1.03)	1.02 (0.95, 1.09)	0.95 (0.88, 1.02)	**0.93 (0.86, 0.99)**
NHS II					
Total vitamin B6(mcg/d)					
Median	1.8	2.3	3.1	4.6	23.6
Cases/person-years	1,352/355,812	1,201/368,418	1,189/368,953	1,109/360,367	1,198/365,874
Model 1	1	**0.88 (0.81, 0.95)**	**0.86 (0.80, 0.93)**	**0.83 (0.76, 0.90)**	**0.85 (0.79, 0.90)**
Model 2	1	**0.92 (0.85, 0.999)**	0.94 (0.87, 1.02)	0.94 (0.87, 1.02)	0.98 (0.90, 1.06)
Model 3	1	0.93 (0.86, 1.01)	0.96 (0.88, 1.04)	0.96 (0.88, 1.04)	1.00 (0.92, 1.08)
Vitamin B6 from food (mcg/d)					
Median	1.6	1.9	2.1	2.3	2.6
Cases/person-years	1,320/363,342	1,197/363,468	1,209/363,681	1,180/364,435	1,143/364,497
Model 4	1	0.96 (0.88, 1.03)	0.995 (0.91, 1.08)	0.94 (0.86, 1.01)	0.97 (0.89, 1.05)
Model 5	1	0.95 (0.88, 1.03)	0.995 (0.91, 1.08)	0.94 (0.86, 1.03)	0.98 (0.89, 1.07)
Vitamin B6 from supplements (mcg/d)					
Median	0	0.3	1	2.3	21.6
Cases/person-years	1,382/404,010	1,143/324,332	1,236/361,770	1,101/364,168	1,187/365,144
Model 6	1	0.88 (0.81, 0.96)	0.94 (0.87, 1.02)	0.92 (0.85, 1.00)	0.98 (0.90, 1.06)
Model 7	1	**0.88 (0.81, 0.96)**	0.94 (0.87, 1.02)	**0.92 (0.85, 0.995)**	0.96 (0.89, 1.04)
HPFS					
Total vitamin B6 (mcg/d)					
Median	1.8	2.3	3.1	4.6	18.9
Cases/person-years	757/168,808	698/179,817	680/173,644	690/174,365	614/174,822
Model 1	1	**0.87 (0.78, 0.96)**	**0.85 (0.76, 0.94)**	**0.85 (0.77, 0.95)**	**0.75 (0.68, 0.84)**

Table 5.2.3. Hazard ratios (95% CIs) for type 2 diabetes according to quintiles of vitamin B6 (continued)

	Q1	Q2	Q3	Q4	Q5
Model 2	1	**0.89 (0.80, 0.98)**	**0.89 (0.80, 0.99)**	0.92 (0.83, 1.03)	**0.84 (0.75, 0.93)**
Model 3	1	**0.90(0.81, 0.998)**	0.92 (0.82, 1.02)	0.95 (0.85, 1.06)	**0.87 (0.77, 0.97)**
Vitamin B6 from food (mcg/d)					
Median	1.7	2	2.2	2.5	2.9
Cases/person-years	744/173,742	736/174,214	671/174,699	630/175,051	658/173,751
Model 4	1	0.98 (0.88, 1.08)	**0.90 (0.81, 0.995)**	**0.86 (0.77, 0.96)**	0.93 (0.84, 1.04)
Model 5	1	0.97 (0.88, 1.08)	**0.90 (0.80, 1.00)**	**0.87 (0.77, 0.98)**	**0.96 (0.85, 1.08)**
Vitamin B6 from supplements (mcg/d)					
Median	0	0.1	0.6	2.2	16.3
Cases/person-years	783/197,397	682/149,920	692/175,026	663/174,321	619/174,792
Model 6	1	1.09 (0.97, 1.22)	1.00 (0.90, 1.12)	0.97 (0.87, 1.08)	0.95 (0.85, 1.06)
Model 7	1	1.10 (0.98, 1.23)	1.02 (0.91, 1.14)	0.96 (0.88, 1.10)	0.95 (0.85, 1.06)
Pooled					
Total vitamin B6(mcg/d)					
Model 1	1	**0.91 (0.87, 0.95)**	**0.88 (0.84, 0.92)**	**0.84 (0.80, 0.87)**	**0.81 (0.78, 0.85)**
Model 2	1	**0.92 (0.88, 0.96)**	**0.93 (0.89, 0.97)**	**0.92 (0.88, 0.96)**	**0.90 (0.86, 0.95)***
Model 3	1	**0.93 (0.89, 0.98)**	0.96 (0.91, 1.00)	**0.95 (0.90, 0.99)**	**0.93 (0.89, 0.98)**
Vitamin B6 from food (mcg/d)					
Model 4	1	0.97 (0.93, 1.02)	0.97 (0.92, 1.01)	**0.94 (0.90, 0.99)**	**0.94 (0.89, 0.99)**
Model 5	1	0.97 (0.93, 1.02)	0.98 (0.93, 1.03)	0.96 (0.91, 1.01)	0.97 (0.92, 1.03)
Vitamin B6 from supplements (mcg/d)					
Model 6	1	**0.95 (0.90, 0.99)***	0.98 (0.94, 1.03)	**0.94 (0.90, 0.99)**	0.95 (0.91, 1.00)
Model 7	1	0.95 (0.91, 1.00)*	0.99 (0.94, 1.04)	**0.94 (0.90, 0.99)**	**0.94 (0.90, 0.99)**

Model 1: adjusted for age (continuously).
Model 2: model 1 + ethnicity (caucasian, black, or other), smoking (never, past, current: 1-14, 15-24, or ≥25 cigarettes/day), marital status (currently married, widowed, divorced or separated, or never married), family history of type 2 diabetes (yes/no), hypertension(yes/no), hypercholesterolemia (yes/no), postmenopausal status and hormone use(premenopausal, or if postmenopausal, never, current, or past postmenopausal hormone use), total energy intake (quintiles), level of physical activity (<3, 3-9, 9-18, 18-27, or ≥27 MET-h/week), alcohol consumption (0, 0.1-5, 5-10, 10-15, or ≥15 g/day), and BMI (<21, 21-23, 23-25, 25-27, 27-30, 30-33, 33-35, 35-40, or ≥40 kg/m2).
Model 3: model 2 + intake of cereal fiber (quintiles), animal protein (quintiles), and PUFA:SFA ratio (quintiles).
Model 4: model 2 + multivitamin use (yes/no).
Model 5: model 3 + multivitamin use (yes/no).
Model 6: model 2 + AHEI index: (quintiles).
Model 7: model 3 + AHEI index: (quintiles).
** p-Value for Q-statistic <0.05, indicating statistically significant heterogeneity among the three cohorts*

Table 5.2.4. Hazard ratios (95% CIs) for type 2 diabetes according to quintiles of vitamin B12

	Q1	Q2	Q3	Q4	Q5
NHS					
Total vitamin B12(mcg/d)					
Median	5	7.7	10	13	21
Cases/person-years	1,543/353,584	1,582/346,079	1,649/360,818	1,715/351,210	1,652/351,736
Model 1	1	1.04 (0.97, 1.12)	1.04 (0.97, 1.12)	**1.10 (1.02, 1.17)**	1.04 (0.97, 1.12)
Model 2	1	1.00 (0.93, 1.07)	1.01 (0.94, 1.08)	1.05 (0.98, 1.12)	1.00 (0.94, 1.08)
Model 3	1	1.00 (0.93, 1.08)	1.01 (0.94, 1.09)	1.05 (0.98, 1.13)	1.01 (0.94, 1.09)
Vitamin B12 from food (mcg/d)					
Median	3.6	4.8	6	8	13.2
Cases/person-years	1,415/351,919	1,519/356,411	1,552/350,468	1,519/353,045	2,136/351,585
Model 4	1	0.98 (0.91, 1.06)	1.06 (0.99, 1.14)	**1.08 (1.01, 1.16)**	**1.12 (1.04, 1.20)**
Model 5	1	0.98 (0.91, 1.06)	1.06 (0.98, 1.14)	**1.08 (1.00, 1.17)**	**1.11 (1.02, 1.19)**
Vitamin B12 from supplements (mcg/d)					
Median	0	0.2	0.9	3.9	8.6
Cases/person-years	1,946/365,271	1,682/338,434	1,736/353,930	1,447/352,380	1,330/353,414
Model 6	1	0.97 (0.90, 1.04)	0.97 (0.91, 1.04)	0.97 (0.91, 1.04)	**0.92 (0.86, 0.99)**
Model 7	1	0.97 (0.91, 1.04)	0.98 (0.91, 1.05)	0.98 (0.91, 1.05)	**0.92 (0.85, 0.98)**
NHS II					
Total vitamin B12(mcg/d)					
Median	4.5	6.4	8.5	12	24
Cases/person-years	1,182/361,038	1,197/376,576	1,192/349,340	1,253/367,927	1,225/364,543
Model 1	1	1.03 (0.94, 1.12)	1.04 (0.96, 1.13)	1.04 (0.96, 1.12)	0.99 (0.91, 1.07)
Model 2	1	0.98 (0.90, 1.06)	0.999 (0.92, 1.08)	0.997 (0.92, 1.08)	0.99 (0.91, 1.07)
Model 3	1	0.98 (0.90, 1.06)	1.00 (0.92, 1.09)	1.00 (0.92, 1.09)	0.99 (0.91, 1.08)
Vitamin B12 from food (mcg/d)					
Median	3.7	4.7	5.6	6.6	9.2
Cases/person-years	1,050/365,697	1,104/363,186	1,213/357,920	1,321/369,028	1,361/363,591
Model 4	1	0.94 (0.86, 1.02)	0.99 (0.91, 1.08)	1.01 (0.93, 1.09)	1.01 (0.93, 1.10)
Model 5	1	0.93 (0.85, 1.02)	0.97 (0.89, 1.06)	0.98 (0.90, 1.08)	0.98 (0.89, 1.07)
Vitamin B12 from supplements (mcg/d)					
Median	0	0.7	2.2	5	15.9
Cases/person-years	1,381/407,570	1,108/309,418	1,162/373,191	1,200/363,351	1,198/365,894
Model 6	1	**0.91 (0.84, 0.99)**	0.95 (0.88, 1.03)	0.98 (0.91, 1.06)	0.99 (0.91, 1.06)
Model 7	1	**0.92 (0.84, 0.99)**	0.96 (0.88, 1.04)	0.98 (0.91, 1.06)	0.97 (0.90, 1.05)
HPFS					
Total vitamin B12 (mcg/d)					
Median	5.5	8	10.5	14.5	25.6
Cases/person-years	620/178,826	701/170,745	688/168,366	771/179,272	659/174,247
Model 1	1	**1.15 (1.03, 1.28)**	1.11 (0.99, 1.24)	**1.18 (1.06, 1.31)**	1.02 (0.91, 1.14)

Table 5.2.4. Hazard ratios (95% CIs) for type 2 diabetes according to quintiles of vitamin B12 (continued)

	Q1	Q2	Q3	Q4	Q5
Model 2		1.07 (0.96, 1.19)	1.03 (0.93, 1.15)	1.10 (0.99, 1.22)	0.95 (0.85, 1.08)
Model 3		1.05 (0.94, 1.17)	1.02 (0.91, 1.14)	1.08 (0.97, 1.20)	0.93 (0.83, 1.04)
Vitamin B12 from food (mcg/d)					
Median	4.5	6	7.3	9.1	13.8
Cases/person-years	717/173,562	656/174,986	707/174,544	709/174,255	650/174,109
Model 4	1	1.11 (0.99, 1.25)	1.08 (0.96,1.21)	**1.19 (1.07, 1.33)**	1.09 (0.97, 1.22)
Model 5	1	1.10 (0.98, 1.24)	1.06 (0.94, 1.19)	**1.15 (1.02, 1.30)**	1.03 (0.91, 1.16)
Vitamin B12 from supplements (mcg/d)					
Median	0	0.3	1.4	4.9	13.9
Cases/person-years	775/194,562	647/152,239	700/174,114	691/175,592	626/174,950
Model 6	1	0.998 (0.89, 1.12)	0.997 (0.89, 1.11)	1.01 (0.91, 1.13)	0.95 (0.85, 1.05)
Model 7	1	1.01 (0.90, 1.12)	1.00 (0.90, 1.12)	1.03 (0.92, 1.14)	0.95 (0.85, 1.06)
Pooled					
Total vitamin B12(mcg/d)					
Model 1	1	**1.06 (1.01, 1.11)**	**1.05 (1.01, 1.11)**	**1.09 (1.04, 1.14)**	1.02 (0.97, 1.07)
Model 2	1	**1.01 (0.96, 1.05)**	**1.01 (0.96, 1.06)**	**1.04 (0.99, 1.09)**	**0.99 (0.94, 1.04)**
Model 3	1	1.00 (0.96, 1.05)	**1.01 (0.96, 1.06)**	**1.04 (0.99, 1.09)**	**0.99 (0.94, 1.04)**
Vitamin B12 from food (mcg/d)					
Model 4	1	0.99 (0.94, 1.04)	1.04 (0.99, 1.09)	**1.08 (1.02, 1.13)***	**1.07 (1.02, 1.13)**
Model 5	1	0.98 (0.93, 1.04)	1.03 (0.97, 1.08)	**1.06 (1.00, 1.12)**	1.05 (0.99, 1.10)
Vitamin B12 from supplements (mcg/d)					
Model 6	1	**0.95 (0.91, 1.00)**	0.97 (0.93, 1.02)	0.98 (0.94, 1.03)	**0.95 (0.91, 1.00)**
Model 7	1	0.96 (0.91, 1.01)	0.98 (0.93, 1.02)	0.99 (0.94, 1.04)	**0.94 (0.90, 0.99)**

Model 1: adjusted for age (continuously).
Model 2: model 1 + ethnicity (caucasian, black, or other), smoking (never, past, current: 1-14, 15-24, or ≥25 cigarettes/day), marital status (currently married, widowed, divorced or separated, or never married), family history of type 2 diabetes (yes/no), hypertension(yes/no), hypercholesterolemia (yes/no), postmenopausal status and hormone use(premenopausal, or if postmenopausal, never, current, or past postmenopausal hormone use), total energy intake (quintiles), level of physical activity (<3, 3-9, 9-18, 18-27, or ≥27 MET-h/week), alcohol consumption (0, 0.1-5, 5-10, 10-15, or ≥15 g/day), and BMI (<21, 21-23, 23-25, 25-27, 27-30, 30-33, 33-35, 35-40, or ≥40 kg/m2).
Model 3: model 2 + intake of cereal fiber (quintiles), animal protein (quintiles), and PUFA:SFA ratio (quintiles).
Model 4: model 2 + multivitamin use (yes/no).
Model 5: model 3 + multivitamin use (yes/no).
Model 6: model 2 + AHEI index: (quintiles).
Model 7: model 3 + AHEI index: (quintiles).
** p-Value for Q-statistic <0.05, indicating statistically significant heterogeneity among the three cohorts*

Table 5.2.5. Hazard ratios (95% CIs) for type 2 diabetes according to quintiles of folate

	Q1	Q2	Q3	Q4	Q5
NHS					
Total folate (mcg/d)					
Median	212	284	363	474	659
Cases/person-years	1,840/350,315	1,770/352,055	1,598/353,694	1,530/353,706	1,403/353,659
Model 1	1	0.94 (0.88, 1.00)	**0.84 (0.78, 0.90)**	**0.80 (0.75, 0.86)**	**0.73 (0.68, 0.78)**
Model 2	1	0.996 (0.93, 1.06)	**0.94 (0.87, 1.00)**	**0.91 (0.85, 0.98)**	**0.89 (0.83, 0.96)**
Model 3	1	1.03 (0.96, 1.10)	0.98 (0.91, 1.05)	0.96 (0.90, 1.03)	0.95 (0.88, 1.02)
Folate from food (mcg/d)					
Median	190.7	239.5	277.5	230	393.7
Cases/person-years	2,007/350,888	1,715/352,983	1,626/352,631	1,498/353,277	1,295/353,649
Model 4	1	**0.996 (0.93, 1.07)**	**0.995 (0.93, 1.07)**	**0.97 (0.90, 1.04)**	**0.92 (0.85, 0.99)**
Model 5	1	1.03 (0.96, 1.10)	1.05 (0.97, 1.12)	1.03 (0.96, 1.11)	1.00 (0.93, 1.08)
Folate from supplements (mcg/d)					
Median	0	4.5	18.5	169.5	379.6
Cases/person-years	1,932/376,167	1,819/323,703	1,723/357,408	1,189/353,242	1,478/352,908
Model 6	1	0.98 (0.91, 1.05)	0.97 (0.90, 1.04)	**0.91 (0.85, 0.98)**	0.94 (0.87, 1.01)
Model 7	1	0.99 (0.92, 1.06)	0.98 (0.92, 1.06)	**0.92 (0.86, 0.99)**	0.94 (0.87, 1.01)
NHS II					
Total folate (mcg/d)					
Median	239.0	318.0	415.0	575.0	821.0
Cases/person-years	1,563/361,937	1,219/363,705	1,161/364,140	1,164/364,834	942/364,808
Model 1	1	**0.77 (0.71, 0.83)**	**0.73 (0.68, 0.79)**	**0.74 (0.69, 0.80)**	**0.63 (0.58, 0.68)**
Model 2	1	**0.92 (0.86, 0.995)**	**0.93 (0.86, 0.999)**	**0.95 (0.88, 1.02)**	**0.88 (0.81, 0.96)**
Model 3	1	0.95 (0.88, 1.02)	0.96 (0.88, 1.04)	0.99 (0.91, 1.07)	0.92 (0.85, 1.00)
Folate from food (mcg/d)					
Median	231.3	290.3	331.7	377.7	465.1
Cases/person-years	1,665/360,362	1,325/363,925	1,188/365,151	1,073/364,235	798/365,751
Model 4	1	**0.94 (0.88, 1.01)**	**0.95 (0.88, 1.02)**	**0.93 (0.86, 1.01)**	**0.84 (0.77, 0.92)**
Model 5	1	0.97 (0.90, 1.04)	0.99 (0.91, 1.07)	0.99 (0.91, 1.08)	0.91 (0.82, 1.00)
Folate from supplements (mcg/d)					
Median	0	2	28.8	142.2	356
Cases/person-years	1,271/418,391	1,288/312,310	1,221/361,331	1,219/365,171	1,050/362,221
Model 6	1	0.94 (0.86, 1.03)	0.997 (0.91, 1.09)	**1.01 (0.93, 1.10)**	0.97 (0.87, 1.10)
Model 7	1	0.95 (0.86, 1.03)	1.00 (0.92, 1.10)	**1.01 (0.93, 1.10)**	0.96 (0.88, 1.05)
HPFS					
Total folate (mcg/d)					
Median	272	360.7	451.5	606.7	862.7
Cases/person-years	784/172,569	728/174,366	675/174,716	628/175,143	624/174,662
Model 1	1	0.91 (0.82, 1.00)	**0.83 (0.75, 0.92)**	**0.77 (0.69, 0.85)**	**0.75 (0.68, 0.83)**

Table 5.2.5. Hazard ratios (95% CIs) for type 2 diabetes according to quintiles of folate (continued)

	Q1	Q2	Q3	Q4	Q5
Model 2	1	0.98 (0.88, 1.08)	0.94 (0.85, 1.05)	**0.88 (0.79, 0.98)**	0.92 (0.83, 1.03)
Model 3	1	1.02 (0.92, 1.13)	0.999 (0.90, 1.11)	0.94 (0.84, 1.05)	0.99 (0.89, 1.11)
Folate from food (mcg/d)					
Median	253	313.6	359.2	412.4	508
Cases/person-years	819/172,813	724/174,098	652/175,009	677/174,510	567/175,026
Model 4	1	0.94 (0.85, 1.04)	**0.86 (0.77, 0.95)**	0.93 (0.84, 1.03)	**0.85 (0.76, 0.95)**
Model 5	1	0.96 (0.87, 1.07)	0.91 (0.81, 1.01)	0.998 (0.90, 1.11)	0.93 (0.83, 1.05)
Folate from supplements (mcg/d)					
Median	0	8.58	38.2	216	435.4
Cases/person-years	823/217,120	591/128,956	724/176,160	655/174,906	646/174,314
Model 6	1	1.07 (0.95, 1.21)	1.10 (0.98, 1.23)	1.03 (0.92, 1.15)	1.01 (0.91, 1.13)
Model 7	1	1.09 (0.96, 1.23)	1.12 (0.98, 1.25)	1.05 (0.94, 1.17)	1.02 (0.91, 1.13)
Pooled					
Total folate (mcg/d)					
Model 1	1	**0.87 (0.83, 0.91)***	**0.80 (0.76, 0.83)***	**0.77 (0.74, 0.81)**	**0.70 (0.66, 0.73)***
Model 2	1	**0.97 (0.92, 1.01)**	**0.93 (0.89, 0.98)**	**0.92 (0.88, 0.96)**	**0.90 (0.85, 0.94)**
Model 3	1	1.00 (0.95, 1.04)	0.98 (0.93, 1.02)	0.97 (0.92, 1.01)	0.95 (0.90, 1.00)
Folate from food (mcg/d)					
Model 4	1	0.97 (0.92, 1.01)	0.95 (0.91, 1.00)	**0.95 (0.90, 0.99)**	**0.88 (0.84, 0.92)**
Model 5	1	0.99 (0.95, 1.04)	1.00 (0.95, 1.05)	1.01 (0.96, 1.06)	0.96 (0.91, 1.01)
Folate from supplements (mcg/d)					
Model 6	1	0.98 (0.93, 1.03)	1.00 (0.95, 1.05)	0.97 (0.92, 1.02)	0.96 (0.92, 1.01)
Model 7	1	0.99 (0.94, 1.04)	1.01 (0.97, 1.07)	0.98 (0.93, 1.02)	0.96 (0.92, 1.01)

Model 1: adjusted for age (continuously).
Model 2: model 1 + ethnicity (caucasian, black, or other), smoking (never, past, current: 1-14, 15-24, or ≥25 cigarettes/day), marital status (currently married, widowed, divorced or separated, or never married), family history of type 2 diabetes (yes/no), hypertension(yes/no), hypercholesterolemia (yes/no), postmenopausal status and hormone use(premenopausal, or if postmenopausal, never, current, or past postmenopausal hormone use), total energy intake (quintiles), level of physical activity (<3, 3-9, 9-18, 18-27, or ≥27 MET-h/week), alcohol consumption (0, 0.1-5, 5-10, 10-15, or ≥15 g/day), and BMI (<21, 21-23, 23-25, 25-27, 27-30, 30-33, 33-35, 35-40, or ≥40 kg/m2).
Model 3: model 2 + intake of cereal fiber (quintiles), animal protein (quintiles), and PUFA:SFA ratio (quintiles).
Model 4: model 2 + multivitamin use (yes/no).
Model 5: model 3 + multivitamin use (yes/no).
Model 6: model 2 + AHEI index: (quintiles).
Model 7: model 3 + AHEI index: (quintiles).
** p-Value for Q-statistic <0.05, indicating statistically significant heterogeneity among the three cohorts*

Table 5.2.6. Hazard ratios (95% CIs) for type 2 diabetes according to quintiles of methionine

	Q1	Q2	Q3	Q4	Q5
NHS					
Methionine (mcg/d)					
Median	1.4	1.6	1.7	1.8	2
Cases/person-years	1,264/347,134	1,426/354,669	1,590/358,789	1,766/351,334	2,095/351,502
Model 1	1	**1.11 (1.03, 1.20)**	**1.23 (1.14, 1.32)**	**1.38 (1.28, 1.48)**	**1.63 (1.51, 1.74)**
Model 2	1	1.02 (0.94, 1.10)	1.03 (0.95, 1.11)	1.04 (0.95, 1.12)	1.03 (0.96, 1.11)
Model 3	1	1.01 (0.90, 1.13)	0.97 (0.84, 1.11)	0.98 (0.83, 1.16)	0.95 (0.78, 1.15)
NHS II					
Methionine (mcg/d)					
Median	1.5	1.8	1.9	2.1	2.4
Cases/person-years	927/364,771	994/364,213	1,128/364,059	1,261/365,380	1,739/361,000
Model 1	1	1.08 (0.99, 1.18)	**1.23 (1.13, 1.34)**	**1.37 (1.26, 1.49)**	**1.86 (1.72, 2.01)**
Model 2	1	0.95 (0.87, 1.04)	0.98 (0.89, 1.07)	0.98 (0.90, 1.07)	1.06 (0.98, 1.15)
Model 3	1	0.92 (0.81, 1.04)	0.96 (0.82, 1.13)	0.95 (0.78, 1.14)	1.03 (0.82, 1.28)
HPFS					
Methionine (mcg/d)					
Median	1.8	2	2.2	2.4	2.6
Cases/person-years	578/173,892	581/175,037	664/174,833	698/173,990	918/173,704
Model 1	1	1.01 (0.90, 1.13)	**1.14 (1.02, 1.28)**	**1.20 (1.07, 1.34)**	**1.57 (1.41, 1.74)**
Model 2	1	0.92 (0.87, 1.09)	**0.98 (0.87, 1.09)**	0.97 (0.86, 1.08)	**1.10 (0.99, 1.23)**
Model 3	1	0.92 (0.79, 1.06)	0.94 (0.79, 1.12)	0.86 (0.71, 1.06)	0.92 (0.72, 1.16)
Pooled					
Methionine (mcg/d)					
Model 1	1	**1.08 (1.03, 1.14)**	**1.21 (1.15, 1.27)**	**1.34 (1.27, 1.40)**	**1.69 (1.61, 1.77)***
Model 2	1	0.98 (0.93, 1.03)	1.00 (0.95, 1.05)	1.00 (0.96, 1.06)	**1.06 (1.01, 1.11)**
Model 3	1	0.95 (0.89, 1.03)	0.96 (0.87, 1.05)	0.94 (0.84, 1.04)	0.97 (0.86, 1.09)

Model 1: adjusted for age (continuously).
Model 2: model 1 + ethnicity (caucasian, black, or other), smoking (never, past, current: 1-14, 15-24, or ≥25 cigarettes/day), marital status (currently married, widowed, divorced or separated, or never married), family history of type 2 diabetes (yes/no), hypertension(yes/no), hypercholesterolemia (yes/no), postmenopausal status and hormone use(premenopausal, or if postmenopausal, never, current, or past postmenopausal hormone use), total energy intake (quintiles), level of physical activity (<3, 3-9, 9-18, 18-27, or ≥27 MET-h/week), alcohol consumption (0, 0.1-5, 5-10, 10-15, or ≥15 g/day), and BMI (<21, 21-23, 23-25, 25-27, 27-30, 30-33, 33-35, 35-40, or ≥40 kg/m2).
Model 3: model 2 + intake of cereal fiber (quintiles), animal protein (quintiles), and PUFA:SFA ratio (quintiles).
** p-Value for Q-statistic <0.05, indicating statistically significant heterogeneity among the three cohorts*

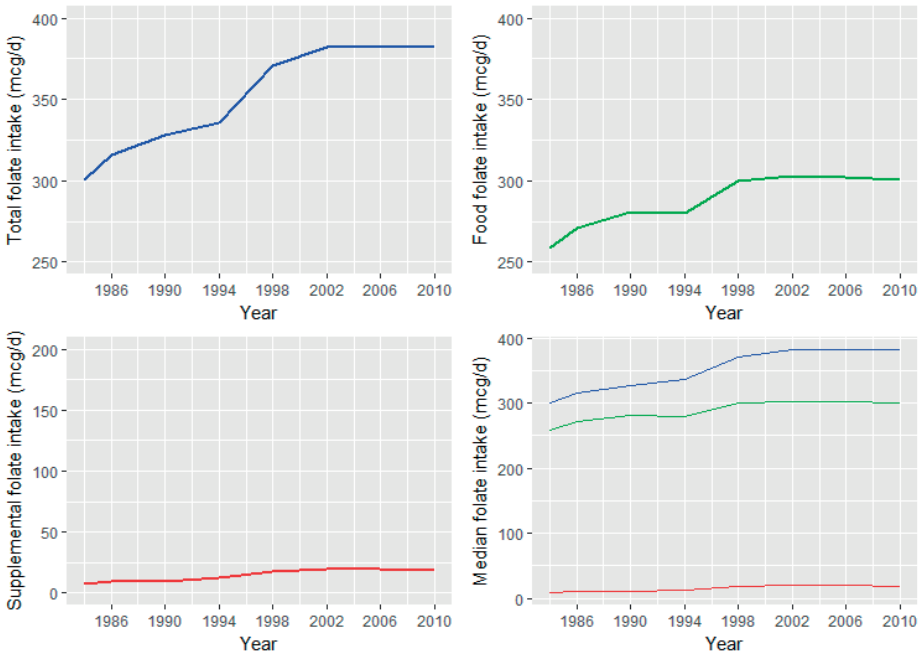


Figure 5.2.1. Median folate intake from 1984-2010

DISCUSSION

Our study suggests that higher intakes of vitamin B2 and vitamin B6 are associated with a lower risk of diabetes. A higher folate intake seemed to be associated with a lower risk of diabetes, but this was explained by intake of other nutrients. A higher vitamin B12 intake from supplements was associated with a lower risk of diabetes. In contrast, higher intake of vitamin B12 from food seems to be associated with a higher risk of diabetes, which may be due to animal products. There was no association between methionine intake and risk of diabetes.

We hypothesized that nutrients involved in one-carbon metabolism may influence the risk of type 2 diabetes through changes in DNA methylation. Of these nutrients, only vitamin B2 and vitamin B6 showed a protective association with the risk of type 2 diabetes. However, total intake of folate and methionine did not seem to be associated with risk of type 2 diabetes. The lack of findings on folate in contrast with findings from previous studies. In a case-control study, Nilsson et al. identified 251 CpG sites that were differentially methylated, of which the majority was hypomethylated, in diabetes cases compared to controls. The difference in DNA methylation could be explained by folate levels as these were positively correlated with the identified CpG sites, and folate levels were lower in participant with diabetes.[18] Results from this study suggest that the association between folate and diabetes may be mediated through changes in DNA methylation. Our null- findings for folate and vitamin B12 are in line with findings from a randomized controlled trial among women at high risk of CVD, where no significant effects of supplementation of vitamin B6, vitamin B12, and folate on the risk of type 2

diabetes were observed.[8] They did observe a reduction in diabetes risk among women with a family history of diabetes. Therefore, we tested in our study population whether associations differed for participant with family history of diabetes. However, no significant interaction was observed, suggesting that there was no difference for these groups in association between methyl donor nutrient intake and risk of type 2 diabetes. Possibly, the protective association observed for vitamin B2 and vitamin B6 on diabetes risk are mediated through one-carbon metabolism, and consequently differential DNA methylation. Another explanation would be that these vitamins lower the risk of diabetes through other mechanisms, as vitamin B2 and vitamin B6 have been proposed to improve glucose metabolism and inflammation markers.[19] So far, not many epidemiological studies have been conducted on the association between vitamin B2 and diabetes, but findings from clinical and animal studies suggest that vitamin B2 may decrease risk of diabetes through reduced inflammation, oxidative stress, and hyperglycemia.[20, 21] Results from animal studies suggest that supplementation of vitamin B2 may be beneficial for glucose and insulin metabolism, but also reduce diabetic complications.[21, 22] In a clinical study Bialy et al. showed that treatment of adipocyte-macrophage co-cultures with vitamin B2 lead to a reduction of pro-inflammatory factors and an increase of anti-inflammatory factors.[23] Our findings regarding vitamin B6 may be explained by its role in glucose metabolism. In an animal study, Maessen at al. reported that intervention with pyridoxine, a vitamin B6 analog, improved glucose metabolism and insulin resistance in obese mice.[24] In an intervention study among overweight and obese subjects, pyridoxine together with leucine had an improving effect on insulin sensitivity.[25] In addition, an increase in advanced glycation end products (AGEs) may lead to diabetes related complications. As vitamin B6 has anti-AGEs properties, it may be a therapeutic target for diabetes complications.[26]

Strengths of this study include the large sample size with a high follow-up rate and its prospective design, which allowed us to asses diet before disease onset. Furthermore, the repeated dietary data enabled us to calculate cumulative average to represent a long-term diet, which have less measurement error than estimated nutrient intake from single assessment.[27] We also had detailed information on nutrients from different sources and a wide range of information on lifestyle factors and other covariates, minimizing the chance of confounding. However, several limitations should be considered. Due to the observational nature of the study-design, residual confounding cannot be excluded. Since dietary intake is self-reported, it is prone to measurement errors. There was low variation in vitamin intake, limiting the ability to detect associations for extreme high or extreme low intakes. Our study population was socioeconomic homogeneous, which may enhance internal validity, however it also reduces external validity. Therefore, results need to be replicated in a more representative group of the total population.

Findings from this study suggest that higher intakes of vitamin B2 and vitamin B6 are associated with a lower risk of type 2 diabetes. Higher intake of vitamin B12 from supplements was also associated with a lower diabetes risk, whereas a higher vitamin B12 intake from food seems to be associated with a higher risk of diabetes, which may be due to animal products. Future studies are needed to replicate these findings in other populations.

REFERENCES

1. Hunter DJ, Reddy KS. Noncommunicable diseases. The New England journal of medicine. 2013; 369(14):1336-43.
2. Wang J, Wu Z, Li D, Li N, Dindot SV, Satterfield MC, et al. Nutrition, epigenetics, and metabolic syndrome. Antioxidants & redox signaling. 2012;17(2):282-301.
3. Barres R, Zierath JR. DNA methylation in metabolic disorders. The American journal of clinical nutrition. 2011;93(4):897S-900S.
4. Keating ST, El-Osta A. Epigenetic changes in diabetes. Clin Genet. 2013;84(1):1-10.
5. McKay JA, Mathers JC. Diet induced epigenetic changes and their implications for health. Acta Physiologica. 2011;202(2):103-18.
6. Anderson OS, Sant KE, Dolinoy DC. Nutrition and epigenetics: an interplay of dietary methyl donors, one-carbon metabolism and DNA methylation. The Journal of nutritional biochemistry. 2012;23(8): 853-9.
7. Krishnaveni GV, Hill JC, Veena SR, Bhat DS, Wills AK, Karat CL, et al. Low plasma vitamin B12 in pregnancy is associated with gestational 'diabesity' and later diabetes. Diabetologia. 2009;52(11):2350-8.
8. Song Y, Cook NR, Albert CM, Van Denburgh M, Manson JE. Effect of homocysteine-lowering treatment with folic Acid and B vitamins on risk of type 2 diabetes in women: a randomized, controlled trial. Diabetes. 2009;58(8):1921-8.
9. Willett WC, Sampson L, Browne ML, Stampfer MJ, Rosner B, Hennekens CH, et al. The use of a self-administered questionnaire to assess diet four years in the past. Am J Epidemiol. 1988;127(1):188-99.
10. Rimm EB, Giovannucci EL, Stampfer MJ, Colditz GA, Litin LB, Willett WC. Reproducibility and validity of an expanded self-administered semiquantitative food frequency questionnaire among male health professionals. Am J Epidemiol. 1992;135(10):1114-26; discussion 27-36.
11. Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. Am J Clin Nutr. 1997;65(4 Suppl):1220S-8S; discussion 9S-31S.
12. Hu FB, Leitzmann MF, Stampfer MJ, Colditz GA, Willett WC, Rimm EB. Physical activity and television watching in relation to risk for type 2 diabetes mellitus in men. Arch Intern Med. 2001;161(12):1542-8.
13. Manson JE, Rimm EB, Stampfer MJ, Colditz GA, Willett WC, Krolewski AS, et al. Physical activity and incidence of non-insulin-dependent diabetes mellitus in women. Lancet. 1991;338(8770):774-8.
14. Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. National Diabetes Data Group. Diabetes. 1979;28(12):1039-57.
15. Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Diabetes Care. 1997;20(7):1183-97.
16. American Diabetes A. Standards of medical care in diabetes--2010. Diabetes Care. 2010;33 Suppl 1: S11-61.
17. Food, Drug A. Food standards: amendment of standards of identity for enriched grain products to require addition of folic acid. Fed Regist. 1996;61:8781-97.
18. Nilsson E, Matte A, Perfilyev A, de Mello VD, Kakela P, Pihlajamaki J, et al. Epigenetic Alterations in Human Liver From Subjects With Type 2 Diabetes in Parallel With Reduced Folate Levels. J Clin Endocrinol Metab. 2015;100(11):E1491-501.
19. Valdes-Ramos R, Guadarrama-Lopez AL, Martinez-Carrillo BE, Benitez-Arciniega AD. Vitamins and type 2 diabetes mellitus. Endocr Metab Immune Disord Drug Targets. 2015;15(1):54-63.
20. Thakur K, Tomar SK, Singh AK, Mandal S, Arora S. Riboflavin and health: A review of recent human research. Crit Rev Food Sci Nutr. 2017;57(17):3650-60.

21. Alam MM, Iqbal S, Naseem I. Ameliorative effect of riboflavin on hyperglycemia, oxidative stress and DNA damage in type-2 diabetic mice: Mechanistic and therapeutic strategies. Arch Biochem Biophys. 2015;584:10-9.
22. Plows JF, Budin F, Andersson RA, Mills VJ, Mace K, Davidge ST, et al. The Effects of Myo-Inositol and B and D Vitamin Supplementation in the db/+ Mouse Model of Gestational Diabetes Mellitus. Nutrients. 2017;9(2).
23. Mazur-Bialy AI, Pochec E. Riboflavin Reduces Pro-Inflammatory Activation of Adipocyte-Macrophage Co-culture. Potential Application of Vitamin B2 Enrichment for Attenuation of Insulin Resistance and Metabolic Syndrome Development. Molecules. 2016;21(12).
24. Maessen DE, Brouwers O, Gaens KH, Wouters K, Cleutjens JP, Janssen BJ, et al. Delayed Intervention With Pyridoxamine Improves Metabolic Function and Prevents Adipose Tissue Inflammation and Insulin Resistance in High-Fat Diet-Induced Obese Mice. Diabetes. 2016;65(4):956-66.
25. Zemel MB, Bruckbauer A. Effects of a leucine and pyridoxine-containing nutraceutical on fat oxidation, and oxidative and inflammatory stress in overweight and obese subjects. Nutrients. 2012;4(6):529-41.
26. Dakshinamurti K. Vitamins and their derivatives in the prevention and treatment of metabolic syndrome diseases (diabetes). Can J Physiol Pharmacol. 2015;93(5):355-62.
27. Hu FB, Stampfer MJ, Rimm E, Ascherio A, Rosner BA, Spiegelman D, et al. Dietary fat and coronary heart disease: a comparison of approaches for adjusting for total energy intake and modeling repeated dietary measurements. Am J Epidemiol. 1999;149(6):531-40.

5.3

Associations between macronutrient intake and coronary heart disease (CHD): The Rotterdam Study

Based on:

Girschik C*, **Braun KVE***, Franco OH, Voortman T. Associations between macronutrient intake and incidence of coronary heart disease (CHD): The Rotterdam Study. *Manuscript in preparation*

**Denotes equal contribution*

ABSTRACT

Background: Substitution of macronutrients may affect the risk of coronary heart disease (CHD) in different manners. Therefore, the aim of this study is to investigate the relationship of macronutrient intake and CHD, by taking into account single macronutrient subgroups and different substitutions, in a large prospective cohort study in the Netherlands.

Methods: This study was performed in 5,905 participants from The Rotterdam Study, a population-based cohort study. Macronutrient intake was measured with use of a semi-quantitative food frequency questionnaire (FFQ). Incidence of CHD was ascertained through medical records. Cox proportional hazard regression analyses were used to study the association between intakes of macronutrients and CHD incidence.

Results: In multivariable-adjusted models plant protein was inversely associated with incident CHD at the expense of animal protein (HR of 0.54; 95%CI 0.30 - 0.97) as well as at the expense of saturated fat (HR of 0.56; 95%CI 0.33 - 0.98). Furthermore, animal protein was related to an increased risk of incident CHD at the expense of plant protein (HR of 1.64; 95%CI 1.01 - 2.68). However, total protein intake showed no association with CHD. Carbohydrates, fat and their subgroups were not associated with CHD incidence at the expense of any other macronutrient.

Conclusion: Findings from this population-based prospective cohort study suggest that a higher plant protein intake at the expense of either animal protein or saturated fat is associated with a lower risk of CHD. These findings support the current guidelines that recommend a dietary pattern providing more plant-based and less animal-based food.

INTRODUCTION

Cardiovascular diseases (CVD) remain the leading global cause of death and its prevalence is still rising.[1, 2] As coronary heart diseases (CHD) account for the major part of this amount, identifying determinants influencing CHD events is of importance.[3]

One of the major determinants for CHD is nutrition. The role of macronutrients in relation to CHD has been studied extensively, however, inconsistent findings have been reported. Although high intakes of saturated fatty acids (SFA) are generally considered a risk factor for CHD, a recent systematic reported that intake of SFAs is not associated with CHD.[4] However, replacement of SFAs with other macronutrients was not taken into account. Another systematic review showed that this association depended on the nutrients for which it was substituted. No association was observed between SFAs and CHD when replaced by carbohydrates or monounsaturated fatty acids (MUFA), but when replaced by polyunsaturated fatty acids (PUFA) a lower risk in CHD was observed.[5] These results suggest that the composition of the nutritional intake rather than single nutrients may be important in the interplay between diet and CHD risk. Besides SFA, other macronutrients and their subtypes have been studied in relation to CHD. In a prospective cohort study following 29,017 post-menopausal women from Iowa for 15 years an isoenergetic replacement of carbohydrates or animal protein by vegetable protein decreased CHD mortality by 30 % among women with the highest protein intake.[6] Within the Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study 21,955 male smokers between 50 and 69 years a replacement of SFA and transfat with carbohydrates showed a decreased risk for CHD, whereas the replacement of MUFA with carbohydrates resulted in an increased risk.[7]

Hence, the macronutrient composition of the diet rather than single macronutrients is of importance when studying the association between nutrition and CHD risk. Nevertheless, studies focusing on all macronutrients and its interplay have not been carried out so far. Therefore, the aim of this study is to investigate the relationship of macronutrient intake and CHD, by taking into account single macronutrient subgroups and substitution in this association, in a large prospective cohort study in the Netherlands.

METHODS

The Rotterdam Study

The current analysis was embedded in the Rotterdam study, a population-based prospective cohort study in the city of Rotterdam, the Netherlands. As a response to the demographic changes of the population, this follow-up study aimed to evaluate the occurrence and progression of chronic diseases and their risk factors that are frequent in elderly persons. The first cohort started in 1990 (RS-I) and included participants of 55 years or over, living in the well-defined Ommoord district of Rotterdam. Until September 1993 all residents of this suburb matching the criteria (n=10,215) were invited and 7,983 people responded (response of 78%). In 2000, a second cohort (RS-II) started with persons that became 55 years or those of 55 years or over that migrated into this district since the start of the study. Of the initially 4,472 invited persons, 3,011 responded (response of 67%). A further extension of the study (RS-III) was initiated in 2006, of the 6,057 invited subjects aged 45-54 years, 3,932 responded

(65% response). In total 14,926 subjects are included in the Rotterdam study, with an overall response of 72%. Briefly, examinations of the participants were administered at baseline enrollment including a home interview (2h) by a trained research assistant and physically examinations (5h) as well as dietary assessment at the research center in their district. All participants gave written informed consent and the medical ethics committee of the Erasmus Medical Center and The Netherlands Ministry of Health, Welfare and Sports approved the study protocol.[8]

Study population

To evaluate the association of the intake of macronutrients and coronary heart disease, data from subjects who responded to the study center at baseline from RS-I (n=7,983), RS-II (n=3,011) and RS-III (n=3,932) of the Rotterdam study were used. Of all 14,926 participants we excluded 5,225 because of incomplete dietary data. In these cases diet was not assessed mainly because of exclusion of subjects suspected of dementia, logistic reasons, implausible reported dietary intake, not returned FFQ, or because dietitians considered reports unreliable. Furthermore, we excluded 86 subjects because of missing follow-up data, or no informed consent on CHD, and 3,710 because of a history or missing data of CHD, stroke, or diabetes mellitus at baseline, which left 5,905 participants for current analysis (**Figure 5.3.1**).

Dietary assessment

Dietary intake of macronutrients for RS-I and RS-II was assessed in a two-step approach. At baseline, participants completed a self-administered questionnaire about the type of food and drinks consumed at least twice a month during the preceding year, as well as about dietary habits, supplement use and prescribed diets. Based on these data, subjects underwent a standardized interview supervised by a trained dietitian during their next visit in the research center after completing a checklist about their nutritional intake using 170-item semi-quantitative food frequency questionnaires (FFQ) to assess the amount and frequency in times per day, week, or month of their nutritional intake. The FFQ included 13 food groups and additional questions on prescribed diets, supplementary use, dietary habits and took seasonal variations in fruit, vegetable, and fish intakes into account.[9] In RS-III, subjects completed a self-administered but more comprehensive FFQ containing of 389 items, based on a validated FFQ for Dutch adults that also took into account preparation methods.[10, 11] Portion sizes in both FFQs were expressed in standardized household measures (e.g., spoon or cup) and natural units (e.g., piece of fruit or slice of bread). To derive total energy and nutrient intakes per day the received dietary data were linked to the Dutch Food Composition Table of 1993 for RS-I and RS-II,[12] and to the Dutch Food Composition Table of 2006 for RS-III.[13]

Ascertainment of CHD

The primary endpoint of this study was incident CHD after the standardized home interview of each individual participant at baseline. CHD was defined as a fatal or non-fatal myocardial infarction (MI), a surgical or percutaneous myocardial revascularization procedure, or fatal CHD.[14] Information on incident CHD was obtained through automated linkage of the study database to medical reports of general practitioners and hospitals in the research area. Data from participants with general practitioners outside the study area was obtained through annual collection of their medical records. All

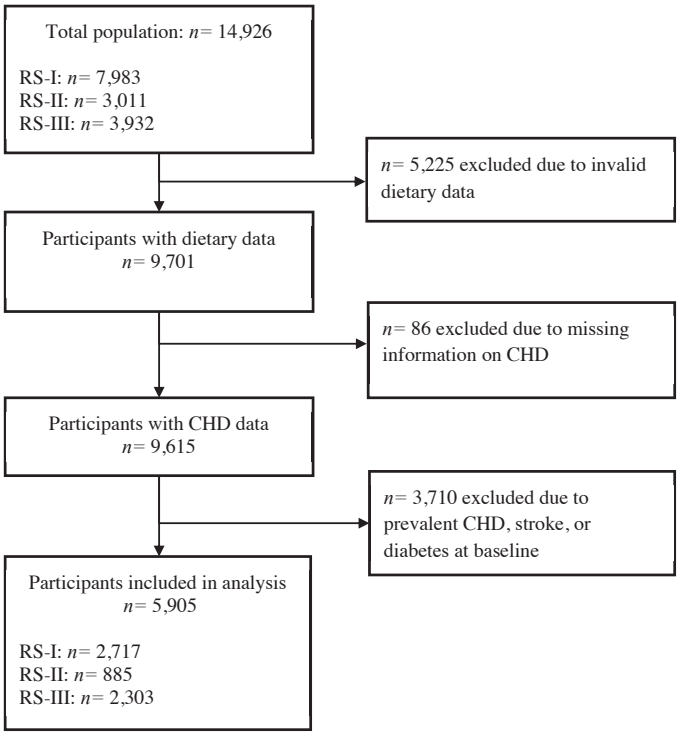


Figure 5.3.1. Flow chart of participants included for analysis

data on potential events were independently adjudicated by two research physicians and subsequently reviewed by a medical specialist in cardiovascular disease, whose judgment was considered decisive. Follow-up information for CHD was available until January 1, 2012.

Assessment of covariates

Information on sociodemographic data, current health status, medical history, medication use, and health behavior was obtained during a home visit by a trained research assistant using a structured questionnaire. Educational level of the subjects was categorized as low (primary education), intermediate (secondary general or vocational education), or high (higher vocational education or university). Smoking behavior was classified in never, former, or current. Data on physical activity were obtained through combinations of an adapted version of the Zutphen Physical Activity Questionnaire and the LASA Physical Activity Questionnaire.[15, 16] Because of differently extended data generation for physical activity between the three cohorts, z-scores were calculated to obtain comparable data. Height and weight of the participants were obtained during physical examinations at the research center while wearing indoor clothes without shoes. Body mass index (BMI) was calculated as weight divided by height squared (kg/m^2). Blood pressure was measured in duplicate at the right brachial artery using a random-zero sphygmomanometer with the participant sitting quietly for at least five minutes. The mean of both measurements was calculated. Participants were categorized as hypertensive when they

had a systolic blood pressure ≥140 mmHg, and/or a diastolic blood pressure ≥90 mmHg, and/or use of antihypertensive medication. Information on prevalent diseases at baseline was obtained during the standardized home interview and reviewed by a study physician using medical records, during physical examinations at the research center and through linkage with data from the Nationwide Medical Registry (LMR, Utrecht, the Netherlands).[14] Current medication use was assessed during the baseline home interview and was coded in accordance with the Anatomical Therapeutic Chemical (ATC) classification system.[14] Subjects were defined as prevalent type 2 diabetes cases if they had a serum glucose concentration ≥ 11.1 mmol/L before or 2 h after a 75-g glucose load or used antidiabetic medication. Stroke was defined as a combination of symptoms showing rapidly developing clinical signs of cerebral impairment for at least 24 h or leading to death caused by vascular origin [17]. Serum total cholesterol was assessed in non-fasting blood samples with an automated enzymatic method using Roche CHODPAP reagent agent.

Statistical analysis

To assess the associations between macronutrients and CHD incidence, a c**ox** (proportional hazards) regression was used. Therefore, person years were calculated starting from the date of baseline examinations until January 1, 2012, or first CHD event, last contact or death. Taking into account the first occurring event, hazard ratios (HRs) were calculated with 95% confidence intervals (CI) for CHD incidence. For evaluation of the association between intakes of macronutrients and CHD incidence, total amounts of the macronutrients carbohydrates, fat and protein as well as their subgroups, i.e. mono- and disaccharides, polysaccharides, animal and plant protein, MUFA, PUFA, and SFA were included in the analysis. In order to account for confounding effects of total energy intake and substitutional influences of other macronutrients nutrient density substitution models were created.[18] In these models intakes of all nutrients that contribute to the total amount of energy intake were included, except for one. Intakes of g per day were calculated for 5% of total energy intake and included in the same model, at the expense of one nutrient (i.e. the nutrient left out of the model). Therefore, coefficients can be interpreted as a 5E% higher intake at the expense of the nutrient left out. Subsequently, three models were constructed including selections of confounders based on literature. The first crude model was adjusted for the energy adjusted macronutrients except one, total energy intake, age, sex, and study cohort. The second model was additionally adjusted for education, smoking, physical activity. And the third fully adjusted model included additionally BMI, hypertension and antihypertensive use, lipid reducing medication, serum cholesterol and postmenopausal status. In order to reduce bias due to missing data and increase statistical power multiple imputation (n=5 imputations) was implemented using the predictive mean matching method. All analyses were executed with IBM SPSS Statistics for Windows, version 21.0, software (IBM Corporation, Armonk, New York).

RESULTS

Baseline characteristics of the study population are shown in **Table 5.3.1**. The total population consists of 5,905 with the highest number of participants in RS-I (n=2,717) followed by RS-III (n=2,303)

Table 5.3.1. Baseline characteristics of 5905 Dutch adults in three different cohorts of the Rotterdam study

		Rotterdam Study cohort (n=5905)					
		RS-I (n=2717)		RS-II (n=885)		RS-III (n=2303)	
		Mean (SD) / Median [95%CI]	Count (%)	Mean (SD) / Median [95%CI]	Count (%)	Mean (SD) / Median [95%CI]	Count (%)
Age at entering the Rotterdam Study		65.32 (6.74)		62.82 (7.64)		56.65 (6.28)	
Sex	male		1030 (37.9)		366 (41.4)		898 (39.0)
	female		1687 (62.1)		519 (58.6)		1405 (61.0)
Person years in the Rotterdam Study		19.26 (4.93)		13.81 (3.17)		7.46 (1.31)	
Incident CHD			410 (15.0)		64 (7.2)		35 (1.5)
Person years of CHD		15 (4)		9 (2)		3 (1)	
BMI		25.98 (3.40)		27.15 (4.08)		27.20 (4.31)	
Highest level of education		1 (1)		2 (1)		2 (1)	
Smoking status	never smoker		957 (35.2)		261 (29.5)		753 (32.7)
	ever smoker		1168 (43.0)		403 (45.5)		997 (43.3)
	current smoker		592 (21.8)		221 (25.0)		553 (24.0)
Physical activity in MET/h		81.24 [18.95 – 199.93]		78.14 [16.88 – 173.57]		43.05 [2.57 – 200.28]	
Energy intake in kcal/d		1983.35 (503.39)		2149.65 (592.90)		2308.52 (713.65)	
Total protein		81.76 (19.47)		88.63 (26.37)		88.39 (26.91)	
Plant protein		28.66 (8.46)		30.98 (11.99)		37.25 (13.67)	
Animal protein		53.17 (15.55)		57.65 (22.37)		51.27 (19.18)	
Total Carbohydrates		214.43 (61.28)		226.26 (88.46)		259.76 (88.81)	
Mono- and disaccharides		107.78 (40.75)		109.75 (52.36)		126.09 (55.69)	
Polysaccharides		106.66 (31.99)		115.91 (48.77)		133.69 (49.20)	
Total fat		80.35 (27.48)		87.81 (25.61)		83.63 (35.24)	
Saturated fat		31.69 (11.76)		34.30 (11.12)		29.52 (13.17)	

Table 5.3.1. Baseline characteristics of 5905 Dutch adults in three different cohorts of the Rotterdam study (continued)

	Rotterdam Study cohort (n=5905)					
	RS-I (n=2717)		RS-II (n=885)		RS-III (n=2303)	
	Mean (SD) / Median [95%CI]	Count (%)	Mean (SD) / Median [95%CI]	Count (%)	Mean (SD) / Median [95%CI]	Count (%)
Transfat	2.31 (1.34)		1.71 (0.84)		1.34 (0.68)	
Mono-unsaturated fat	27.45 (10.25)		26.62 (9.87)		29.00 (13.17)	
Poly-unsaturated fat	15.26 (7.65)		18.97 (5.16)		17.50 (8.58)	
Alcohol intake in grams/d	4.27 [0.00 – 49.73]		7.22 [0.00 – 63.06]		8.08 [0.00 – 54.37]	
Hypertension or antihypertensive use		1245 (45.8)		216 (24.4)		1038 (45.1)
Serum lipid reducing agents		54 (2.0)		83 (9.4)		393 (17.1)
Serum cholesterol	10.79 (3.04)		5.76 (0.98)		5.65 (1.03)	
Postmenopausal		2635 (97.0)		816 (92.2)		1673 (72.6)

and the fewest in RS-II (n=885). Mean ± SD age of subjects in RS-I was 65.3 ± 6.7 years, subjects in RS-II had a mean age of 62.8 ± 7.6 years, and in RS-III 56.7 ± 6.3 years. The person years of follow up since baseline differed between the cohorts depending on the start of the study with a mean of 19.3 ± 4.9 person years in RS-I, 13.8 ± 3.2 years in RS-II and 7.5 ± 1.3 years in RS-III. During this time 410 (15.0 %) CHD events in RS-I, 64 (7.2 %) in RS-II and 35 (1.5 %) in RS-III occurred.

Table 5.3.2 presents the associations between total, plant and animal protein intake at the expense of the other macronutrients with the incidence of CHD. In fully adjusted models (model 3), intake of plant protein was negatively associated with incident CHD when consumed at the expense of animal protein (HR of 0.54; 95%CI 0.30 - 0.97) or SFA (HR of 0.56; 95%CI 0.33 - 0.98), but not at the expense of the other macronutrients. Animal protein intake was positively associated with incident CHD when consumed at the expense of plant protein (HR of 1.64; 95%CI 1.01 - 2.68), but not at the expense of other macronutrients. Total protein intake was not related to CHD incidence when consumed at the expense of fat or carbohydrates and its subgroups.

Table 5.3.2. Association between incident CHD and total, plant and animal protein intake at the expense of the other macronutrients and their subgroups

	Total Protein HR[a] (95% CI[b])	**Plant Protein** HR[a] (95% CI[b])	**Animal Protein** HR[a] (95% CI[b])
Expense of			
-Plant Protein	/	/	1.643 (1.007 - 2.682)*
-Animal Protein	/	0.541 (0.302 - 0.967)*	/
Total Fat	1.080 (0.903 - 1.292)	0.587 (0.341 - 1.012)	1.131 (0.943 - 1.358)
-Saturated Fat	0.940 (0.727 - 1.216)	0.563 (0.325 - 0.976)*	1.045 (0.792 - 1.380)
-Mono-Unsaturated Fat	1.293 (0.974 - 1.715)	0.697 (0.362 - 1.344)	1.279 (0.962 - 1.700)
-Poly-Unsaturated Fat	1.041 (0.827 - 1.310)	0.572 (0.312 - 1.052)	1.058 (0.840 - 1.332)
Total Carbohydrates	1.092 (0.917 - 1.299)	0.717 (0.443 - 1.163)	1.098 (0.921 - 1.309)
-Mono- and Disaccharides	1.089 (0.914 - 1.299)	0.610 (0.342 - 1.088)	1.123 (0.938 - 1.346)
-Polysaccharides	1.105 (0.909 - 1.342)	0.564 (0.288 - 1.105)	1.038 (0.846 - 1.275)

[a]Hazard ratio (HR) estimated from Cox proportional hazard regression model. [b]Confidence interval of the estimated HR. Multivariate model were adjusted for macronutrients, energy intake (continuous, in kcal/day), age (continuous), sex (reference = male), cohort (RS-I, RS-II, RS-III), education (low, intermediate, high), smoking (current, former, never), physical activity (continuous), BMI (continuous), hypertension or antihypertensive use, lipid reducing medication, serum cholesterol (continuous), postmenopausal status. *p<0.05.

Associations between total carbohydrate, mono- and disaccharide and polysaccharide intake and CHD incidence of the fully adjusted models are shown in **Table 5.3.3**. No association was found for mono- and disaccharides at the expense of polysaccharides (HR of 0.92; 95%CI 0.80 - 1.07). Intake of total carbohydrates, mono- and disaccharides and polysaccharides were not related to CHD incidence when consumed at the expense of protein or fat and its subgroups.

Table 5.3.3. Association between incident CHD and total carbohydrates, mono- and disaccharides and polysaccharides at the expense of the other macronutrients and their subgroups

	Total Carbohydrates HR[a] (95% CI[b])	**Mono- and Disaccharides** HR[a] (95% CI[b])	**Polysaccharides** HR[a] (95% CI[b])
Expense of			
-Mono- and Disaccharides	/	/	1.084 (0.934 - 1.258)
-Polysaccharides	/	0.924 (0.796 - 1.072)	/
Total Protein	0.928 (0.780 - 1.104)	0.925 (0.773 - 1.107)	0.912 (0.748 - 1.112)
-Plant Protein	1.367 (0.917 - 2.038)	1.463 (0.898 - 2.384)	1.560 (0.885 - 2.749)
-Animal Protein	0.916 (0.771 - 1.089)	0.893 (0.745 - 1.070)	0.968 (0.789 - 1.188)
Total Fat	0.990 (0.916 - 1.069)	1.001 (0.919 - 1.089)	1.103 (0.951 - 1.279)
-Saturated Fat	0.859 (0.694 - 1.063)	0.930 (0.736 - 1.175)	1.009 (0.771 - 1.320)
-Mono-Unsaturated Fat	1.186 (0.954 - 1.475)	1.141 (0.911 - 1.429)	1.236 (0.968 - 1.577)
-Poly-Unsaturated Fat	0.953 (0.794 - 1.143)	0.941 (0.784 - 1.129)	1.021 (0.816 - 1.276)

[a]Hazard ratio (HR) estimated from Cox proportional hazard regression model. [b]Confidence interval of the estimated HR. Multivariate model were adjusted for macronutrients, energy intake (continuous, in kcal/day), age (continuous), sex (reference = male), cohort (RS-I, RS-II, RS-III), education (low, intermediate, high), smoking (current, former, never), physical activity (continuous), BMI (continuous), hypertension or antihypertensive use, lipid reducing medication, serum cholesterol (continuous), postmenopausal status. *p<0.05.

Table 5.3.4 presents the associations between total, MUFA, PUFA, and SFA intake with incident CHD. MUFA was not related to CHD at the expense of SFA (HR of 0.81; 95%CI 0.53 - 1.24) or PUFA (HR of 0.82; 95%CI 0.59 - 1.14). PUFA was not related to CHD at the expense of SFA (HR of 0.99; 95%CI 0.77 - 1.28) or MUFA(HR of 1.21; 95%CI 0.88 - 1.65), as well as SFA showed no association at the expense of MUFA (HR of 1.22; 95%CI 0.82 - 1.82) or PUFA (HR of 1.02; 95%CI 0.79 - 1.32). Furthermore, total, MUFA, PUFA and SFA intake were not related to CHD incidence when consumed at the expense of any of the other examined macronutrients.

Table 5.3.4. Association between incident CHD and total, mono-unsaturated, poly-unsaturated and saturated fat at the expense of the other macronutrients and their subgroups

	Total Fat HR[a] (95% CI[b])	**Mono-Unsaturated Fat** HR[a] (95% CI[b])	**Poly-Unsaturated Fat** HR[a] (95% CI[b])	**Saturated Fat** HR[a] (95% CI[b])
Expense of				
-Saturated Fat	/	0.810 (0.530 - 1.237)	0.990 (0.766 - 1.280)	/
-Mono-Unsaturated Fat	/	/	1.206 (0.882 - 1.648)	1.220 (0.820 - 1.815)
-Poly-Unsaturated Fat	/	0.818 (0.585 - 1.143)	/	1.016 (0.785 - 1.315)
Total Carbohydrates	1.013 (0.939 - 1.093)	0.848 (0.672 - 1.070)	1.054 (0.882 - 1.258)	1.072 (0.852 - 1.348)
-Mono- and Disaccharides	1.010 (0.929 - 1.097)	0.870 (0.686 - 1.102)	1.063 (0.889 - 1.271)	1.078 (0.856 - 1.358)
-Polysaccharides	1.023 (0.917 - 1.142)	0.801 (0.620 - 1.036)	0.985 (0.791 - 1.226)	0.999 (0.766 - 1.302)
Total Protein	0.940 (0.788 - 1.122)	0.770 (0.570 - 1.039)	0.975 (0.778 - 1.221)	1.078 (0.836 - 1.389)
-Plant Protein	1.343 (0.922 - 1.956)	1.276 (0.713 - 2.285)	1.536 (0.928 - 2.541)	1.577 (0.998 - 2.493)
-Animal Protein	0.904 (0.755 - 1.082)	0.774 (0.574 - 1.045)	0.952 (0.759 - 1.194)	0.966 (0.734 - 1.271)

[a]Hazard ratio (HR) estimated from Cox proportional hazard regression model. [b]Confidence interval of the estimated HR. Multivariate model were adjusted for macronutrients, energy intake (continuous, in kcal/day), age (continuous), sex (reference = male), cohort (RS-I, RS-II, RS-III), education (low, intermediate, high), smoking (current, former, never), physical activity (continuous), BMI (continuous), hypertension or antihypertensive use, lipid reducing medication, serum cholesterol (continuous), postmenopausal status. *p<0.05.

DISCUSSION

In this population-based study among 5,905 Dutch persons aged 45 or older, plant protein was inversely associated with incident CHD at the expense of animal protein as well as at the expense of SFA. Furthermore, animal protein was related to an increased risk of incident CHD at the expense of plant protein. However, total protein intake showed no association with CHD. Carbohydrates, fat and their subgroups were not associated with CHD incidence at the expense of any other macronutrient.

Findings of the present study show that higher plant protein intake is related to a lower CHD risk. This beneficial association might be driven by reducing CHD risk factors. Postulated effects of an increased dietary protein intake are a weight loss due to a greater satiety and lower subsequent energy intake, improved lipid and lipoprotein profile by decreasing triacylglycerol and oxidized LDL cholesterol and reduced blood pressure.[19-21] However, these findings remain controversial and seem to depend on the dietary source of protein as a critical determinant of the outcome. A systematic review focusing on the protein source has shown beneficial effects on blood pressure for plant protein, and no association for animal protein.[22] Similarly, a review of observational and interventional studies regarding CVD risk underlines the favorable effects of a diet containing mainly plant-derived protein sources as well as unprocessed animal-based protein foods that are low in SFA,[23] which is in line with the present study. However, no significant association was observed between total protein intake and CHD risk. Due to the opposite directions in which plant and animal protein are associated with CHD risk, it is possible that both protein subgroups outweigh each other, resulting in a null-association for total protein intake. In line with several previous studies, we observed that a higher plant protein intake at the expense of SFA was associated with lower CHD incidence.[6, 24] Beneficial effects by replacing SFA in the diet have also been shown in former studies. In contrast to the present study, this association was evaluated in replacing SFA by PUFAs as reported in a pooled analysis of 11 European and American cohort studies, which analyzed the association of MUFAs, PUFAs and carbohydrates with CHD. However, carbohydrate subgroups and dietary protein intake were not included in the analysis.[5] Another study combining data from the Nurses´ Health Study following 84,628 women and Health Professionals Follow-up Study following 42,908 men took different sources of carbohydrates into account and found that both PUFAs and high-quality carbohydrates are associated with a decreased risk of CHD if they replace dietary SFA.[25]

Strengths of this study contain the prospective study design with extensive data collection, allowing the adjustment for many possible confounders and a long follow-up period. Furthermore, in this study we used a comprehensive approach that combines all macronutrients, including subgroups and evaluated all possible combinations of substitution by each other in association with CHD incidence. Due to application of the nutrient density substitution model, the amount of energy intake as well as influencing effects of other nutrients were taken into account and allow a more practical interpretation of the results.[18] A limitation of this study is the observational nature, therefore, no causal relationship can be concluded. Furthermore, residual confounding cannot be ruled out, as in any observational study. Moreover, using an FFQ relies on self-report an can lead to measurement error. Therefore, it may not have been possible to estimate accurate absolute amounts of nutritional intakes Nevertheless, this instrument has been shown to be an appropriate method for ranking individuals by their nutritional intakes in epidemiologic studies and has been validated against 15 24-h food records

among participants in the Rotterdam Study obtaining good correlations between both methods.[9, 26] A further limitation of this study consists in the lack of repeated data on dietary intakes. Changes in the diet between baseline measurement and the outcome could not be assessed and included in the analysis. Furthermore, the generalizability of these results should be taken into account. As subjects with prevalent CHD, stroke and diabetes at baseline were excluded from the analysis suggesting a higher risk of developing CHD in these groups, the evaluated study population is a relatively healthy group. Therefore, results can only be generalized to a relatively healthy older population.

In conclusion, findings from this population based prospective cohort study suggest that a higher plant protein intake at the expense of either animal protein or SFA is associated with a lower risk of CHD. This is in line with current dietary guidelines that recommend a shift towards a higher plant-based diet and lower animal-based diet. The current study has shown substitutional associations of macronutrients and contributes to an evaluation of the whole diet in context of CHD risk. Nevertheless, further studies that investigate the underlying mechanisms are needed.

REFERENCES

1. Fuster V, Mearns BM. The CVD paradox: mortality vs prevalence. Nat Rev Cardiol. 2009;6(11):669.
2. Laslett LJ, Alagona P, Jr., Clark BA, 3rd, Drozda JP, Jr., Saldivar F, Wilson SR, et al. The worldwide environment of cardiovascular disease: prevalence, diagnosis, therapy, and policy issues: a report from the American College of Cardiology. J Am Coll Cardiol. 2012;60(25 Suppl):S1-49.
3. Mackay J, Mensah GA. The atlas of heart disease and stroke: World Health Organization; 2004.
4. Chowdhury R, Warnakula S, Kunutsor S, Crowe F, Ward HA, Johnson L, et al. Association of dietary, circulating, and supplement fatty acids with coronary risk: a systematic review and meta-analysis. Ann Intern Med. 2014;160(6):398-406.
5. Jakobsen MU, O'Reilly EJ, Heitmann BL, Pereira MA, Balter K, Fraser GE, et al. Major types of dietary fat and risk of coronary heart disease: a pooled analysis of 11 cohort studies. Am J Clin Nutr. 2009;89(5): 1425-32.
6. Kelemen LE, Kushi LH, Jacobs DR, Jr., Cerhan JR. Associations of dietary protein with disease and mortality in a prospective study of postmenopausal women. Am J Epidemiol. 2005;161(3):239-49.
7. Simila ME, Kontto JP, Mannisto S, Valsta LM, Virtamo J. Glycaemic index, carbohydrate substitution for fat and risk of CHD in men. Br J Nutr. 2013;110(9):1704-11.
8. Hofman A, Brusselle GG, Darwish Murad S, van Duijn CM, Franco OH, Goedegebure A, et al. The Rotterdam Study: 2016 objectives and design update. Eur J Epidemiol. 2015;30(8):661-708.
9. Klipstein-Grobusch K, den Breeijen JH, Goldbohm RA, Geleijnse JM, Hofman A, Grobbee DE, et al. Dietary assessment in the elderly: validation of a semiquantitative food frequency questionnaire. Eur J Clin Nutr. 1998;52(8):588-96.
10. Feunekes GI, Van Staveren WA, De Vries JH, Burema J, Hautvast JG. Relative and biomarker-based validity of a food-frequency questionnaire estimating intake of fats and cholesterol. Am J Clin Nutr. 1993;58(4):489-96.
11. Goldbohm RA, van den Brandt PA, Brants HA, van't Veer P, Al M, Sturmans F, et al. Validation of a dietary questionnaire used in a large-scale prospective cohort study on diet and cancer. Eur J Clin Nutr. 1994;48(4):253-65.
12. Table DFC. Nevo tabel 1993 Voorlichtingsbureau voor de Voeding The Hague. The Netherlands. 1993.
13. Nevo-tabel N. Nederlands Voedingsstoffenbestand. Dutch Nutrient database. 2006.
14. Leening MJ, Kavousi M, Heeringa J, van Rooij FJ, Verkroost-van Heemst J, Deckers JW, et al. Methods of data collection and definitions of cardiac outcomes in the Rotterdam Study. Eur J Epidemiol. 2012; 27(3):173-85.
15. Caspersen CJ, Bloemberg BP, Saris WH, Merritt RK, Kromhout D. The prevalence of selected physical activities and their relation with coronary heart disease risk factors in elderly men: the Zutphen Study, 1985. Am J Epidemiol. 1991;133(11):1078-92.
16. Stel VS, Smit JH, Pluijm SM, Visser M, Deeg DJ, Lips P. Comparison of the LASA Physical Activity Questionnaire with a 7-day diary and pedometer. J Clin Epidemiol. 2004;57(3):252-8.
17. Wieberdink RG, Ikram MA, Hofman A, Koudstaal PJ, Breteler MM. Trends in stroke incidence rates and stroke risk factors in Rotterdam, the Netherlands from 1990 to 2008. Eur J Epidemiol. 2012;27(4): 287-95.
18. Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. Am J Clin Nutr. 1997;65(4 Suppl):1220S-8S; discussion 9S-31S.
19. Hu FB. Protein, body weight, and cardiovascular health. Am J Clin Nutr. 2005;82(1 Suppl):242S-7S.
20. Leidy HJ, Clifton PM, Astrup A, Wycherley TP, Westerterp-Plantenga MS, Luscombe-Marsh ND, et al. The role of protein in weight loss and maintenance. Am J Clin Nutr. 2015.

21. Vasdev S, Stuckless J. Antihypertensive effects of dietary protein and its mechanism. Int J Angiol. 2010; 19(1):e7-e20.
22. Altorf-van der Kuil W, Engberink MF, Brink EJ, van Baak MA, Bakker SJ, Navis G, et al. Dietary protein and blood pressure: a systematic review. PLoS One. 2010;5(8):e12102.
23. Richter CK, Skulas-Ray AC, Champagne CM, Kris-Etherton PM. Plant protein and animal proteins: do they differentially affect cardiovascular disease risk? Adv Nutr. 2015;6(6):712-28.
24. Song M, Fung TT, Hu FB, Willett WC, Longo VD, Chan AT, et al. Association of Animal and Plant Protein Intake With All-Cause and Cause-Specific Mortality. JAMA Intern Med. 2016;176(10):1453-63.
25. Li Y, Hruby A, Bernstein AM, Ley SH, Wang DD, Chiuve SE, et al. Saturated Fats Compared With Unsaturated Fats and Sources of Carbohydrates in Relation to Risk of Coronary Heart Disease: A Prospective Cohort Study. J Am Coll Cardiol. 2015;66(14):1538-48.
26. Kipnis V, Subar AF, Midthune D, Freedman LS, Ballard-Barbash R, Troiano RP, et al. Structure of dietary measurement error: results of the OPEN biomarker study. Am J Epidemiol. 2003;158(1):14-21; discussion 2-6.

Chapter 6

General discussion & summary

6.1

General discussion

AIMS

The aim of this thesis was to investigate the interplay between nutritional factors, DNA methylation, and cardiometabolic risk factors and diseases in children as well as adults. We were mainly interested in nutrients that are known to be involved in DNA methylation through their role in the one-carbon metabolism, including vitamin B2, vitamin B6, vitamin B12, folate, and methionine. Other nutrients that may be involved in DNA methylation were examined as well, such as protein and fatty acids. The cardiometabolic risk factors of interest for both children and adults were body composition and lipid levels. Additionally, in adults we explored CHD events and type 2 diabetes.

MAIN FINDINGS

Chapter 2: Nutrition and DNA methylation

In the past years, there has been an increasing interest in the effect of nutrition on DNA methylation.[1-3] Altered DNA methylation may have an effect on several cardiometabolic outcomes; the role of nutrition in this mechanism may therefore be of importance in strategies for prevention and treatment of several chronic diseases.[4] Results from animal studies have shown promising results,[5-7] and the past decade there has been an incredible increase in publications of human studies on this topic as well. However, findings from studies in humans on the role of nutrition in DNA methylation are still inconsistent.[1] In order to gain more insight in effects of several nutrients on DNA methylation across the life-course in humans, a clear overview of the current evidence is of importance. Identifying which nutrients affect DNA methylation of which CpG sites will provide insight in the mechanisms that are responsible for the effect of nutrition on several health outcomes. The aim of this chapter was to systematically review the literature on the associations between different nutrients with DNA methylation in humans across all stages of the life course (**Chapter 2.1**). The majority of the publications included in this review explored the role of methyl donor nutrients, including folate and other B-vitamins, and fatty acids in DNA methylation. The literature review showed that several nutrients, including folate, fatty acids, and vitamin D, are associated with DNA methylation on a global, gene-specific, or genome-wide level. However, the directions of these associations are inconsistent. Further large-scale studies of high quality are needed to elucidate the role of nutrition in DNA methylation.

Chapter 3: DNA methylation and cardiometabolic health

DNA methylation has been linked to several health outcomes, including cardiometabolic factors such as dyslipidemia and obesity. There is promising evidence that DNA methylation is related to cholesterol and triacylglycerol. In order to obtain a better understanding of the role of DNA methylation on dyslipidemia, a comprehensive overview of the current evidence is of importance. Therefore, we conducted a systematic review on studies investigating the association between DNA methylation (global, gene-specific, and genome-wide) and lipid levels (triacylglycerol, total cholesterol, LDL-cholesterol, and HDL-cholesterol) (**Chapter 3.1**). In this review we included 31 publications, based on 23 unique studies. There were no consistent associations observed for global DNA methylation and lipid levels. Gene-specific and genome-wide studies showed significant associations between lipid levels and DNA

methylation at *ABCG1*, *CPT1A*, *SREBF1*, *TNNT1*, *MIR33B*, and *TNIP*, which were replicated. Of all studies included in our review, only three performed an EWAS.[8-10] As a hypothesis-free approach was used in these types of studies, replication is important. Therefore, we performed an EWAS on lipid levels using data from the Rotterdam Study (**Chapter 3.2**). In this EWAS we replicated previously reported associations for methylation of *ABCG1*, *CPT1A*, and *SREBF1* with triacylglycerol and HDL-cholesterol.[9, 10] In addition, we observed a novel association between DNA methylation of a CpG site near *DHCR24* and HDL-cholesterol. The *DHCR24* gene encodes for 3-hydroxysterol-24 reductase, a cholesterol biosynthesis enzyme that catalyzes the conversion of desmosterol to cholesterol.[11, 12] Furthermore, we studied the association between DNA methylation and anthropological measures, such as BMI and waist circumference. We conducted an EWAS on BMI and WC in the Rotterdam Study and replicated these results in the ARIC study (**Chapter 3.3**). We identified and replicated 12 CpG sites associated with BMI and 13 CpG sites associated with WC. These included several CpG sites known to be related to cardiometabolic outcomes, such as *ABCG1* and *CPT1A*. In addition to these known CpG sites, we identified two novel CpG sites related to both BMI and WC, namely *MSI2* and *LARS2*. *MSI2* encodes RNA-binding proteins and plays a central role in posttranscriptional gene regulation,[13] and has been suggested to be associated with eating behaviors.[14] *LARS2* encodes an enzyme that catalyzes aminoacylation of mitochondrial tRNALeu and may be linked to bipolar disorder.[15, 16] Considering that bipolar disorder is associated with obesity, overweight, and abdominal obesity,[17] methylation of *MSI2* and *LARS2* could play a role in disturbances in eating behaviors, and consequently can affect BMI and WC. In conclusion, in our EWASs we replicated previously reported associations between DNA methylation and cardiometabolic outcomes, but also identified novel CpG sites related to lipids and obesity-related traits.

Chapter 4: Nutrition and cardiometabolic health in children

In chapter 4 we studied the associations between several nutrients, which may be involved in DNA methylation and cardiometabolic outcomes in children. Some of these nutrients are known to be directly involved in the one-carbon metabolism, such as vitamin B6, vitamin B12 and folate. Besides these nutrients, other nutrients, including protein and fatty acids, have been suggested to affect DNA methylation as well. These nutrients could therefore also be associated with cardiometabolic health in children through epigenetic mechanisms. In **Chapter 4.1** we started by investigating the association between the known methyl donor nutrients, including vitamin B6, vitamin B12, folate and methionine with body composition. We observed that high intake of folic acid at the age of 1 year was associated with a lower weight and BMI at the age of 6 years, whereas a high methionine intake was associated with an unfavorable body composition, including a higher BMI and body fat percentage. Furthermore, we studied the association between protein intake ate the age at 1 year with repeatedly measured growth and detailed body composition. We observed that a higher protein intake, in particular animal protein, at 1 year was associated with a higher height, weight, and BMI, up to the age of 9 years (**Chapter 4.2**), and we found that the positive association between protein intake and BMI was driven by a higher fat mass index (**Chapter 4.3**). Intake of fat or fatty acids at the age of 1 year, however, were not associated with these growth and body composition outcomes or with other aspects of cardiometabolic health in childhood (**Chapter 4.4).**

Chapter 5: Nutrition and cardiometabolic health in adults

The aim of this chapter was to investigate associations between nutrition and cardiometabolic health in different large cohort studies adults. First, we studied the association of folate and vitamin B12 levels with body composition in the B-PROOF study (**Chapter 5.1**). We observed that higher levels of folate were associated with a lower BMI. This association was most pronounced in participants with prevalent cardiometabolic diseases, and was mainly driven by FFMI rather than FMI among these participants. In contrast, higher levels of vitamin B12, measured by HoloTC and MMA, were associated with a higher FMI. Second, we studied the role of methyl donor nutrients and the risk of diabetes using data from the Nurses' Health Study and Health Professionals Follow-up Study (**Chapter 5.2**). We observed that higher intake of vitamin B2 and vitamin B6 was associated with a lower risk of type II diabetes. This association was attenuated slightly, but remained significant, after adjustment of several sociodemographic and lifestyle covariates. Higher intake of vitamin B12 from supplements was also associated with a lower diabetes risk, whereas a higher vitamin B12 intake from food seems to be associated with a higher risk of diabetes, which may be due to animal products. Finally, we investigated the association between macronutrient intake and the risk of CHD (**Chapter 5.3**). Although we observed no association between total protein intake and risk of CHD, when we separated analyses for animal and vegetable protein we observed significant associations in opposite directions. A higher animal protein intake was associated with a higher risk of CHD, whereas a higher plant protein intake was associated with a lower risk of CHD. The discrepancy in findings for animal and plant protein could be due to a difference in amino acid composition. Certain amino acids, such as methionine and lysine, may affect DNA methylation, and consequently cardiometabolic health.[3]

METHODOLOGICAL CONSIDERATIONS

Study design & study populations

For the studies in this thesis data from different cohorts were used. The majority of the studies included in this thesis were carried out in two Dutch population-based prospective cohort studies: The Generation R Study and the Rotterdam Study. The Generation R Study included offspring of women from Rotterdam who were recruited during pregnancy. Data on follow up in early childhood was available up to the age of 9 years. The Rotterdam Study included residents from Ommoord, aged 45 years and older. In addition, for a few studies in this thesis we used data from the Nurses' Health Study, Nurses' Health Study II, Health Professionals Follow-Up Study, ARIC, and B-PROOF Study. All the data we used from these studies were of observational nature. When interpreting the results of our studies, both internal validity and external validity should be considered. Internal validity refers to the accuracy of the measurements of the variables of interest in our studies. Regarding internal validity, three different types of bias should be taken into account, i.e. selection bias, information bias, and confounding. The specific types of information bias and confounding will be discussed per variable of interest separately later in this chapter. Selection bias occurs when associations between the exposure and outcome of interest are different for the population that participate in the study and the population that was eligible for inclusion, but did not participate. In this thesis, the studied populations tented toward a selection of a more healthy population with a higher social-economic status.

However, previous studies have shown that selective non-participation at baseline are not likely to be related to future risk of diseases and therefore do not strongly influence associations, making bias due to selection unlikely.[18] Nevertheless, it may affect the external validity of our results, which should be taken into account when extrapolating our findings. External validity, or generalizability, means to what extent findings can be extrapolated to the population of interest. In order to determine the generalizability of findings, one should always take into account the source population that is used to select subjects included in the studies. There are some differences in external validity between cohorts included in this thesis, as the Generation R Study and Rotterdam Study are both population-based cohort studies, whereas the Nurses' Health Study and Health Professionals Follow-Up Study exclusively consist of health professionals. These different source populations have several strengths and limitations. For example, the internal validity of studies including health professionals may be higher, as this is a well-educated group of health professionals and may be more capable of estimating their health behaviors, such as diet, as opposed to the general population. However, findings from these studies may not be generalizable to the total population. In contrast, the population-based studies may have a higher external validity, but internal validation may not be as high compared to the cohorts involving health professionals.

DNA methylation

For our studies involving DNA methylation, we used an epigenome-wide association study approach. This method provides the opportunity to identify novel CpG sites that are involved in regulation of a wide range of phenotypes, in our case, cardiometabolic health. Nevertheless, with this type of approach caution must be taken when interpreting the results. As EWAS follows a hypothesis-free approach, significant findings from these types of studies could be the result of false positives. This type of error, also known as Type I error, can be caused by multiple testing. Considering the high number of tests being carried out, the generally used cut-off of a p-value of <0.05 is not desirable as this would result in a high number of chance findings. One way of dealing with this issue is the use of the Bonferroni correction, i.e. dividing the cut-off for statistical significance (e.g. $P<0.05$) by number of test that are carried out. This limits the possibility of false-positive findings. However, this method is also considered to be too conservative as the number of effective tests may not be as high as the number of CpGs studied, due to the correlations among CpGs.[19] This correction may therefore be too stringent, which may lead to false-negative findings, also known as Type II error. Usually, this is caused by lack of power, mainly due to a small sample size. A common approach to overcome this power issue is the collaborations between studies. Furthermore, hypothesis-free approaches are in fact generating a hypothesis that need replication. Therefore, to ensure that our findings are true-positives, we made use of a discovery and a replication cohort. In the discovery cohort, we apply the EWAS approach, which fulfills hypothesis-generating purposes. Consequently, we investigate whether the significant results from the discovery cohort were also associated within the replication cohort, serving as a hypothesis-testing purpose. When findings replicate across two different cohorts, the chance that this is due to Type I error is minimal. Two different types of replication are usually applied, internal replication (within same study population) or external replication (across different study populations). When an internal replication is applied, populations of the discovery and replication cohorts are from the same source population and thus more homogeneous, making it less likely that associations will

be distorted by cohort-specific differences. When external replication is applied, heterogeneity across studies may limit the possibility to replicate findings across cohorts. On the other hand, external replication provides the opportunity to explore whether identified CpG sites replicate across different populations. In contrast to genetics, DNA methylation is dynamic over time and prone to change in response to environmental factors. Therefore, it is crucial to consider potential confounders when exploring the association between DNA methylation and cardiometabolic outcomes. Considering this, we adjusted our analyses for lifestyle factors, such as smoking, which has been shown to be strongly associated with DNA methylation.[20] Furthermore, the DNA methylation profile varies across different cell types. Therefore, confounding by differences in cell types should be taken into account when analyzing DNA methylation. Since whole blood was used for DNA methylation measurements, we adjusted all our models for leukocyte proportions,[21] as these are strongly correlated with DNA methylation. In addition, DNA methylation measurements are subject to batch effects. Therefore, we included technical covariates, i.e. array number and position on array, in our models.

Nutrition assessments

The dietary intake data from all cohorts used in this thesis were measured using FFQs. The strength of using an FFQ is its ability to measure habitual intake and that they are easily completed by participants, making them particularly appealing for large-scale studies. Limitations of the FFQ are that it is self-reported and it is therefore prone to measurement error, which may result in to misclassification of the exposure. In general, subjects tend to overestimate healthy food items, such as fruits and vegetables, but underestimate unhealthy products.[22, 23] If these measurement errors are non-differential, i.e. these are random and not related to the outcome of interest, this may attenuate associations between nutrient intake and the outcome of interest. However, misclassification could also be differential, i.e. it is related to the outcome of interest. For instance, obese subjects are more likely to underreport their food intake, especially fat intake.[24] Due to measurement error, it is not feasible to estimate the exact amount of intake of a certain nutrient with use of an FFQ. However, it has been shown to be a good method for ranking participants by intake of nutrients, making it a suitable method to examine associations with health outcomes.[25] The FFQs used in our studies were evaluated against other methods that measure dietary intake which demonstrated the ability to adequately rank participants according to their intake. In the first two cohorts of the Rotterdam Study dietary intake was assessed with an FFQ of 170 food items. This FFQ was validated against fifteen 24h food records and four 24h urinary urea excretion samples in a subsample of 80 participants of the Rotterdam Study. Pearson's correlations for nutrient intakes with the food records ranged between 0.44 and 0.85 and Spearman's correlation for protein intake against urinary urea was 0.67.[26] For the third cohort, dietary intake was assessed using a self-administered 389 item semi quantitative FFQ. This FFQ was previously validated in two other Dutch populations using a 9-day dietary record[27] and a 4 week dietary history,[28] which showed Pearson's correlations for intakes of different nutrients varying from 0.40 to 0.86. In The Generation R Study the FFQ was validated against three 24h-recalls, obtained by trained nutritionists, in a representative sample of 32 Dutch children aged 14 months.[29] Intraclass correlation coefficients for nutrient intakes ranged from 0.36 to 0.74.[30] The FFQ was only validated for Dutch children, while we had a multi-ethnic study population. Therefore, we performed sensitivity analyses in Dutch children only, which showed that all results were similar to those of the

total study population. In our study populations the variation of intakes of nutrients was relatively low. This made it difficult to detect an association between dietary intake of certain nutrients and any of the cardiometabolic outcomes evaluated. Considering that intake of many nutrients are correlated with total energy intake, associations with cardiometabolic outcomes may be confounded by energy intake. To account for this potential bias, we adjusted our nutrients of interest for energy using two methods: the residual method and the macronutrient substitution model.[31] Besides the adjustment for energy intake, these methods will also reduce the measurement error.[25, 32] In addition to energy intake, we adjusted for several medical, demographic and lifestyle variables. However, as in all observational studies, residual confounding may still be present.

Cardiometabolic health

For the studies included in this thesis our outcomes of interest consisted of various cardiometabolic risk factors in both children and adults, including body composition and lipid levels. In addition, in adults we studied incidence of cardiometabolic diseases, i.e. diabetes and coronary heart disease. When studying disease incidence as an outcome of interest, a large sample size is needed to provide sufficient statistical power and enable us to detect associations. On the other hand, studying cardiometabolic risk factors as outcome of interest requires less power, and provides the opportunity to investigate underlying mechanisms. Furthermore, by identifying determinants of risk factors, early preventive measures can be taken, i.e. before disease incidence occurs. Nevertheless, the extent to which cardiometabolic diseases can be predicted by biomarkers and risk factors should be taken into account. Obesity, defined as having a BMI of >30, is one of the major risk factors for cardiometabolic diseases.[33] Therefore, we aimed to study determinants that lead to an increase of BMI. This will not only provide insights in potential preventive strategies for reducing obesity, it may consequently prevent diabetes and CHD. The use of BMI as outcome of interest is practical and low in costs, making it ideal for large-scale studies. However, this measure is limited regarding prediction of cardiometabolic health because it does not take into account body composition or distribution of body fat.[34, 35] A higher fat mass is detrimental for cardiometabolic health, whereas a higher fat-free mass is associated with beneficial cardiometabolic health.[36] Furthermore, distribution of body fat is of importance when predicting cardiometabolic health. Abdominal fat may lead to an increased risk of cardiometabolic diseases, but this may not be the case for gynoid fat. One solution could be waist circumference or waist-to-hip ratio, as these measurements provide a better estimate of abdominal fat.[37] However, using detailed measures of body composition is preferred over indirect measures.[38] Therefore, a strength of our studies is the assessment of body composition and distribution of body fat with the use of DXA. By making a distinction between fat mass and fat-free mass, which may have different effects on later health outcomes, we were able to show that the positive association between protein intake and BMI was mainly driven by FMI. Still, to what extent body composition during childhood predicts future risk of cardiometabolic diseases could be debated, as it is more challenging to demonstrate this due to the time-lapse. Nevertheless, previous studies have shown that cardiometabolic risk factors most certainly track from childhood to adulthood and increase the risk of cardiometabolic diseases.[39-41] Hence, targeting factors associated with cardiometabolic health during childhood could be a good strategy to prevent later cardiometabolic diseases.

Clinical implications and directions for future research

Considering that our studies are based on observational data, the possibility to determine a causal relation is limited. However, some of our results confirm previously reported effects shown in trials. The causal effect of a higher protein intake in early childhood on a higher weight and BMI has been confirmed in a randomized controlled trial among 1138 children comparing the effect of higher- and lower-protein formula during the first year of life on growth up to 6 years of age.[42, 43] Our findings regarding protein intake and growth and body composition, confirms these results. In addition, we have shown that this association persists up to the age of 9 years. Furthermore, the association between a higher animal protein intake and BMI was driven by a higher fat mass and not fat-free mass. This increase of fat mass in response to animal protein could lead to detrimental health outcomes in later life. Promoting a lower protein intake, especially animal protein, in early life could be an effective strategy to prevent childhood adiposity, and consequently lower the risk of cardiometabolic diseases later in life. In line with this, in adults higher animal protein intake may also lead to unfavorable cardiometabolic health, as it was associated with a higher risk of CHD. In contrast, plant protein was associated with a lower risk. These findings are in line with current guidelines stating that consumption of animal-based food items should be limited. Furthermore, findings from our studies suggest that higher intakes of certain methyl donor nutrients, including vitamin B2, vitamin B6, and folate are associated with favorable cardiometabolic health outcomes, such as a lower BMI and lower risk of type 2 diabetes. However, evidence from epidemiological studies is still scarce, thus future studies are needed to replicate these findings in other populations. In contrast to our hypothesis, a higher intake of total vitamin B12 was associated with unfavorable cardiometabolic health outcomes in both children and adults. Since vitamin B12 naturally occurs in animal products, this could explain these unexpected findings. Interestingly, when we separated our analyses for vitamin B12 from foods and supplements in relation to type 2 diabetes (chapter 5.2), results were in opposite direction. A higher vitamin B12 intake from food was associated with a higher risk of type 2 diabetes, whereas a higher intake of vitamin B12 from supplements was associated with a lower risk of type 2 diabetes. This suggests that the detrimental associations we observe for vitamin B12 may indeed be due to the foods in which it is present: animal based food. This could be elucidated by implementing randomized controlled trials with vitamin B12 supplementation, in which the food source of vitamin B12 not an issue.

One of the mechanisms that may mediate the association between nutrition and cardiometabolic health is DNA methylation. In our studies on DNA methylation and cardiometabolic risk factors we have replicated several CpG sites, which were also previously reported by several studies, namely *ABCG1*, *CPT1A*, and *SREBF1*. These CpG sites could be used as a marker to predict risk of cardiometabolic diseases and may act as therapeutic targets. Future studies should further investigate how nutrition, including methyl donor nutrients, affect DNA methylation at these CpG sites. In other cases, our findings create new hypotheses, such as methylation of *DHCR24* that was associated with HDL-C. Although it is biologically plausible that this CpG site plays a role in cholesterol metabolism, replication is still required to confirm whether this newly identified CpG sites is indeed associated with HDL-C. Furthermore, the use of different study designs are required to determine potential causality, for instance with the use of the Mendelian randomization approach.[44, 45] As we did not have any longitudinal information on DNA methylation, we could not examine DNA methylation changes

over time. In addition, when investigating DNA methylation at a genome-wide level a large sample size is required to obtain sufficient statistical power. Therefore, collaboration of cohort studies in consortia is required to achieve this.

CONCLUSION

In conclusion, findings from this thesis suggest that higher intakes of protein, especially animal protein, may be detrimental for cardiometabolic health in children as well as adults. Furthermore, we observed that intake of some methyl donor nutrients, such as vitamin B2, vitamin B6, and folate, could be beneficial for cardiometabolic health. Epigenetic mechanisms may be involved in these associations, as some of these nutrients have been shown to be associated with DNA methylation. In addition, we have identified and replicated CpG sites involved in cardiometabolic pathways which were differentially methylated in relation to cardiometabolic risk factors. These CpG sites might serve as biomarkers to identify subjects at high risk of cardiometabolic diseases and could be used as therapeutic targets. Still, further large-scale studies are needed to replicate these findings and further elucidate the role of nutrition in DNA methylation and its effects on cardiometabolic health.

REFERENCES

1. Anderson OS, Sant KE, Dolinoy DC. Nutrition and epigenetics: an interplay of dietary methyl donors, one-carbon metabolism and DNA methylation. The Journal of nutritional biochemistry. 2012;23(8):853-9.
2. Park LK, Friso S, Choi SW. Nutritional influences on epigenetics and age-related disease. Proc Nutr Soc. 2012;71(1):75-83.
3. Glier MB, Green TJ, Devlin AM. Methyl nutrients, DNA methylation, and cardiovascular disease. Mol Nutr Food Res. 2014;58(1):172-82.
4. Milagro FI, Mansego ML, De Miguel C, Martinez JA. Dietary factors, epigenetic modifications and obesity outcomes: progresses and perspectives. Mol Aspects Med. 2013;34(4):782-812.
5. Altmann S, Murani E, Schwerin M, Metges CC, Wimmers K, Ponsuksili S. Maternal dietary protein restriction and excess affects offspring gene expression and methylation of non-SMC subunits of condensin I in liver and skeletal muscle. Epigenetics. 2012;7(3):239-52.
6. Altmann S, Murani E, Schwerin M, Metges CC, Wimmers K, Ponsuksili S. Dietary protein restriction and excess of pregnant German Landrace sows induce changes in hepatic gene expression and promoter methylation of key metabolic genes in the offspring. J Nutr Biochem. 2013;24(2):484-95.
7. Cordero P, Campion J, Milagro FI, Martinez JA. Transcriptomic and epigenetic changes in early liver steatosis associated to obesity: effect of dietary methyl donor supplementation. Molecular genetics and metabolism. 2013;110(3):388-95.
8. Guay S-P, Voisin G, Brisson D, Munger J, Lamarche B, Gaudet D, et al. Epigenome-wide analysis in familial hypercholesterolemia identified new loci associated with high-density lipoprotein cholesterol concentration. Epigenomics. 2012;4(6):623-39.
9. Irvin MR, Zhi D, Joehanes R, Mendelson M, Aslibekyan S, Claas SA, et al. Epigenome-wide association study of fasting blood lipids in the genetics of lipid-lowering drugs and diet network study. Circulation. 2014;130(7):565-72.
10. Pfeiffer L, Wahl S, Pilling LC, Reischl E, Sandling JK, Kunze S, et al. DNA methylation of lipid-related genes affects blood lipid levels. Circ Cardiovasc Genet. 2015;8(2):334-42.
11. Drzewinska J, Walczak-Drzewiecka A, Ratajewski M. Identification and analysis of the promoter region of the human DHCR24 gene: involvement of DNA methylation and histone acetylation. Molecular biology reports. 2011;38(2):1091-101.
12. Zerenturk EJ, Sharpe LJ, Ikonen E, Brown AJ. Desmosterol and DHCR24: unexpected new directions for a terminal step in cholesterol synthesis. Progress in lipid research. 2013;52(4):666-80.
13. Sakakibara S, Nakamura Y, Satoh H, Okano H. Rna-binding protein Musashi2: developmentally regulated expression in neural precursor cells and subpopulations of neurons in mammalian CNS. J Neurosci. 2001;21(20):8091-107.
14. Do DN, Strathe AB, Ostersen T, Jensen J, Mark T, Kadarmideen HN. Genome-wide association study reveals genetic architecture of eating behavior in pigs and its implications for humans obesity by comparative mapping. PLoS One. 2013;8(8):e71509.
15. Sohm B, Sissler M, Park H, King MP, Florentz C. Recognition of human mitochondrial tRNALeu(UUR) by its cognate leucyl-tRNA synthetase. Journal of Molecular Biology. 2004;339(1):17-29.
16. Munakata K, Iwamoto K, Bundo M, Kato T. Mitochondrial DNA 3243A>G mutation and increased expression of LARS2 gene in the brains of patients with bipolar disorder and schizophrenia. Biol Psychiatry. 2005;57(5):525-32.
17. McElroy SL, Kotwal R, Malhotra S, Nelson EB, Keck PE, Nemeroff CB. Are mood disorders and obesity related? A review for the mental health professional. J Clin Psychiatry. 2004;65(5):634-51, quiz 730.

18. Nohr EA, Frydenberg M, Henriksen TB, Olsen J. Does low participation in cohort studies induce bias? Epidemiology. 2006;17(4):413-8.
19. Rakyan VK, Down TA, Balding DJ, Beck S. Epigenome-wide association studies for common human diseases. Nature Reviews Genetics. 2011;12(8):529-41.
20. Joubert BR, Felix JF, Yousefi P, Bakulski KM, Just AC, Breton C, et al. DNA Methylation in Newborns and Maternal Smoking in Pregnancy: Genome-wide Consortium Meta-analysis. Am J Hum Genet. 2016;98(4):680-96.
21. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC bioinformatics. 2012;13(1):86.
22. Amanatidis S, Mackerras D, Simpson JM. Comparison of two frequency questionnaires for quantifying fruit and vegetable intake. Public Health Nutr. 2001;4(2):233-9.
23. Scagliusi FB, Polacow VO, Artioli GG, Benatti FB, Lancha AH, Jr. Selective underreporting of energy intake in women: magnitude, determinants, and effect of training. J Am Diet Assoc. 2003;103(10):1306-13.
24. Goris AH, Westerterp-Plantenga MS, Westerterp KR. Undereating and underrecording of habitual food intake in obese men: selective underreporting of fat intake. Am J Clin Nutr. 2000;71(1):130-4.
25. Kipnis V, Subar AF, Midthune D, Freedman LS, Ballard-Barbash R, Troiano RP, et al. Structure of dietary measurement error: results of the OPEN biomarker study. American Journal of Epidemiology. 2003;158(1):14-21.
26. Klipstein-Grobusch K, Den Breeijen JH, Goldbohm RA, Geleijnse JM, Hofman A, Grobbee DE, et al. Dietary assessment in the elderly: Validation of a semiquantitative food frequency questionnaire. European journal of clinical nutrition. 1998;52(8):588-96.
27. Goldbohm RA, van den Brandt PA, Brants HA, van't Veer P, Al M, Sturmans F, et al. Validation of a dietary questionnaire used in a large-scale prospective cohort study on diet and cancer. Eur J Clin Nutr. 1994;48(4):253-65.
28. Feunekes GI, Van Staveren WA, De Vries JH, Burema J, Hautvast JG. Relative and biomarker-based validity of a food-frequency questionnaire estimating intake of fats and cholesterol. Am J Clin Nutr. 1993;58(4):489-96.
29. Kiefte-de Jong JC, de Vries JH, Bleeker SE, Jaddoe VWV, Hofman A, Raat H, et al. Socio-demographic and lifestyle determinants of 'Western-like'and 'Health conscious' dietary patterns in toddlers. British Journal of Nutrition. 2013;109(01):137-47.
30. Voortman TV, Kiefte-de Jong JC, Geelen A, Villamor E, Hofman A, Jaddoe VWV, et al. A diet score to assess diet quality in preschool children: Development and determinants in the Generation R study. Submitted.
31. Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. The American journal of clinical nutrition. 1997;65(4):1220S-8S.
32. Freedman LS, Schatzkin A, Midthune D, Kipnis V. Dealing with dietary measurement error in nutritional cohort studies. Journal of the National Cancer Institute. 2011;103(14):1086-92.
33. Berrington de Gonzalez A, Hartge P, Cerhan JR, Flint AJ, Hannan L, MacInnis RJ, et al. Body-mass index and mortality among 1.46 million white adults. The New England journal of medicine. 2010;363(23):2211-9.
34. Nevill AM, Stewart AD, Olds T, Holder R. Relationship between adiposity and body size reveals limitations of BMI. Am J Phys Anthropol. 2006;129(1):151-6.
35. Freedman DS, Wang J, Ogden CL, Thornton JC, Mei Z, Pierson RN, et al. The prediction of body fatness by BMI and skinfold thicknesses among children and adolescents. Ann Hum Biol. 2007;34(2):183-94.

36. Srikanthan P, Karlamangla AS. Relative muscle mass is inversely associated with insulin resistance and prediabetes. Findings from the third National Health and Nutrition Examination Survey. J Clin Endocrinol Metab. 2011;96(9):2898-903.
37. Klein S, Allison DB, Heymsfield SB, Kelley DE, Leibel RL, Nonas C, et al. Waist circumference and cardiometabolic risk: a consensus statement from shaping America's health: Association for Weight Management and Obesity Prevention; NAASO, the Obesity Society; the American Society for Nutrition; and the American Diabetes Association. Diabetes Care. 2007;30(6):1647-52.
38. Duren DL, Sherwood RJ, Czerwinski SA, Lee M, Choh AC, Siervogel RM, et al. Body composition methods: comparisons and interpretation. J Diabetes Sci Technol. 2008;2(6):1139-46.
39. Morrison JA, Glueck CJ, Woo JG, Wang P. Risk factors for cardiovascular disease and type 2 diabetes retained from childhood to adulthood predict adult outcomes: the Princeton LRC Follow-up Study. International journal of pediatric endocrinology. 2012;2012(1):6.
40. Katzmarzyk PT, Perusse L, Malina RM, Bergeron J, Despres JP, Bouchard C. Stability of indicators of the metabolic syndrome from childhood and adolescence to young adulthood: the Quebec Family Study. J Clin Epidemiol. 2001;54(2):190-5.
41. Morrison JA, Friedman LA, Gray-McGuire C. Metabolic syndrome in childhood predicts adult cardiovascular disease 25 years later: the Princeton Lipid Research Clinics Follow-up Study. Pediatrics. 2007;120(2):340-5.
42. Koletzko B, von Kries R, Closa R, Escribano J, Scaglioni S, Giovannini M, et al. Lower protein in infant formula is associated with lower weight up to age 2 y: a randomized clinical trial. Am J Clin Nutr. 2009;89(6):1836-45.
43. Weber M, Grote V, Closa-Monasterolo R, Escribano J, Langhendries JP, Dain E, et al. Lower protein content in infant formula reduces BMI and obesity risk at school age: follow-up of a randomized trial. Am J Clin Nutr. 2014;99(5):1041-51.
44. Wahl S, Drong A, Lehne B, Loh M, Scott WR, Kunze S, et al. Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. Nature. 2017;541(7635):81-6.
45. Dekkers KF, van Iterson M, Slieker RC, Moed MH, Bonder MJ, van Galen M, et al. Blood lipids influence DNA methylation in circulating cells. Genome Biol. 2016;17(1):138.

6.2

Summary

SUMMARY

Chapter 1 provides a general introduction on nutrition, DNA methylation and cardiometabolic health and a description of the studies on which this thesis is based.

In **Chapter 2** we discuss the associations between nutrition and DNA methylation. **Chapter 2.1** provides an overview of the current evidence on the associations between different nutrients with DNA methylation in humans across all stages of the life course (i.e. during pregnancy, infancy, childhood, adolescence, and adulthood). In this systematic review we identified 3774 references, of which 98 studies met all inclusion criteria. The majority examined associations in adults and folate was the main nutrient of interest. Several candidate gene and epigenome-wide association studies reported differential DNA methylation of CpG sites in response to folate (e.g. IGF2, H19, HOX), fatty acids (e.g. PPRAGC1A, TNFα), and vitamin D (CYP24A1). Some of these observed associations were specific to life course stage (e.g. IGF2 in early life) and tissue (e.g. opposite directions for PPRAGC1A in muscle vs. fat tissue). To date, promising results have been reported in the field of nutrition and DNA methylation in humans at different stages across the life-course. Especially for nutrients known to be involved in one-carbon metabolism, such as folate, but also others, such as fatty acids and vitamin D. Studies on other nutrients, such as other macronutrients and several minerals are still scarce. Further large-scale studies of high quality are needed to expand our understanding on the role of nutrition in DNA methylation and its effects on health and disease.

In **Chapter 3** we discuss the associations of DNA methylation and cardiometabolic risk factors. **Chapter 3.1** provides an overview of studies investigating the association between DNA methylation (global, gene-specific, and genome-wide) and lipid levels (triacylglycerol, total cholesterol, LDL-cholesterol, and HDL-cholesterol). In this systematic review we included 31 publications, based on 23 unique studies. There were no consistent associations observed for global DNA methylation and lipid levels. Gene-specific and genome-wide studies showed significant associations between lipid levels and DNA methylation at *ABCG1*, *CPT1A*, *SREBF1*, *TNNT1*, *MIR33B*, and *TNIP*, which were replicated. In **Chapter 3.2** presents results of an EWAS on lipid levels in the Rotterdam Study. In this EWAS we replicated previously reported associations for methylation of *ABCG1*, *CPT1A*, and *SREBF1* with triacylglycerol or HDL-cholesterol. In addition, we observed a novel association between DNA methylation of a CpG site near *DHCR24* and HDL-cholesterol. The *DHCR24* gene encodes for 3-hydroxysterol-24 reductase, a cholesterol biosynthesis enzyme that catalyzes the conversion of desmosterol to cholesterol. **Chapter 3.3** presents the findings of an EWAS on BMI and WC in the Rotterdam Study and replication of these findings in a USA-based study. We identified and replicated 12 CpG sites associated with BMI and 13 CpG sites associated with WC. These included several CpG sites, which are known to be related to cardiometabolic outcomes, such as *ABCG1* and *CPT1A*. In addition to these known CpG sites, we identified two novel CpG sites related to both BMI and WC, namely *MSI2* and *LARS2*. *MSI2* encodes RNA-binding proteins and plays a central role in posttranscriptional gene regulation, and has been suggested to be associated with eating behaviors. *LARS2* encodes an enzyme that catalyzes aminoacylation of mitochondrial tRNALeu and may be linked to bipolar disorder.

Chapter 4 focuses on associations of nutrition in early childhood with cardiometabolic health in children at school age from The Generation R Study. **Chapter 4.1** describes associations between

intake of methyl donor nutrients, including vitamin B6, vitamin B12, folate, and methionine with growth and body composition. We observed that high intake of folic acid at the age of 1 year was associated with a lower weight and BMI at the age of 6 years, whereas a high methionine intake was associated with an unfavorable body composition, including a higher BMI and fat percentage. **Chapter 4.2** shows the association between protein intake with repeatedly measured growth. We observed that a higher protein intake, in particular animal protein, at 1 year was associated with a higher height, weight, and BMI, up to the age of 9 years. **Chapter 4.3** presents associations between protein intake with detailed body composition. We found that the positive association between protein intake and BMI was driven by a higher fat mass index. **Chapter 4.4** describes associations between different types of fatty acids with body composition and cardiometabolic health in childhood. Intake of fat or fatty acids at the age of 1 year, however, were not associated with body composition or with other aspects of cardiometabolic health in childhood**.**

Chapter 5 focuses on associations between nutrition and cardiometabolic health in adults. **Chapter 5.1** presents the association between folate and vitamin B12 with body composition in the B-PROOF study. higher folate levels were associated with a lower BMI. This association was most pronounced in participants with prevalent cardiometabolic diseases, and was mainly driven by FFMI. In **Chapter 5.2** we studied the role of methyl donor nutrients and the risk of diabetes using data from the Nurses' Health Study, Nurses' Health Study II, and Health Professionals Follow-Up Study. We observed that higher intakes of vitamin B2 and vitamin B6 are associated with a lower risk of diabetes. This association was attenuated slightly, but remained significant, after adjustment of several sociodemographic and lifestyle covariates. **Chapter 5.3** presents the association between macronutrient intake and the risk of CHD in The Rotterdam Study. Although we observed no association between total protein intake, when we separated analyses for animal and vegetable protein we observed significant associations in opposite directions. A higher animal protein intake was associated with a higher risk of CHD, whereas a higher plant protein intake was associated with a lower risk of CHD. Carbohydrates, fat and their subgroups were not associated with CHD incidence at the expense of any other macronutrient.

In **Chapter 6**, interpretation of the main findings are discussed as well as the methodological considerations, implications, and recommendation for future studies.

6.3

Nederlandse samenvatting

NEDERLANDSE SAMENVATTING

In **Hoofdstuk 1** van dit proefschrift wordt een algemene introductie van voeding, DNA methylatie en cardiometabole gezondheid gegeven en worden de studies waarop dit proefschrift is gebaseerd beschreven.

In **Hoofdstuk 2** beschrijft de associaties tussen voeding en DNA methylatie. **Hoofdstuk 2.1** geeft een overzicht van het huidige bewijs van de associaties tussen verschillende nutriënten met DNA methylatie. In dit systematische literatuuronderzoek hebben we 3774 referenties geïdentificeerd, waarvan 98 aan alle inclusiecriteria voldeden. De meerderheid onderzocht de associaties in volwassenen. Van alle nutriënten was foliumzuur het meest bestudeerd. Meerdere gen-specifieke en epigenoom-brede associatie studies rapporteerden differentiële DNA methylatie van CpG sites in relatie tot foliumzuur (o.a. IGF2, H19, HOX), vetzuren (e.g. PPRAGC1A, TNFα), en vitamine D (CYP24A1). Een aantal van deze geobserveerde associaties waren specifiek voor bepaalde levensloopfases, zoals IGF2 in het vroege levensstadium. Daarnaast waren sommige associaties specifiek voor weefsels, zo werden associaties geobserveerd tussen vetinname en methylatie van PPRAGC1A, maar deze waren in tegengestelde richtingen voor vetweefsel en spierweefsel. Veelbelovende resultaten zijn gerapporteerd op het gebied van voeding en DNA methylatie in mensen op verschillende momenten van de levensloop. Vooral voor nutriënten waarvan bekend is dat deze betrokken zijn in het "one-carbon metabolism", zoals foliumzuur, maar ook voor andere nutriënten, zoals vetzuren en vitamine D. Studies die zich richten op andere nutriënten, zoals andere macronutriënten en verschillende mineralen zijn schaars. Verdere grootschalige studies van hoge kwaliteit zijn nodig om onze kennis te vergoten van de rol van voeding in DNA methylatie en de effecten hiervan op gezondheid en ziekte.

In **Hoofdstuk 3** worden de associaties tussen DNA methylatie en cardiometabole risicofactoren besproken. **Hoofdstuk 3.1** geeft een overzicht van de studies die de associatie tussen DNA methylatie, zowel globaal, gen-specifiek en epigenoom-breed, met lipide gehalte (inclusief triglyceride, totaal cholesterol, LDL-cholesterol en HDL-cholesterol). In dit systematische literatuuronderzoek includeerden we 31 publicaties, gebaseerd op 23 unieke studies. Er waren geen consistente associaties geobserveerd voor globale DNA methylatie en lipide gehalte. Gen-specifieke en epigenoom-brede studies lieten significante associaties zien tussen lipide gehalte en DNA methylatie op *ABCG1*, *CPT1A*, *SREBF1*, *TNNT1*, *MIR33B* en *TNIP.* Deze bevindingen werden ook gerepliceerd. **Hoofdstuk 3.2** presenteert de resultaten van een epigenoom-brede associatie studie op lipide gehalte. In deze studie hebben we eerder gerapporteerde resultaten gerepliceerd voor methylatie van *ABCG1, CPT1A* en *SREBF1* met triglyceriden of HDL-cholesterol. Daarnaast observeerden we een nieuwe associatie tussen methylatie van een CpG-site die gelokaliseerd is bij *DHCR24* en HDL-cholesterol. Het *DHCR24* gen codeert voor 3-hydroxysterol-24 reductase, een cholesterolbiosynthese enzym dat de omzetting van desmosterol naar cholesterol katalyseert. In **Hoofdstuk 3.3** presenteren we de bevindingen van een epigenoom-brede associatie studie op BMI en middelomtrek in de Rotterdam Study en replicatie van deze bevindingen in een VS-gebaseerde studie. We identificeerden en repliceerden 12 CpG sites geassocieerd met BMI en 13 CpG sites geassocieerd met middelomtrek. Tot deze bevindingen hoorden enkele CpG sites waarvan bekend is dat deze geassocieerd zijn met cardiometabole uitkomsten, zoals *ABGG1* en *CPT1A*. Hiernaast identificeerden we twee nieuwe CpG sites welke beide gerelateerd waren aan BMI en middelomtrek, namelijk *MSI2* en *LARS2.* MSI2 codeert voor RNA-bindende pro-

teïnen en speelt een belangrijke rol in posttranscriptionele genregulatie, en is mogelijk geassocieerd met eetgedrag. LARS2 codeert voor een enzym dat de aminoacylatie van mitochondriale tRNALeu katalyseert en is mogelijk gerelateerd aan bipolaire stoornis.

Hoofdstuk 4 richt zich op de associatie tussen voeding in de vroege jeugd en cardiometabole gezondheid in kinderen op schoolleeftijd van de Generation R Study. In **Hoofdstuk 4.1** worden associaties tussen inname van methyl donor nutriënten, inclusief vitamine B6, vitamine B12, foliumzuur, en methionine met groei en lichaamssamenstelling beschreven. We observeerden dat een hoge inname van foliumzuur op de leeftijd van 1 jaar geassocieerd was met een ongunstige lichaamssamenstelling, inclusief een hogere BMI en vetpercentage. **Hoofstuk 4.2** laat de associaties zien tussen eiwitinname en herhaaldelijk gemeten groei. We observeerden dat een hogere eiwitinname, met name dierlijk eiwit, op 1 jaar geassocieerd was met een hogere lengte, gewicht en BMI tot en met de leeftijd van 9 jaar. **Hoofdstuk 4.3** presenteert de associaties tussen eiwitinname en gedetailleerde lichaamssamenstelling. We vonden dat de positieve associatie tussen eiwitinname en BMI voornamelijk werd gedreven door een hogere vetmassa. **Hoofdstuk 4.4** beschrijft de associaties tussen verschillende soorten vetzuren met lichaamssamenstelling en cardiometabole gezondheid in kinderen. Echter was de inname van vet of vetzuren op de leeftijd van 1 jaar niet geassocieerd met groei, lichaamssamenstelling of andere aspecten van cardiometabole gezondheid.

Hoofdstuk 5 richt zich op de associatie tussen voeding en cardiometabole gezondheid in volwassenen. **Hoofdstuk 5.1** presenteert de associaties tussen bloedwaarde van foliumzuur en vitamine B12 met lichaamssamenstelling in de B-PROOF Study. We observeerden dat hogere waarden van foliumzuur geassocieerd waren met een lagere BMI. Deze associatie was het sterkst in deelnemers met cardiometabole ziekten, en werd voornamelijk gedreven door vetvrije massa. In **Hoofdstuk 5.2** bestudeerden we de rol van methyl donor nutriënten en het risico op diabetes met gebruik van de data van de Nurses' Health Study, Nurses' Health Study II, and Health Professionals Follow-Up Study. We observeerden dat een hogere inname van vitamine B2 en vitamine B6 geassocieerd was met een verlaging op het risico van type II diabetes. Deze associatie was enigszins verzwakt, maar bleef significant na het adjusteren voor verschillende sociodemografische en levensstijl-covariaten. **Hoofdstuk 5.3** presenteert de associatie tussen inname van macronutriënten en het risico op CHD in de Rotterdam Study. Alhoewel we geen associatie observeerden voor totale eiwitinname, waren dierlijk en plantaardig eiwit in tegenovergestelde richtingen geassocieerd met CHD. Een hogere inname van dierlijk eiwit was geassocieerd met een hoger risico op CHD, terwijl een hogere inname van plantaardig eiwit geassocieerd was met een lager risico op CHD. Koolhydraten, vetten, en de subgroepen hiervan waren niet geassocieerd met CHD incidentie ten koste van een van de andere macronutriënten.

In **Hoofdstuk 6** worden de interpretaties van de belangrijkste bevindingen besproken en de methodologische aspecten, implicaties en aanbevelingen voor vervolgstudies.

Chapter 7

Appendices

List of manuscripts
PhD portfolio
About the author
Dankwoord

LIST OF MANUSCRIPTS

Braun KVE*, Portilla E*, Chowdhury R, Nano J, Troup J, Voortman T, Franco OH, Muka T. "*The role of epigenetic modifications in cardiometabolic diseases*", in: Moskalev A, Vaiserman AM. "*Epigenetics of Aging and Longevity: Translational Epigenetics*", Academic Press; 2017; p. 347-64.

Dhana K*, **Braun KVE***, Nano J, Voortman T, Demerath EW, Guan W, Fornage M, van Meurs JBJ, Uitterlinden AG, Hofman A, Franco OH, Dehghan A. Epigenome-wide association study (EWAS) on obesity-related traits. *American Journal of Epidemiology. 2018. Accepted for publication in American Journal of Epidemiology.*

Braun KVE, Dhana K, de Vries PS, Voortman T, van Meurs JBJ, Uitterlinden AG, Hofman A, Hu FB, BIOS consortium, Franco OH, Dehghan A. Epigenome-wide association study (EWAS) on lipids: the Rotterdam Study. *Clinical Epigenetics. 2017;9:15.*

V Jen, NS Erler, MJ Tielemans, **KVE Braun**, VWV Jaddoe, OH Franco, T Voortman. Mothers' intake of sugar-containing beverages during pregnancy and body composition of their children during childhood: The Generation R Study. *The American Journal of Clinical Nutrition. 2017;105(4):834-841.*

Braun KVE*, Voortman T*, Dhana K, Troup J, Bramer WM, Troup J, Chowdhury R, Dehghan A, Muka T, Franco OH. The role of DNA methylation in dyslipidaemia: A systematic review. *Progress in Lipid Research. 2016;64:178-91.*

Stroobant W*, **Braun KVE***, Kiefte-de Jong JC, Moll HA, Jaddoe VWV, Brouwer IA, Franco OH, Voortman T. Intake of Different Types of Fatty Acids in Infancy Is Not Associated with Growth, Adiposity, or Cardiometabolic Health up to 6 Years of Age. *Journal of Nutrition. 2017;147(3):413-20.*

Braun KVE, Erler NS, Kiefte-de Jong JC, Jaddoe VW, van den Hooven EH, Franco OH, Voortman T. Dietary Intake of Protein in Early Childhood Is Associated with Growth Trajectories between 1 and 9 Years of Age. *Journal of Nutrition. 2016;146(11):2361-7.*

Voortman T, **Braun KVE**, Kiefte-de Jong JC, Jaddoe VW, Franco OH, van den Hooven EH. Protein intake in early childhood and body composition at the age of 6 years: The Generation R Study. *International Journal of Obesity (Lond). 2016;40(6):1018-25.*

T Muka*, J Nano*, T Voortman, **KVE Braun**, S Ligthart, S Stranges, WM Bramer, J Troup, R Chowdhury, A Dehghan, OH Franco. The role of Global and Regional DNA Methylation and Histone Modifications in Type 2 Diabetes: A Systematic Review. *Nutr Metab Cardiovasc Dis. 2016;26(7):553-66.*

RMA van Gijssel*, **KVE Braun***, JC Kiefte-de Jong, VWV Jaddoe, OH Franco, T Voortman. Association between dietary fibre intake in infancy and cardiometabolic health at school age: The Generation R Study. *Nutrients. 2016;8(9).*

Braun KVE*, Voortman T*, Kiefte-de Jong JC, Jaddoe VWV, Hofman A, Franco OH, van den Hooven EH. Dietary Intakes of Folic Acid and Methionine in Early Childhood Are Associated with Body Composition at School Age. *Journal of Nutrition. 2015;145(9):2123-9.*

T Voortman, EH van den Hooven, **KVE Braun**, M van den Broek, WM Bramer, R Chowdhurry, OH Franco. Effects of polyunsaturated fatty acid intake and status during pregnancy, lactation, and early childhood on cardiometabolic health: a systematic review. *Progress in Lipid Research. 2015;59:67-87.*

T Voortman, AH Garcia*, **KVE Braun***, SK Thakkar, MJ Tielemans, W Stroobant, G Mutungi, WM Bramer, JD Schoufour, F Giuffrida, I Silva-Zolezzi, OH Franco. Effects of maternal nutrition on quantity and nutritional quality of breast milk: a systematic review. *Submitted*

Oliai Araghi S*, **Braun KVE***, van der Velde N, van Dijk SC, van Schoor NM, Zillikens C, de Groot L, Uitterlinden A, Stricker B, Voortman T*, Kiefte-de Jong JC*. Associations of serum folate and vitamin B12 with body composition in elderly: The B-PROOF study. Submitted for publication. *Submitted for publication.*

Braun KVE, Satija A, Voortman T, Franco OH, Sun Q, Bhupathiraju SN, Hu FB. Methyl donor nutrient intake and incidence of type 2 diabetes mellitus: results from the Nurses' Health Study and Health Professionals Follow-Up Study. *Manuscript in preparation*

Girschik C*, **Braun KVE***, Franco OH, Voortman T. Association between macronutrient intake and incidence of coronary heart disease (CHD): The Rotterdam Study. *Manuscript in preparation*

Braun KVE*, Mandaviya P*, Franco OH, Nano J, Girschik C, Bramer WM, Muka T, Troup J, van Meurs JBJ, Heil SG, Voortman T. Nutrients and DNA methylation across the life course: a systematic review. *Submitted for publication.*

V Jen, **KVE Braun**, LG Karagounis, AN Nguyen, VWV Jaddoe, JD Schoufour, OH Franco, T Voortman. Longitudinal association of dietary protein intake in infancy and adiposity throughout childhood. *Submitted for publication*

J Merino*, M Guasch-Ferré*, C Ellervik*, HS Dashti*, SJ Sharp, P Wu, K Overvad, C Sarnowski, M Kuokkanen, RN Lemaitre, AE Justice, U Ericson, **KVE Braun**, Y Mahendran, AC Frazier-Wood, D Sun, AY Chu, T Tanaka, J Luan, J Hong, A Tjønneland, M Ding, A Lundqvist, K Mukamal, R Rohde, C Schulz, OH Franco, N Grarup, YI Chen, L Bazzano, PW Franks, JE Buring, C Langenberg, C Liu, T Hansen, M Jensen, K Sääksjärvi, BM Psaty, KL Young, G Hindy, M Ghanbari, CH Sandholt, PM Ridker, JM Ordovas, JB Meigs, O Pedersen, P Kraft, M Perola, D Mozaffarian, KE North, Marju Orho-Melander, T Voortman, U Torf, JI Rotter, L Qi, NG Forouhi, TIA Sørensen, MJ Stampfer, S Männistö, E Selvin, F Imamura, V Salomaa, FB Hu, NJ Wareham, J Dupuis, CE Smith*, TO Kilpeläinen*, DI Chasman*, JC Florez*. Fat intake, genetic risk, and type 2 diabetes incidence. *Submitted for publication.*

V Jen, AN Nguyen*, **KVE Braun***, E Castañeda-Gutiérrez, S Almoosawi, VWV Jaddoe, JD Schoufour, OH Franco, T Voortman. Longitudinal associations of dietary glycaemic index and glycaemic load during infancy with growth and body composition in childhood: The Generation R Study. *Submitted for publication*

M. Glisic*, E. Asllanaj*, C Ochoa Rosales, E Portilla Fernandez, V Gonzalez, **KVE Braun**, J Nano, T Voortman, M Ghanbari, OH Franco, T Muka. Sex differences of epigenetic mechanisms in cardiovascular diseases. *Submitted for publication*

**Denotes equal contribution*

PHD PORTFOLIO

Name PhD student: Kim V.E. Braun
Erasmus MC department: Epidemiology
Research School: NIHES
PhD period: July 2014 – January 2018
Promotors: Prof.dr. Oscar H. Franco
Co- promotors: Dr.ir. Trudy Voortman and Dr. Abbas Dehghan

1. PhD training	Year	ECTs
General courses		
Master of Science in Epidemiology, NIHES	2014-2015	70
Doctor of Science in Genetic Epidemiology, NIHES	2015-2016	70
Specific courses		
EndNote, Medical Library, Erasmus MC	2014	0.3
Systematic Literature Search, Medical Library, Erasmus MC	2014	0.6
Basic course on R, MolMed	2015	1
Nutrition and physical activity, University of Cambridge	2015	1
Basic Human Genetics course, MolMed	2015	1
English Biomedical writing, Erasmus MC	2016	2
Radiation protection level 5R	2016	1
Seminars and meetings		
ErasmusAGE research meetings, Erasmus MC	2014-2017	1
2020 meetings, Erasmus MC	2014-2017	1
MolEpi meetings, Erasmus MC	2014-2017	1
PhD days, Erasmus MC	2014-2017	1
Seminars at the department of Epidemiology, Erasmus MC	2014-2017	1
Seminars at the department of Nutrition, Harvard T.H. Chan School of Public Health	2016-2017	0.5
Nurses' Health Study and Health Professionals Follow-up Study meetings, Harvard T.H. Chan School of Public Health	2016-2017	0.5
Conferences		
Sophia research day, Rotterdam	2015	0.3
Rotterdam science festival, Rotterdam	2015	0.3
Nutritional Sciences Days, Heeze	2015	1
IDEAL congress, Leiden	2015	1
FENS, Berlin	2015	1
NAV public lecture, The Hague	2016	0.3
CHARGE meeting, Houston	2016	1
Clinical Epigenetics International Meeting, Dusseldorf	2016	1
Nutrition and Growth Conference, Vienna	2016	1
NAV public lecture, The Hague	2017	0.3
CHARGE meeting, New York City	2017	1
CHARGE meeting, Boston	2017	1
IUNS 21st International Congress of Nutrition, Buenos Aires	2017	1

Presentations		
Sophia research day, Rotterdam – Oral presentation	2015	0.3
European Congress of Epidemiology – Poster presentation	2015	0.3
Rotterdam science festival, Rotterdam – Oral presentation	2015	0.3
Nutritional Sciences Days, Heeze – Oral presentation	2015	0.3
FENS, Berlin – Poster presentation	2015	0.3
Clinical Epigenetics International Meeting, Dusseldorf – Poster presentation	2016	0.3
Nutrition and Growth Conference, Vienna – Oral presentation	2016	0.3
IUNS 21st International Congress of Nutrition, Buenos Aires – Oral & poster presentation	2017	0.3
Scholarships and awards		
Dr. Hendrik Muller's Vaderlandsch Fonds	2016	
Stichting Fundatie van de Vrijvrouwe van Renswoude	2016	
Erasmus Trustfonds – research visit grant	2016	
Albert Renold Travel Fellowship	2017	
Erasmus Trustfonds – conference participation grant	2017	
Other		
Reviewer of international peer-reviewed journals: JAMA, Nutrients, European Journal of Epidemiology, BMC nutrition, Diabetes Care, Diabetes, PLOS ONE	2015-2017	1
Visiting scientist Harvard T.H.Chan School of Public Health	2016-2017	1
2. Teaching		
Supervising students		
Wendy Stroobant, MSc thesis	2015	2
Rafaelle van Gijssel, MSc thesis	2015	2
Carolin Thomas, MSc thesis	2017	2
Ehab Atweh, MSc thesis	2017	2
Lectures & supervising exercises		
Science café at Rotterdam Science Festival: "The impact of nutrition on our genes"	2015	1
Guest lecture: Public Health in low and middle income countries (LMICs): Impact of health programmes – general principles, NIHES	2016	1
Online lecture: "Sustainable development goals" for medical students, Erasmus MC	2016	1
Supervising exercises: Biostatistical Methods I: basic principles, NIHES	2016	0.5
Lecture: "Global nutrition", Minor Global Health for medical students, Erasmus MC	2017	1

ABOUT THE AUTHOR

Kim Braun was born on April 13th 1989 in Rotterdam, the Netherlands. She studied Nutrition and Dietetics at The Hague University and obtained her degree as Dietician in 2012. During her Bachelor, she completed a clinical internship at the department of Dietetics at the Harbor Hospital (Havenziekenhuis) in Rotterdam. She conducted research for her bachelor thesis at the department of Internal Medicine of Erasmus Medical Center in Rotterdam where she worked on the B-PROOF study. To be able to combine her two greatest interests, nutrition and science, she decided to continue with a Master of Science in Nutrition and Health at the VU University Amsterdam. As part of the Master of Science program she performed her research project at the department of Epidemiology of Erasmus Medical Center in the ErasmusAGE group. After she graduated in 2014, Kim continued with a postgraduate research fellowship at the ErasmusAGE group and completed a Master of Science in Epidemiology at the Netherlands Institute of Health Sciences (NIHES) in 2015. Subsequently, she continued to work at the department of Epidemiology and expanded her studies in her current PhD research entitled "Nutrition and Cardiometabolic Health: The role of DNA methylation". The results of this work are presented in this thesis. In 2016, she completed the NIHES Doctor of Science program in Genetic Epidemiology. In the final year of her PhD program, Kim worked as a visiting scientist at the department of Nutrition at the Harvard T.H. Chan School of Public Health in Boston, MA, USA.

DANKWOORD

Het voltooien van dit proefschrift heb ik niet alleen gedaan, en had ik ook zeker nooit alleen kunnen doen. Ik heb hier ontzettend veel mensen om voor te bedanken. Ook al heb ik het idee dat woorden hier tekort zullen schieten, doe ik toch een poging!

Om te beginnen wil ik de deelnemers van de studies bedanken. Ik heb de privilege gehad om met data van maar liefst 7 verschillende cohorts te werken, en daarmee heb ik een enorm aantal mensen om te bedanken, welgeteld 319.380 deelnemers (ja, ik heb het uitgerekend). Hierbij wil ik alle deelnemers van The Generation R Study, Rotterdam Study, B-PROOF study, ARIC study, Nurses' Health Study, Nurses' Health Study II, en Health Professionals Follow-up Study hartelijk bedanken voor hun onmisbare bijdrage. Uiteraard gaat daarbij ook mijn dank uit naar alle medewerkers van de onderzoekscentra en datamanagement.

I would like to thank my promotor and co-promotors. Dear Oscar, thank you for this incredible opportunity. I am very grateful for all the guidance and support you have provided over the past few years. Trudy, ik had me geen betere copromotor kunnen wensen. Bij het tot stand komen van dit proefschrift en alles wat daarbij hoort heb ik ongelofelijk veel aan jou te danken. Ik heb veel bewondering en respect voor hoe jij alles doet en dat je zelfs met een overvolle agenda toch altijd voor je studenten klaar staat. Bedankt voor al je hulp zowel binnen als buiten dit proefschrift. Abbas, I feel very lucky you joined my team of supervisors and I got the chance to learn from you. Thank you for you for all the insightful discussions and patience when I just started in the field of epigenetics.

Geachte prof. Boersma, dankuwel voor het beoordelen van mijn proefschrift alsmede de taak van secretaris op u te nemen. Daarmee gaat mijn dank ook uit naar prof. van Rossum en prof. Geleijnse voor het deelnemen in de leescommissie en het beoordelen van mijn proefschrift. To the members of the large doctoral committee, thank you very much for participating in my PhD defense. It is a great honor for me to have you as members of my doctoral committee.

As the work in this thesis is mainly based on collaborations, I have many co-authors to thank. Pooja, it was great working with you, you are an excellent review-partner. It was nice to work with someone who does not only tolerate my OCD, but actually appreciates it – and may even be worse than I am! Sadaf, ik vind het geweldig om jaren na mijn afstudeerstage bij de B-PROOF study weer met je samen te werken. Nicole, thank you so much for all your support in statistical challenges. Klodian, I was very happy we were able to pick up our projects in Boston where we left of in Rotterdam. Thank you for all the discussions, brainstorm sessions on R codes, and many coffee breaks. My gratitude also goes out to all the other co-authors, both from Erasmus MC as well as the CHARGE consortium, in particular Jessica Kiefte-de Jong, Wichor Bramer, Ellen Demerath, Weihua Guan, Myriam Fornage, Joyce van Meurs, Sandra Heil, Paul de Vries, Taulant Muka, and Vincent Jaddoe. Thank you for all your help and input in the manuscripts of this thesis.

In the past few years, I have had the pleasure to be surrounded by very "gezellige" roommates. First I would like to thank my current roommates, Marija, Chantal, Jelena, Valentina, Xiaofang, Magda, and Eralda. A special thank you goes out to my two direct neighbors, who are true life-savers! Both for very different reasons, yet equally important. Magda, thank you for the many many times you saved my life when I needed you for technical support while I was in Boston. Eralda, your sense of my afternoon-dip due to cookie deficiency is impeccable. Thank you for always feeding me cookies and chocolate at crucial times! Also to my previous neighbor, Oscar Leonel, *muchas gracias por todas las lecciones de español*. *Eres un gran maestro*! You changed language barriers to language opportunities.

I would also like to thank those from "room" 2901-K, where it all started near the giraffe. Rebecca, dankjewel voor al je hulp gedurende de NIHES master. Ik ben blij dat het ons nog steeds lukt de traditie van de 2901-K-uitjes gaande te houden. Chantal, fijn dat we in hetzelfde schuitje zitten en nu samen kunnen sparren over thesis printing stress. Ik heb altijd opgekeken naar jouw georganiseerde en structurele werkwijze. Waar ik altijd achter de feiten aanloop, heb jij altijd alles ruim voor de deadline klaar staan. Je moet het me echt eens leren! Debora, van proosten op het behalen van onze eerste Master-course in Rotterdam, naar proosten op een PhD plek in Bonn. Wie weet waar ter wereld we kunnen proosten op de volgende stap. Bedankt voor al je steun (onder andere in de vorm van de zeer gewaardeerde kattenfilmpjes) die me vooral in de laatste fase op de been hielden!

My colleagues from the ErasmusAGE, CVD, and other groups at the department of Epidemiology (probably the most international department of EMC): thank you, dankjewel, muchas gracias, faleminderit, and all other languages I still need to learn to say thank you in, to Ester, Myrte, Loes, Lisanne, Audrey, Lisan, Anna, Jolien, Zhangling, Astrid, Mohsen, Carolina, Irma, Eliana, Adela, Silvana, Arjola, Isabel, Michelle, Maryam, Jelena, Elif, Lyda, Oscar (the third), Hoyan, Juna, Ashley, Kate, Nathalie, Genna, Janine, Guannan, Suzanne, Gavro, Tamara, and all whose names I have missed. Thanks for all the great meetings, conference visits, lunches, coffees, and cookie breaks. Anh Nhi, thank you for all your help in this last stage and for bringing me to Schiphol and waiving me goodbye when I went to Boston last time. Hamid, thank you for taking over the cookie breaks. I am glad to see that this responsibility is in hands of someone who takes cookies at least as serious as I do. Sorry for all the forgotten chess-matches, let's pick it up again soon! Mirjana, thank you so much for all your help in the final stage of my thesis. Dina, I was so happy we randomly crossed paths when we were planning our trip for the CHARGE meeting in Houston. It was nice we were able to repeat this in New York and Boston as well. I hope we can explore more cities together this way in the future! Mirjam, thank you for all you valuable help! To my students: Wendy, Rafaelle, Carolin, and Ehab. Thank you for all your hard work, it was great supervising each of you.

A special thank you to Frank Hu for hosting me at the department of Nutrition at the Harvard School of Public Health. Shilpa, thank you for all your help and insightful meetings during my stay. Ambika, thank you very much for helping me get started with the analyses in SAS. If it wasn't for you, I would probably still be trying to work out the errors in my program! Also, thank you to the rest of the group; it was great working with you. My thanks also goes out to all my other colleagues from Boston: Jordi and Martha (whose story is probably the academic equivalent of a fairytale), Jacky, Ane, Sarah,

Manja, Geng, Gang, Shafqat, Hassan, Jiantoa, Yanping, and Reem. And to my group of colleagues who made everyone wonder whether there were any people at Harvard who did not come from Erasmus: Klodian, Josje, Jana, Layal, Sirwan, Daniel, Frank, and Tim, thank you for all the nice lunches, coffee breaks, dinners, and the occasional beer(s). Boston would not have been as much fun without you.

Ook wil ik graag mijn voormalig begeleiders bedanken die mij op weg naar dit promotietraject hebben geholpen. Anke, mijn afstudeerstage bij B-PROOF was mijn introductie in de wetenschap. Bedankt voor je begeleiding gedurende deze periode, dit heeft mij ongetwijfeld een goede start geboden. Edith, dankjewel voor de kans te starten bij ErasmusAGE. Iets wat begon als een master thesis, is uiteindelijk een proefschrift geworden. Bedankt voor je hulp en aanmoediging hier in verder te gaan.

Vincent en Mathilde, bedankt dat jullie mijn paranimfen willen zijn! Vincent, *bruh,* "*Started from the bottom, now we're here*". Opgegroeid in dezelfde wijk, op dezelfde middelbare school gezeten, al kruisten onze paden niet tot een paar jaar geleden. Maar al snel was het duidelijk – vooral na de development of the infamous BraunJen-method in R – wij zijn een top-team! En misschien kunnen Bacon en Dikkie Dik ons team versterken in de toekomst. Math, van samen op de middelbare school, naar samen de NIHES master, en uiteindelijk beide op naar de PhD. Ik vond het echt gezellig toen je ook hier op de afdeling kwam werken, en vind het supertof dat je verder bent gegaan met een promotietraject hebt dat echt bij je past. Ik ben blij dat je naast me zult staan tijdens mijn verdediging!

Mijn lieve vriendinnen, Cait, Joyce, Ils, Jess, Jos, Mees, en Mir, bedankt voor alle nodige afleiding in de vorm van gezellige dinnerdates, stapavondjes en weekendjes weg. Ook gaat mijn dank uit naar mijn hockeyteams, zowel voormalig dames 3 van OMHC als het bedrijfshockeyteam bij Erasmus MC, die altijd voor zeer welkome sportieve afleiding en gezelligheid hebben gezorgd.

Mijn lieve familie, inclusief alle Braunies, Peterse, Bosmans, Wulfjes, en alle "koude kant" die daarbij hoort, waar ik zo ontzettend veel bewondering voor heb. Met wat vast een collectief record moet zijn in tegenslagen, weten jullie toch altijd moed en veerkracht te tonen. Nick, bedankt voor alle schaakpotjes die je toch elke keer weer met me aangaat ondanks dat ik je zelden weet te verslaan. Ik ga mijn best doen in de toekomst een gelijkwaardige tegenstander te worden, of op zijn minst een uitdaging te bieden! Ik ben zo trots op je dat je besloten hebt verder te studeren en voel me vereerd dat je altijd aan mij denkt als je daar een ingewikkelde vraag erover hebt (ookal kan ik ze de helft van de tijd niet beantwoorden). Steph, bedankt voor het o zo vaak zitting nemen als test publiek als ik weer een presentatie moest voorbereiden. Jouw enige taak was dan ook "alleen maar geïnteresseerd kijken", en wat doe je dat altijd goed! Bedankt voor je instant begrip voor alles wat ik doe en zeg. Hiermee lukt het ons altijd, tot grootste ergernis en onbegrip van anderen om ons heen, hele gesprekken te voeren met halve woorden en afgebroken zinnen. Ik ben blij dat (ondanks we niet meer samen wonen) we nog steeds onze *mojo* hebben. Eddy, fijn om je bij de familie te hebben. Bedankt dat je zo'n goede invloed op mijn lieve zus bent – we hebben haar nog nooit zo snel een keuken zien opruimen – en op de rest van ons, volgens mij was onze vakantie in Thailand het meest actief toen we met jullie aan het reizen waren. Mam, bedankt voor alle enthousiaste aanmoediging en vrijheid die je ons hebt gegeven om onze eigen weg te vinden. En de enorme woordenschat die je ons hebt bijgebracht. De meerder-

heid zal niemand ooit doorgaans in een zin gebruiken (behalve jij dan natuurlijk), maar mochten ze die ene keer aan bod komen, dan ben ik op de hoogte!

Aan mijn schoonfamilie, zowel de Henneps als Dubois. In het bijzonder Carmen, Wendel, en – voor altijd in onze herrinnering – Ruben (al zal dit voor mij altijd "meneer" blijven), bedankt dat jullie me altijd thuis hebben laten voelen. Carmen, dankuwel voor alle saoto soep, pom, moksi alesi, nasi, bami, roti, bakabana, pitjel, en noem maar op. Ik kijk elke week weer uit naar uw traditionele zondagse maaltijden, ze geven me altijd de boost die ik nodig heb om weer vol goede moed de week in te gaan. Soso lobi!

Lieve dikkie, ik weet niet hoe ik moet beginnen je te bedanken voor alle steun die je me hebt gegeven. Ik weet niet hoe ik het zonder je had overleefd. Bedankt voor je eindeloze support in alles wat ik doe, je onzelfzuchtigheid hier in, je geduld (die ik soms misschien iets teveel op de proef stel), en dat je me aan het einde van de dag altijd weer aan het lachen weet te krijgen. Je bent awesome ;) Waar heb ik je toch aan verdiend!

En dan sluit ik dit dankwoord af met degene waarbij het allemaal begon. Lieve papa, wat heb ik jou de afgelopen jaren gemist. Ironisch genoeg denk ik dat een van de redenen dat ik jou miste, dezelfde reden is dat het succesvol afronden van mijn proefschrift gelukt is. Zonder al jouw wijze lessen was dit proefschrift nooit tot stand gekomen. Je hebt ons altijd aangemoedigd vragen te stellen, en nog belangrijker, te beantwoorden! Want als we het niet weten, "dan zoeken we het toch op?". Op de strenge regel van tijdens het eten niet van tafel mogen, was er een uitzondering: als de encyclopedie of de Dikke van Dale er bij gehaald moest worden. Ik geloof dat hier de ontwikkeling van mijn wetenschappelijke nieuwsgierigheid is ontstaan. Dit is slechts een klein voorbeeld, hiernaast heb ik nog zo ontzettend veel waarvoor ik je wil bedanken, dat de laatste paragraaf van mijn dankwoord minstenst de lengte van mijn proefschrift zou kunnen krijgen (wat ik geheel terecht zou vinden, maar geen mens die dat leest). Dus laat ik het samenvatten:

Pap, bedankt voor ALLES!

(p.s. dankzij jouw briljante golfbal oplossing ben ik mijn sleutels al 15 jaar niet meer verloren!)

“Netheid is voor de dommen, het genie beheerst de chaos”

– Albert Einstein