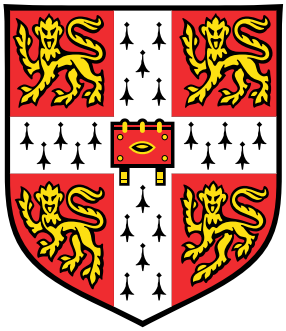
Dairy products and cardio-metabolic health: aspects from nutritional, molecular and genetic epidemiology



Eirini Trichia

MRC Epidemiology Unit

University of Cambridge

This dissertation is submitted for the degree of

Doctor of Philosophy

Wolfson College October 2018

Name: Eirini Trichia

Title: Dairy products and cardio-metabolic health: aspects from nutritional, molecular and genetic epidemiology

Summary

There is accumulating evidence on differences in the link between types of dairy products and cardio-metabolic health, but inconsistent findings limit the field. In my PhD project, I undertook an epidemiological investigation comprising inter-related but distinct themes evaluating aspects of nutritional, molecular and genetic epidemiology to advance scientific understanding.

I undertook research to describe dairy consumption patterns over time by evaluating nationally-representative data of the United Kingdom National Diet and Nutrition Survey. I observed significant time trends for specific dairy types and groups, which were different among different groups of people e.g. adults younger than 65 years or elderly people. Using data from the large Fenland (n~12,000) and EPIC Norfolk (n~25,000) studies, I investigated associations of total and types of dairy consumption with markers of metabolic risk and adiposity as potential pathways to cardio-metabolic disease. The analyses showed differential associations of dairy types and groups mainly with markers of adiposity and lipidaemia. I explored the potential of objective markers to assess dairy consumption, by examining metabolomics profiles and blood fatty acids to identify a set of biomarkers predicting dairy consumption and prospective associations of the identified biomarkers with type 2 diabetes risk. I was able to develop and validate metabolite scores reflecting consumption of some dairy products and observed inverse associations between some of these scores and type 2 diabetes incidence. I analysed genetic determinants of dairy consumption, using a genome-wide association study in the UK Biobank (n~500,000) and identified single nucleotide polymorphisms predicting milk, cheese and total dairy consumption.

Overall, this PhD work contributed towards (1) a more precise description of dairy consumption patterns in the UK, (2) hypothesis formulation for potential biological pathways linking to cardio-metabolic disease, (3) discovery of metabolite scores as potential dairy biomarkers and (4) hypothesis formulation for potential genetic predictors of dairy consumption.

Της παιδείας την μεν ρίζαν είναι πικράν του δε καρπόν γλυκύν...

(The root of education is bitter, but the fruit is sweet…)

Isocrates

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text.

It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text

This dissertation contains less than 60,000 words (excluding appendices, references, tables and figures).

Eirini Trichia
October 2018

Acknowledgements

I undertook this PhD project in the MRC Epidemiology Unit of the University of Cambridge School of Clinical Medicine and it would be impossible to complete without the contribution of some people. This experience has substantially contributed to my personal and professional development and I feel very lucky and grateful for it.

First and foremost, I would like to offer my special thanks to my two supervisors in the MRC Epidemiology Unit, namely Professor Nita Forouhi (primary supervisor) and Dr Fumiaki Imamura (co-supervisor), who have not only been supervisors to me, but real mentors giving me constructive feedback, advice, guidance, help and encouragement to be the best I can be and contributing significantly to my personal and academic development.

I would like to express my great appreciation to Dr Albert Koulman, Dr Felix Day, Dr Birdem Amoutzopoulos and Stephen Sharp for sharing their expert knowledge with me and added the very important multi-disciplinary component to my project that I aimed for combining elements from nutrition, genetics and metabolomics.

Special thanks also to Robert Luben, Connie Tang, Dr Marleen Lentjes and Angela Mulligan from the EPIC Norfolk study, and Inge Loudon and Susie Boatman from the data management team of the Fenland study, who provided the data and helped with all the inquiries I had.

I would also like to thank my colleagues from the MRC Epidemiology Unit including Dr Isobel Stewart, Jian'an Luan and Dr John Perry for their help with my genetic project; my fellow PhD students for creating a stimulating student community with special thanks to Tom White, Sherly Li, Laura Wittemans and Tammy Tong for their great support, help and stimulating discussions; Meriel Smith for all the administrative help and support; nutrition group members for the interesting discussions and useful feedback.

Finally, I am particularly grateful to the Cambridge Trust and the Vergottis family, who funded my PhD project and gave me the opportunity to acquire an education to such a high level.

Contributions

Chapters 2, 4, 5, 6, 7, and 8 include analyses and are a result of collaborative work. I conducted the research described in this thesis, supervised by Professor Nita Forouhi and Dr Fumiaki Imamura from the MRC Epidemiology Unit. Others contributed to this work and their contributions are mentioned below. To indicate the collaborative work, the pronouns "I" and "we" are used interchangeably throughout the document. I did not take part in the design or data collection of the cohort studies used.

Dr Birdem Amoutzopoulos from the Elsie Widdowson Laboratories disaggregated composite foods and recipes in the National Diet and Nutrition Survey (NDNS) and provided me with the dairy and nutrient content of composite foods, and Darren Cole from the Elsie Widdowson Laboratories managed the disaggregation datasets (Chapter 2).

Inge Loudon and Susie Boatman managed and provided the data from the Fenland study (Chapters 4 and 6). Robert Louben and Connie Tang managed and provided the data from the EPIC Norfolk study (Chapters 5, 6 and 7). Dr Marleen Lentjes and Angela Mulligan contributed to the dietary data management of the EPIC Norfolk study (Chapters 5, 6 and 7).

Metabolomics profiling in the Fenland study was perfomed by Dr Albert Koulman, Larissa Richardson, Dr Luke Marney, and Professor Julian Griffin at the Medical Research Council Human Nutrition Research, Cambridge (Chapter 6). Dr Fumiaki Imamura, Dr Albert Koulman and Nicola Kerrison processed and cleaned the raw data. Dr Fumiaki Imamura performed batch correction in collaboration with Stephen Sharp. Metabolomics profiling in the EPIC Norfolk study was performed by Metabolon under the responsibility of Dr Claudia Langenberg (Chapters 6 and 7).

Data from the UK Biobank were managed and provided by Dr Felix Day and Dr John Perry (Chapter 8). Data for the positive control analysis were provided by Jian'an Luan (Chapter 8). The genome-wide association study was performed with the high performance computing resources by Dr Isobel Stewart guided by Dr Felix Day and Dr John Perry (Chapter 8).

Abstract

There is accumulating evidence on differences in the link between types of dairy products and cardio-metabolic health, but inconsistent findings limit the field. In my PhD project, I undertook an epidemiological investigation comprising inter-related but distinct themes evaluating aspects of nutritional, molecular and genetic epidemiology to advance scientific understanding.

I undertook research to describe dairy consumption patterns over time by evaluating nationally-representative data of the United Kingdom National Diet and Nutrition Survey. I observed significant time trends for specific dairy types and groups, which were different among different groups of people e.g. adults younger than 65 years or elderly people. Using data from the large Fenland (n 12,000) and EPIC Norfolk (n 25,000) studies, I investigated associations of total and types of dairy consumption with markers of metabolic risk and adiposity as potential pathways to cardio-metabolic disease. The analyses showed differential associations of dairy types and groups mainly with markers of adiposity and lipidaemia. I explored the potential of objective markers to assess dairy consumption, by examining metabolomics profiles and blood fatty acids to identify a set of biomarkers predicting dairy consumption and prospective associations of the identified biomarkers with type 2 diabetes risk. I was able to develop and validate metabolite scores reflecting consumption of some dairy products and observed inverse associations between some of these scores and type 2 diabetes incidence. I analysed genetic determinants of dairy consumption, using a genome-wide association study in the UK Biobank (n 500,000) and identified single nucleotide polymorphisms predicting milk, cheese and total dairy consumption.

Overall, this PhD work contributed towards (1) a more precise description of dairy consumption patterns in the UK, (2) hypothesis formulation for potential biological pathways linking to cardio-metabolic disease, (3) discovery of metabolite scores as potential dairy biomarkers and (4) hypothesis formulation for potential genetic predictors of dairy consumption.

Publications

In revision at the Journal of Nutrition

Trichia E, Imamura F, Brage S, De Lucia Rolfe E, Griffin SJ, Wareham NJ, Forouhi N G. Associations of types of dairy consumption with cardio-metabolic risk and adiposity: cross-sectional findings from over 12,000 adults in the Fenland Study, UK, Journal of Nutrition (Results from Chapter 4)

Under review by co-authors

Trichia E, Imamura F, Khaw KT, Wareham NJ, Forouhi N G. Prospective associations of total and types of dairy products with markers of metabolic risk and adiposity: findings from the EPIC Norfolk study, UK (Results from Chapter 5)

Under preparation

- Trichia E, Imamura F, Koulman A, Brage S, Griffin SJ, Langenberg C, Khaw KT, Wareham NJ, Forouhi N G. Development and validation of dairy prediction models using metabolomics in two UK cohorts and the associations of derived metabolite scores with type 2 diabetes risk (Results from Chapters 6 and 7)
- Trichia E, Amoutzopoulos B, Imamura F, Forouhi N G. Dairy consumption patterns and their contribution to nutrient intakes: findings from the National Diet and Nutrition Survey 2008-2016 (Results from Chapter 2)

Analyses ongoing

Trichia E, Day FR, Imamura F, Perry J, Wareham NJ, Forouhi N G. Associations of total and types of dairy consumption with type 2 diabetes risk: a Mendelian Randomisation study (Results from Chapter 8)

Conference/meeting presentations

- What can metabolomics tell us about dairy consumption and its associations with type 2 diabetes?, **Oral presentation**, Rank Prize Symposium, September 2018, Grasmere, UK (Results from Chapters 6 and 7)
- Metabolomics as a tool for nutritional biomarker discovery to understand the links between dairy consumption and type 2 diabetes, **Poster presentation**, Strategic Research Review Meeting, September 2018, Institute of Public Health, Cambridge, UK (Results from Chapters 6 and 7)

- Prospective associations of total and types of dairy products with markers of metabolic risk in the EPIC Norfolk study, UK, **Poster presentation**, International Diabetes Federation Congress, December 2017, Abu Dhabi, UAE (Results from Chapter 5)

- Associations of dairy consumption with fat and lean mass distribution in the Fenland study, UK, **Poster presentation**, European Association for the Study of Diabetes (EASD) Annual Meeting, September 2017, Lisbon, Portugal (Results from Chapter 4)

- Dairy products intake, diabetes, obesity and body composition - is there a link?, **Oral presentation**, Wolfson Research Event, February 2016, Cambridge, UK (Results from Chapter 4)

Public engagement

- European Association for the Study of Diabetes (EASD) Annual Meeting 2017 press release “Body fat mass distribution: A possible explanation for lower diabetes risk associated with dairy food consumption” (https://www.eurekalert.org/pub_releases/2017-09/d-bfm091117.php, https://www.sciencedaily.com/releases/2017/09/170913193146.htm; Results from Chapter 4)

- Conversation “Dairy got the all-clear this week - but was it justified?” (https://theconversation.com/dairy-got-the-all-clear-this-week-but-was-it-justified-77423)

- Entry to Max Perutz writing competition: "A milk-way round trip: From genes to foods to disease and back"

List of abbreviations

ALT: Alanine Aminotransferase
ApoA1: Apolipoprotein A1
ApoB: Apolipoprotein B
AST: Aspartate Transaminase
AUC: Area Under the Curve
BMI: Body Mass Index
CLA: Conjugated Linoleic Acid
CHD: Coronary Heart Disease
CRP: C-Reactive Protein
CVD: Cardiovascular Disease
DEXA: Dual Energy X-ray Absorptiometry
DBP: Diastolic Blood Pressure
ESI:Electrospray Ionisation
FAO:Food and Agriculture Organisation
FDR: False Discovery Rate
FFQ: Food Frequency Questionnaire
FIA: Flow Injection Analysis
FSA: Food Standards Agency
GBD: Global Burden of Disease
GGT:Gamma Glutamyl Transferase
GWAS: Genome-Wide Association Study
HbA1c: Haemoglobin A1c
HESI-II: Heated Electrospray Ionisation - II
HILIC: Hydrophilic Interaction Liquid Chromatography
HOMA-IR: Homeostasis Model Assessment for Insulin Resistance
HOMA-%B: Homeostasis Model Assessment for β-cell Function
LASSO: Least Absolute Shrinkage and Selection Operator
LDL-C: Low-Density Lipoprotein Cholesterol
LP: Lactase Persistence
LPC: Lyso-Phospatidylcholine
MRC: Medical Research Council
MS: Mass Spectrometry
MUFA: Monounsaturated Fatty Acids
NAFLD: Non-Alcoholic Fatty Liver Disease
NDNS: National Diet and Nutrition Survey
NEFA: Non-Esterified Fatty Acids
NMR: Nuclear Magnetic Reasonance

OCSFA: Odd-Chain Saturated Fatty Acid
PC: Phosphatidylcholine
PCA: Principal Component Analysis
PUFA: Polyunsaturated Fatty Acids
QC: Quality Control
RCT: Randomised Controlled Trial
ROC: Receiver Operating Characteristic
SBP: Systolic Blood Pressure
SCAT: Subcutaneous Adipose Tissue
SFA: Saturated Fatty Acids
SGD: Stochastic Gradient Descent
SM: Sphingomyelin
SNP: Single Nucleotide Polymorphism
T2D: Type 2 Diabetes
TFA: Trans Fatty Acids
TRIPOD: Transparent Reporting of a multivariable Prediction model for individual Prognosis Or Diagnosis
UPLC: Ultra Performance Liquid Chromatography
VAT: Visceral Adipose Tissue

Table of contents

List of figures

List of tables

Chapter 1

Introduction

1.1 Burden of cardio-metabolic disease

According to the most recent Global Burden of Disease reports of 2016 (GBD 2016), cardiovascular disease (CVD) is still the leading cause of death worldwide, and also in the UK with the ischaemic heart disease in the first place and stroke in the third place (with lung cancer in the second place)[1]. The International Diabetes Federation, in their latest release of the Diabetes Atlas, reported that the number of adults with type 2 diabetes (T2D) worldwide almost tripled from 151 million in 2000 to 425 million in 2017[2]. The prevalence of T2D in the UK was 5.9% in 2017[2]. According to the latest GBD 2016 report on disease risk factors, low dietary quality is the second risk factor globally for disability adjusted life years following child and maternal malnutrition (**Figure 1.1**)[3]. It is also the leading risk factor in countries of middle or middle-high socio-demographic index[3]. High body mass index (BMI) is the seventh top risk factor globally, fifth among countries of middle or middle-high socio-demographic index and fourth among countries of high socio-demographic index (Figure 1.1)[3]. Mortality increased by 11.2% and disability adjusted life years increased by 8.6% from 2006 to 2016 due to suboptimal diet, while increases due to high BMI were 28.6% for both over this decade[3].

The substantial contribution of cardio-metabolic disease to the global mortality and the interplay between nutrition, metabolic risk factors, and cardio-metabolic disease suggest that studying the role of nutrition in relation to disease endpoints, but also intermediate endpoints is of high importance. The World Health Organisation suggests 12 steps to healthy eating for the prevention of non-communicable diseases e.g. CVD, T2D, cancer and obesity (http://www.euro.who.int/en/health-topics/disease-prevention/nutrition/a-healthy-lifestyle). These steps include the consumption of fruit, vegetables, whole grains, replacement of saturated fatty acids (SFAs) with unsaturated fat, replacement of fatty meat with lean meat, fish, beans and legumes, consumption of low-fat dairy products and low consumption of sugar and salt. Dairy products comprise a widely consumed food

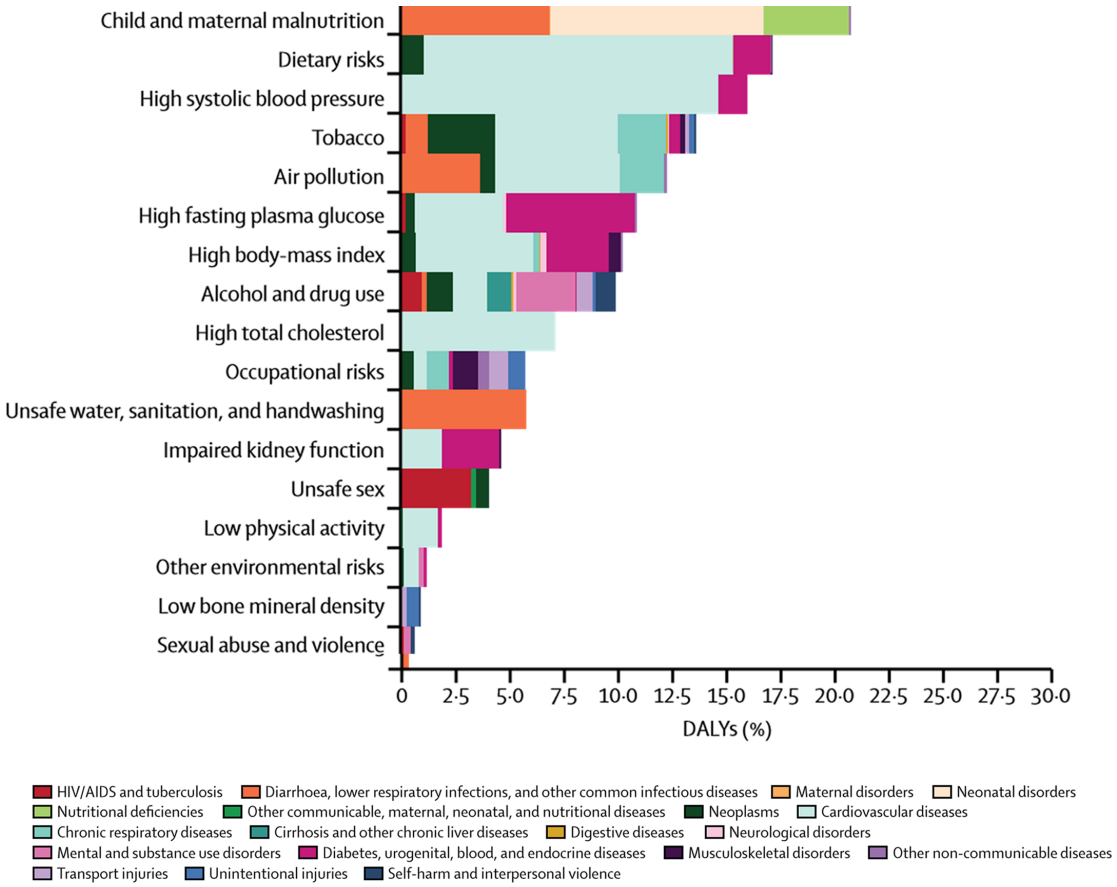


Fig. 1.1 Top risk factors for disability adjusted life years (DALYs) globally, based on the latest report of the Global Burden of Disease (GBD) 2016. Adapted from Gakidou et al[3]

group, which has been linked with bone health and prevention of hip fractures[4], and also cardio-metabolic health, but with substantial heterogeneity in research evidence[5].

1.2 Dairy products and cardio-metabolic disease

After a literature search and extraction of relevant papers from references, I identified at least 17 systematic reviews with meta-analyses of prospective cohort studies evaluating associations of total and types of dairy products with cardio-metabolic disease[6–22] including mortality (n=5 meta-analyses of 2-29 studies)[7, 13, 14, 20, 22], CVD (n=8 meta-analyses of 3-29 studies)[6, 7, 9, 10, 14, 15, 20, 22], coronary heart disease (CHD; n=8 meta-analyses of 3-29 studies)[6, 7, 9, 10, 13–15, 20], stroke (n=10 meta-analyses of 3-18 studies)[6, 9–11, 13–15, 17, 20, 22] and T2D (n=8 meta-analyses of 3-21 studies)[8, 12, 14, 16, 18, 19, 21, 22]. I extracted information on the direction of associations reported from these meta-analyses and summarised it in **Table 1.1**.

Table 1.1 Directions* of associations between total and types of dairy products and cardio-metabolic disease outcomes as reported from meta-analyses of prospective cohort studies

Type of dairy products	Author	Year	Ref.	CVD	CHD	Stroke	T2D	Mortality
Milk	Guo	2017	[7]	↔	↔			↔
	Alexander	2016	[10]	↔	↔	↔		
	de Goede	2016	[11]			↓		
	Gijsbers	2016	[12]				↔	
	Mullie	2016	[13]		↔	↔		↔
	Hu	2014	[17]			↓		
	Aune	2013	[18]				↔	
	Gao	2013	[19]				↔	
	Soedamah-Muthu	2011	[20]	↓	↔	↔		↔
	Elwood	2010	[22]	↓			↓	
Full-fat milk	de Goede	2016	[11]			↑		
	Gijsbers	2016	[12]				↔	
	Aune	2013	[18]				↔	
	Gao	2013	[19]				↔	
	Tong	2011	[21]				↔	
Low-fat milk	de Goede	2016	[11]			↔		
	Gijsbers	2016	[12]				↔	
	Aune	2013	[18]				↓	
	Gao	2013	[19]				↓	
Yoghurt	Guo	2017	[7]	↔	↔			↔
	Wu	2017	[9]	↔	↔	↔		
	Alexander	2016	[10]	↔	↔	↔		
	de Goede	2016	[11]			↔		
	Gijsbers	2016	[12]				↓	
	Qin	2015	[15]		↔	↔		
	Chen	2014	[16]				↓	
	Aune	2013	[18]				↓	
	Gao	2013	[19]				↓	
	Tong	2011	[21]				↓	
Cheese	Chen	2017	[6]	↓	↓	↓		
	Guo	2017	[7]	↓	↔			↔
	Alexander	2016	[10]	↔	↓	↓		
	de Goede	2016	[11]			↔		
	Gijsbers	2016	[12]				↔	
	Qin	2015	[15]		↔	↓		
	Hu	2014	[17]			↓		
	Aune	2013	[18]				↓	
	Gao	2013	[19]				↓	
Butter	de Goede	2016	[11]			↔		
	Pimpin	2016	[14]	↔	↔	↔	↓	↔
	Qin	2015	[15]		↔	↔		
	Hu	2014	[17]			↔		
	Elwood	2010	[22]	↔				
Ice-cream	Gijsbers	2016	[12]				↓	

Table 1.1 (continued)

Type of dairy products	Author	Year	Ref.	CVD	CHD	Stroke	T2D	Mortality
Cream	Gijsbers	2016	[12]				↔	
	Hu	2014	[17]			↔		
Fermented dairy products	Guo	2017	[7]	↓	↔			↔
	de Goede	2016	[11]			↔		
	Gao	2013	[19]				↔	
High-fat dairy products	Guo	2017	[7]	↔	↔			↔
	Alexander	2016	[10]		↔			
	de Goede	2016	[11]			↓		
	Gijsbers	2016	[12]				↔	
	Qin	2015	[15]		↔	↔		
	Chen	2014	[16]				↔	
	Aune	2013	[18]				↔	
	Gao	2013	[19]				↔	
	Soedamah-Muthu	2011	[20]		↔			
	Tong	2011	[21]				↔	
Low-fat dairy products	Guo	2017	[7]	↔	↔			↔
	Alexander	2016	[10]		↓			
	de Goede	2016	[11]			↓		
	Gijsbers	2016	[12]				↓	
	Qin	2015	[15]		↔	↓		
	Chen	2014	[16]				↔	
	Aune	2013	[18]				↓	
	Gao	2013	[19]				↓	
	Soedamah-Muthu	2011	[20]		↔			
	Tong	2011	[21]				↓	
Total dairy products	Guo	2017	[7]	↔	↔			↔
	Schwingshackl	2017	[8]				↓	
	Alexander	2016	[10]	↔	↔	↓		
	de Goede	2016	[11]			↔		
	Gijsbers	2016	[12]				↓	
	Qin	2015	[15]	↓	↔	↓		
	Chen	2014	[16]				↔	
	Hu	2014	[17]			↓		
	Aune	2013	[18]				↓	
	Gao	2013	[19]				↓	
	Soedamah-Muthu	2011	[20]		↔			
	Tong	2011	[21]				↓	
	Elwood	2010	[22]			↓	↓	↓

*↔: no association, ↑: positive association/increase in risk, ↓: inverse association/decrease in risk
Abbreviations: CVD: Cardiovascular disease; CHD: Coronary heart disease; T2D: Type 2 diabetes

Total and types of dairy products were not associated with all-cause mortality[7, 13, 14, 20, 22].

Results on associations with CVD were consistently null for yoghurt[7, 9, 10], fermented dairy products[7], low-fat[7] and high-fat dairy products[7], but mixed for milk[7, 10, 20, 22], cheese[6, 7, 10] and total dairy products[7, 10, 15] indicating inverse or null associations. Specifically, the latest meta-analyses for milk showed null associations with CVD[7, 10] and included five additional studies[7] compared to the older meta-analyses, which showed inverse associations[20, 22]. The latest meta-analyses for cheese and cardiovascular disease showed inverse associations[6, 7] and included six additional studies compared to a former meta-analysis, which showed null associations[10].

Associations between dairy products and risk of CHD were consistently null for milk[7, 10, 13], yoghurt[7, 9, 10, 15], butter[14, 15], fermented[7], high-fat[7, 10, 15, 20]

and total[10, 15, 20] dairy products, while results for cheese[6, 7, 10, 15] and low-fat dairy products[7, 10, 15, 20] showed either inverse or null associations.

No associations were reported between yoghurt[9–11, 15], butter[11, 14, 15, 17], cream[17] or fermented dairy products[11] and risk of stroke, while inverse associations were reported between low-fat dairy consumption and risk of stroke[11, 15]. Either null or inverse associations were reported between milk[10, 11, 13, 17, 20], cheese[6, 10, 11, 15, 17], high-fat[11, 15] or total[10, 11, 15, 17, 22] dairy products and stroke incidence. However, most of the meta-analyses indicated that cheese was associated with a lower stroke incidence by a range of 6% to 13%[6, 10, 15, 17] and total dairy consumption was associated with a lower risk of stroke by a range of 9% to 21%[10, 15, 17, 22].

Finally, null associations were reported between full-fat milk[12, 18, 19, 21], cream[12], fermented[19] or high-fat[12, 16, 18, 19, 21] dairy products and T2D risk and inverse associations were consistently reported between yoghurt (risk reduction range: 9% - 18%)[12, 16, 18, 19, 21] or butter (4% risk reduction per 45 g/day)[14] and T2D risk. Findings were more mixed for milk (low-fat[12, 18, 19] or total milk[12, 18, 19, 22]), cheese[12, 18, 19] or dairy products (low-fat[12, 16, 18, 19, 21] or total dairy products[8, 12, 16, 18, 19, 21, 22]) and T2D risk. Most of the meta-analyses on total milk showed null associations with T2D risk[12, 18, 19]. The meta-analyses also reported an 18% risk reduction from higher consumption of low-fat milk[18, 19] and 8% from higher consumption of cheese[18, 19]. The risk reduction was 4% - 12% per 200 g daily consumption of low-fat dairy products[12, 18, 19, 21] and 3% - 7% per 200 g daily consumption of total dairy products[8, 12, 18, 19, 21, 22].

The diversity of the results among the different types of dairy products and the significant heterogeneity observed in some of the meta-analyses[11–13, 16, 19] raise two key issues: one on what the contributing mechanisms and pathways to disease incidence are; the other on the methodological aspects of relevant studies such as the observational study design and the measurement error of self-reported methods of dietary assessment. Understanding these key issues could further elucidate the link between dairy consumption and cardio-metabolic health and enhance appropriate translation to public health messages about dairy consumption. This is even more relevant for dairy types such as yoghurt or cheese, for which evidence is more limited compared with total dairy products.

1.3 Dietary guidelines

The results of the meta-analyses reviewed above indicate that total and types of dairy products including their high-fat alternatives were associated with either a reduction or no change in the risk of cardio-metabolic disease outcomes. Yet, dietary guidelines continue to recommend the consumption of low-fat dairy products over their high-fat versions[23, 24]. For example, the latest release of the Eatwell Guide by Public Health England in 2016,

recommends the daily consumption of some dairy products, with a preference for low-fat products (**Figure 1.2**)[23]. This apparent discrepancy between the evidence from meta-analyses of prospective studies and the dietary guidelines can be understood in the context of the classical diet-heart hypothesis[25].

According to this hypothesis, SFA food sources, such as dairy products, are thought to be associated with a higher cardiovascular risk, because SFA intake has been associated with higher cholesterol levels[25]. However, this hypothesis has been increasingly controversial. For instance, a meta-analysis of prospective cohort studies failed to report any association between SFA intake and total or cardiovascular mortality, risk of CHD, stroke or T2D[26]. This meta-analysis did not take into account nutrient substitutions, which is an important consideration in the context of an isocaloric diet. Evidence from a meta-analysis of 11 prospective cohort studies[27], a meta-analysis of eight randomised controlled trials (RCTs)[28], and a meta-analysis of 15 RCTs[29], which did account for isocaloric nutrient substitution, showed that substitution of SFAs with polyunsaturated fatty acids (PUFA) leads to lower risk of CHD. Specifically, pooled results from the RCTs showed a 10% CHD risk reduction from the substitution of 5% energy from SFA with 5% energy from PUFA [RR=0.90 (95% CI: 0.83, 0.97)][28]. It is of note that no significant risk change was observed for substitution with monounsaturated fatty acids (MUFA) [27–29], carbohydrates or protein[29].

Overall SFA intake does not seem to be harmful for cardio-metabolic health, when considered alone, but there is a coronary benefit if it is replaced with PUFA. Based on this type of evidence, the World Health Organisation and the Scientific Advisory Committee on Nutrition drafted a public consultation document on evidence and recommendations for SFA intake[30]. This document recommends SFA intake reduction and maintenance to less than 10% energy and replacement of SFAs with PUFA, but the existing evidence was identified as of low or moderate quality[30]. Though the topic of the health effects of SFAs remains contentious, two specific issues are relevant in the context of dairy products: (1) whether the diet-heart hypothesis in case of dairy products is valid and, (2) whether it would be more appropriate to reach conclusions after accounting for the dairy "food matrix" rather than a single nutrient (SFA).

1.4 Dairy food matrix

The food matrix is defined by the United States Department of Agriculture as "the nutrient and non-nutrient compounds of foods and their chemical relationships" (https://definedterm.com/a/definition/197890). Using a more holistic approach by studying the overall dairy food matrix instead of individual nutrients might provide a better insight into the link between dairy consumption and cardio-metabolic health. Dairy products comprise a very

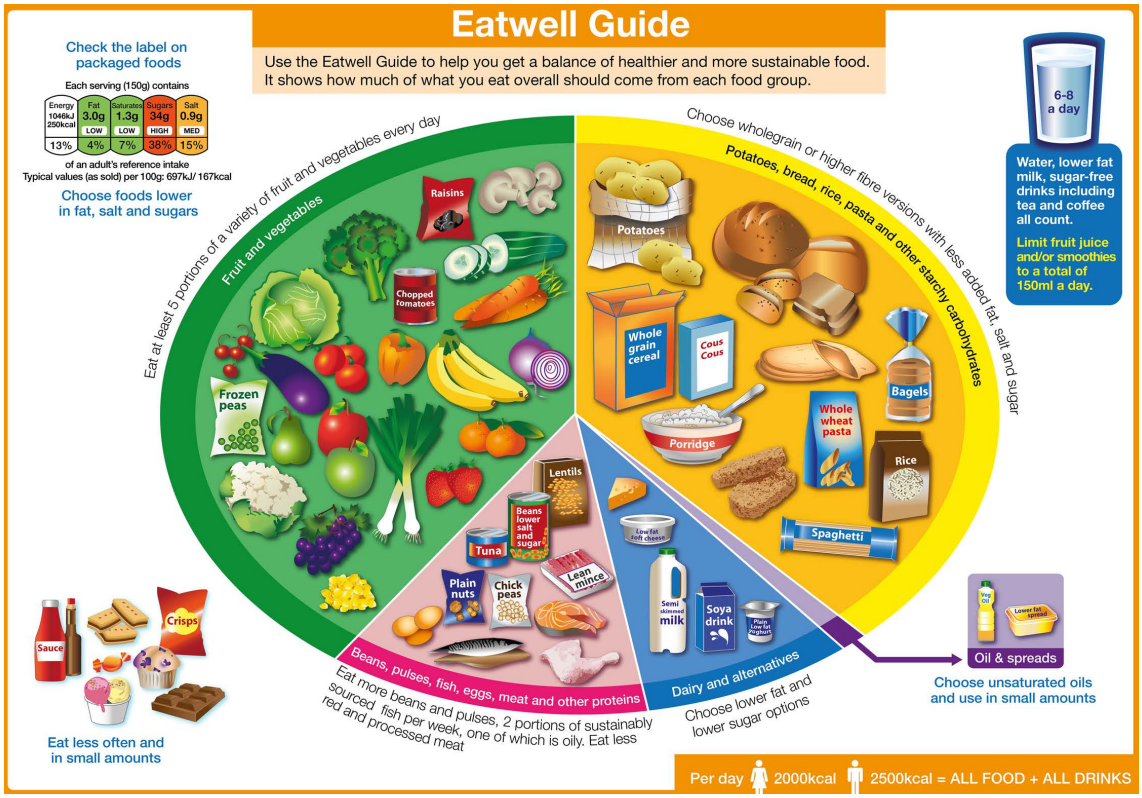


Fig. 1.2 Eatwell Guide 2016, Public Health England, UK[23]

diverse food group with different dairy types varying by nutrient content (e.g. fat), form (e.g. solid, semi-solid, liquid) and processing (e.g. fermentation).

All dairy products are derived from milk, which is defined by Codex Alimentarius as "the normal mammary secretion of milking animals intended for consumption as liquid milk or further processing"[31]. Milk processing can result in products with very different food matrices.

According to Codex Alimentarius, yoghurt is the fermented milk product, which is produced by the starter symbiotic cultures of Streptococcus Thermophilus and Lactobacillus delbrueckii subsp. bulgaricus, while alternate culture yoghurt is produced by cultures of Streptococcus Thermophilus and any Lactobacillus species[31]. This process results in a nutritional profile of yoghurt with less water than milk and thus higher nutrient density with also higher bioavailability. Yoghurt contains lower amounts of lactose, which is also more effectively digested than that from milk due to the presence of lactic acid bacteria[32].

Cheese is a food item with many distinct subtypes. Codex Alimentarius defines general standards for cheese, but also standards for groups of cheese e.g. cheese in brine, unripened cheese including fresh cheese, extra hard grating cheese and specific standards for each type of cheese falling within these groups e.g. mozzarella, cheddar, edam etc[31]. According to the general standards, for a milk product to be considered cheese, the ratio of whey protein to casein should not exceed that of milk and it should be produced through the coagulation of milk protein[31].

Butter is a product with a minimum fat content of 80%, maximum fat content of 90% and a maximum water content of 16%[31].

The heterogeneity of the different dairy types is evident from the pie charts of **Figures 1.3- 1.6**. The water contribution to the total food weight is about 15% for butter and increases to 37% for high-fat cheese, 82-87% for yoghurt and 87-89% for milk.

In terms of fat content, butter has the highest (82%), followed by high-fat cheese (35%), whereas the fat content of the rest of dairy types is below 4%. From the total fat in different dairy types, SFA is consistently the main contributor constituting 62-66% of total dairy fat, followed by 22-24% of MUFA, whereas trans fatty acids (TFA) contribute by just 2-4%. The most abundant TFA is vaccenic acid, but trans fat also consists of conjugated linoleic acid and trans-palmitoleate[33]. For most of the dairy products, contribution of different fatty acids is consistent, with a few exceptions due to different processing. For example, MUFA and TFAs contribute less than 0.1% to total fat of low-fat cottage cheese. The majority of SFAs consists of long-chain fatty acids with 13 carbon atoms or more (62-79%), of which 1.3-1.9% is C15:0 and 0.8-1.1% is C17:0. Medium chain fatty acids (6-12 carbon atoms) constitute 10-16% of dairy SFAs, while short chain fatty acids (up to 5 carbon atoms) constitute 5-6.5%. It should be noted that this is a simplified presentation of the nutrient composition of dairy products, especially for fat, as milk contains over 400 different fatty acids, but most of them are in trace amounts[33]. In addition, fat in dairy products takes several different forms. The majority of it is in the form of triglycerides (98%), but there is also diacylglycerol (<2%), cholesterol (<0.5%), phospholipids (1%) and free fatty acids (0.1%)[33]. Part of triglycerides are surrounded by the milk fat globule membrane, which also contains some lipid classes such as phospholipids[34]. Depending on the process of homogenisation, the milk fat globule membrane in butter might be mostly absent[34].

Concerning protein, cheese has the highest content (25%), followed by yoghurt (5-6%) and milk (3.5%), whereas the contribution to the weight of butter is just 0.6%. Casein constitutes 80% of milk protein and whey constitutes 20%.

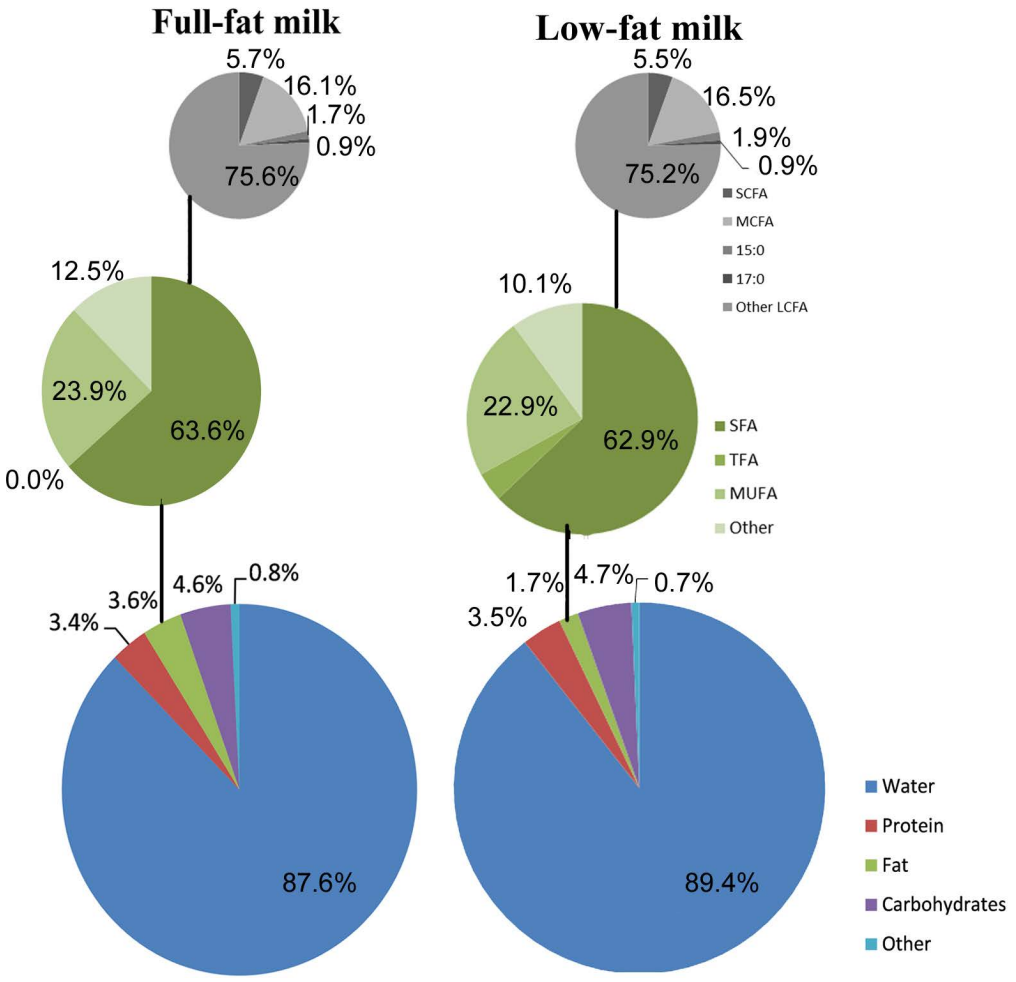


Fig. 1.3 Macronutrient composition of full-fat and low-fat milk. Source: McCance and Widdowson's Food Composition Tables[35]. The food item selected to represent full-fat milk composition was "Milk, whole, pasteurised, average" (food code:12-596) and the item selected for low-fat milk composition was "Milk, semi-skimmed, pasteurised, average" (food code: 12-313) Abbreviations: LCFA: Long-chain fatty acids; MCFA: Medium-chain fatty acids; MUFA: Monounsaturated fatty acids; SCFA: Short-chain fatty acids; SFA: Saturated fatty acids; TFA: Trans fatty acids

Finally, carbohydrates do not contribute more than 8% to the weight of dairy products, when there are no added sugars. The highest contribution of carbohydrates to the total food weight is for yoghurt (7.8%), followed by milk (4.6%) and some types of low-fat cheese (3.3%), while for butter and other types of cheese, it is 0.6% or lower. Carbohydrates in milk, cheese and butter are mainly lactose, whereas in yoghurt 60% of the carbohydrates are lactose and 40% are galactose due to the action of the lactic acid bacteria.

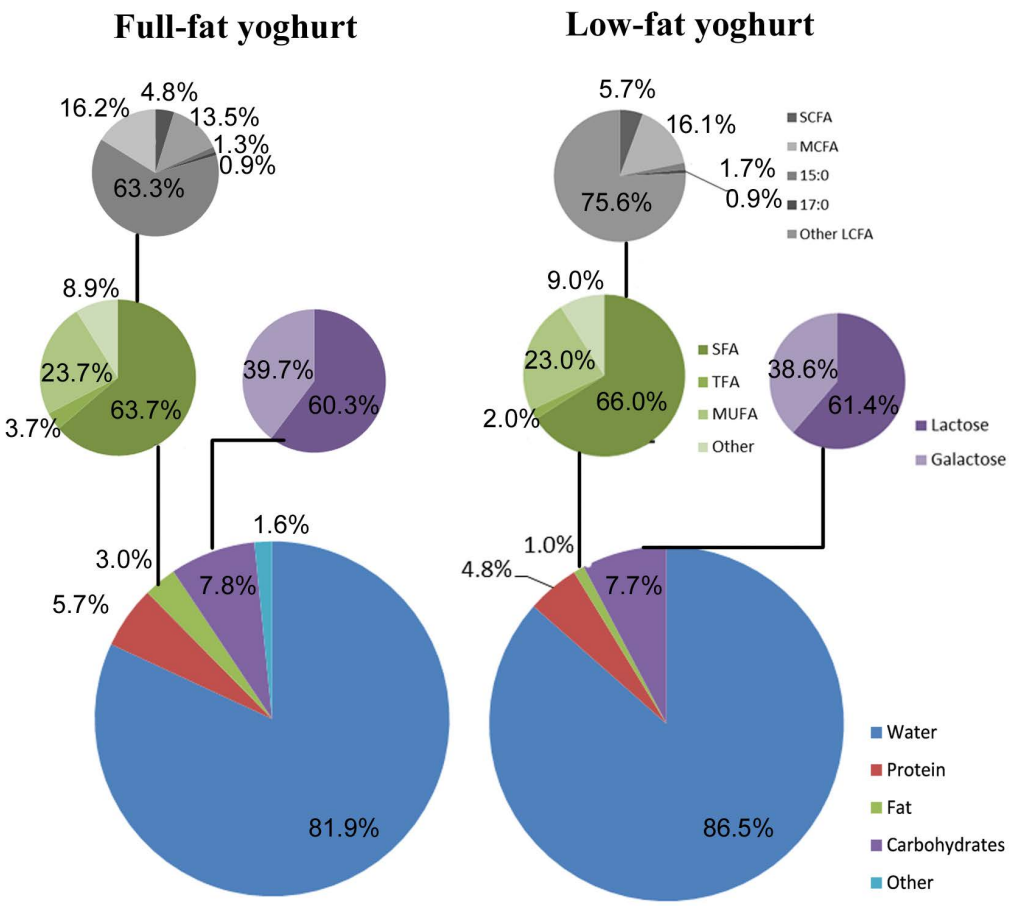


Fig. 1.4 Macronutrient composition of full-fat and low-fat yoghurt. Source: McCance and Widdowson's Food Composition Tables[35]. The food item selected to represent full-fat yoghurt composition was "Yoghurt, whole milk, plain" (food code: 12-184) and the item selected for low-fat yoghurt composition was "Yoghurt, low fat, plain" (food code: 12-379). Abbreviations: LCFA: Long-chain fatty acids; MCFA: Medium-chain fatty acids; MUFA: Monounsaturated fatty acids; SCFA: Short-chain fatty acids; SFA: Saturated fatty acids; TFA: Trans fatty acids

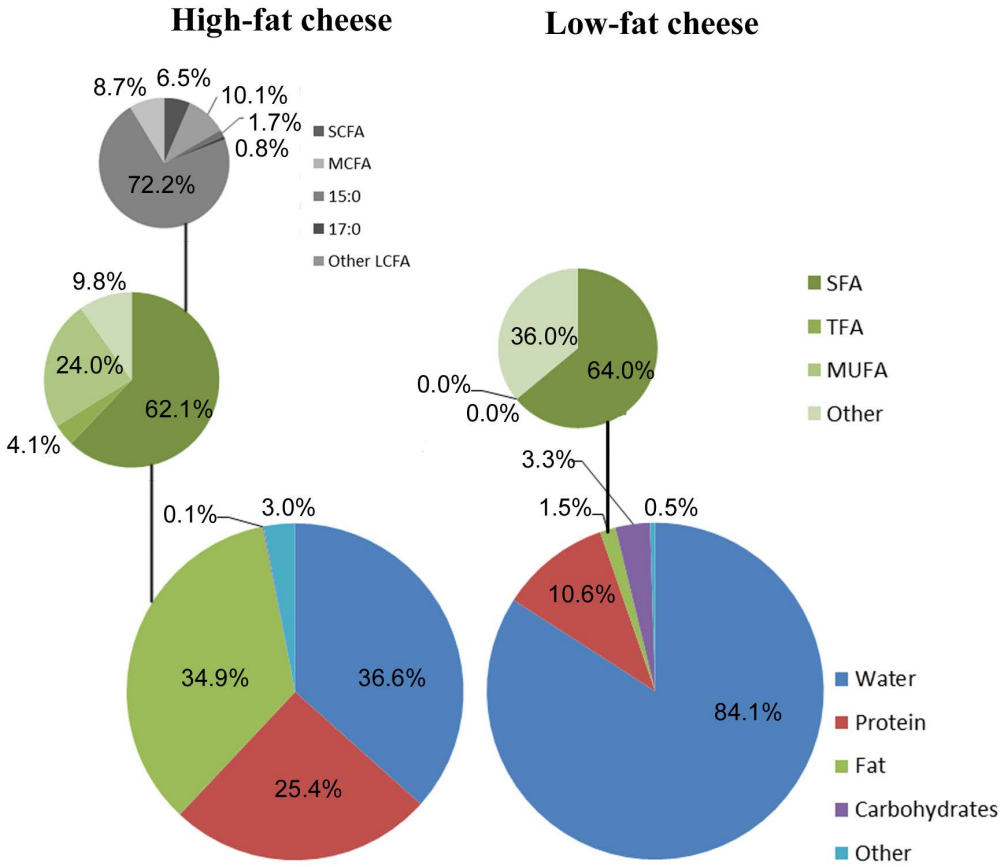


Fig. 1.5 Macronutrient composition of high-fat and low-fat cheese. Source: McCance and Widdowson's Food Composition Tables[35]. The food item selected to represent high-fat cheese composition was "Cheese, Cheddar, English" (food code: 12-346) and the item selected for low-fat cheese composition was "Cheese, cottage, plain, reduced fat" (food code: 12-550). Abbreviations: LCFA: Long-chain fatty acids; MCFA: Medium-chain fatty acids; MUFA: Monounsaturated fatty acids; SCFA: Short-chain fatty acids; SFA: Saturated fatty acids; TFA: Trans fatty acids

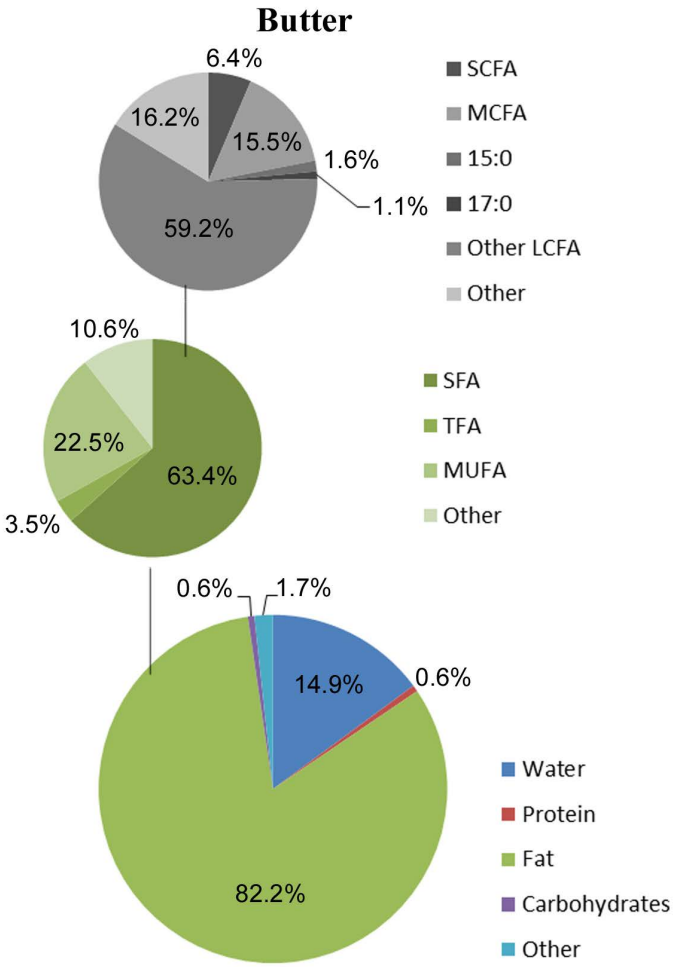


Fig. 1.6 Macronutrient composition of butter. Source: McCance and Widdowson's Food Composition Tables[35]. The food item selected to represent butter composition was "Butter, unsalted" (food code: 17-661). Abbreviations: LCFA: Long-chain fatty acids; MCFA: Medium-chain fatty acids; MUFA: Monounsaturated fatty acids; SCFA: Short-chain fatty acids; SFA: Saturated fatty acids; TFA: Trans fatty acids

Information on the micronutrient content of dairy types is presented in **Table 1.2**. Full-fat and low-fat yoghurt contain higher amounts of minerals compared with full-fat and low-fat milk. Butter has the lowest mineral content, whereas cheese has the highest calcium, magnesium, phosphorus and zinc content of all the dairy types. Regarding vitamin content, butter contains the highest amount of fat-soluble vitamins i.e. vitamin A, vitamin D and vitamin K, followed by cheese. Cheese has the highest content of vitamin B_{12}.

Table 1.2 Content of selected micronutrients in 100 g of dairy food for selected dairy types †

	Minerals (mg)					Vitamins (μg)			
Dairy products ‡	**Potassium**	**Calcium**	**Magnesium**	**Phosphorus**	**Zinc**	**Vitamin A**	**Vitamin D**	**Vitamin B_{12}**	**Vitamin K1**
Full-fat milk	157	120	11	96	0.5	38			0.6
Low-fat milk	156	120	11	94	0.4	20		0.9	
Full-fat yoghurt	280	200	19	170	0.7	32		0.2	
Low-fat yoghurt	228	162	16	143	0.6	8	0.1	0.3	0.03
High-fat cheese	75	739	29	505	4.1	388	0.3	2.4	4.7
Low-fat cheese	161	127	13	171	0.6	17		0.6	
Butter	27	18	2	23	0.1	1060	0.9	0.3	7.4

†Information extracted from the McCance and Widdowson's Food Composition Tables[35]

‡ The food items selected from the Food Composition Table to represent the different dairy types are "Milk, whole, pasteurised, average" (food code: 12-596) for full-fat milk, "Milk, semi-skimmed, pasteurised, average" (food code: 12-313) for low-fat milk, "Yoghurt, whole milk, plain" (food code: 12-184) for full-fat yoghurt, "Yoghurt, low fat, plain" (food code: 12-379) for low-fat yoghurt, "Cheese, Cheddar, English" (food code: 12-346) for high-fat cheese, "Cheese, cottage, plain, reduced fat" (food code: 12-550) for low-fat cheese and "Butter, unsalted" (food code: 17-661) for butter

The complexity of the dairy food matrix, but also its heterogeneity across different dairy types suggest the need for an in depth investigation of the links between dairy consumption and cardio-metabolic health using a multi-disciplinary approach.

1.5 Project aims

The overall aim of this PhD was to develop an understanding of the associations of total and types of dairy products with cardio-metabolic health by incorporating aspects of nutritional, molecular and genetic epidemiology. A mindmap of the specific sub-aims of the PhD and the rationale behind them is presented in **Figure 1.7**.

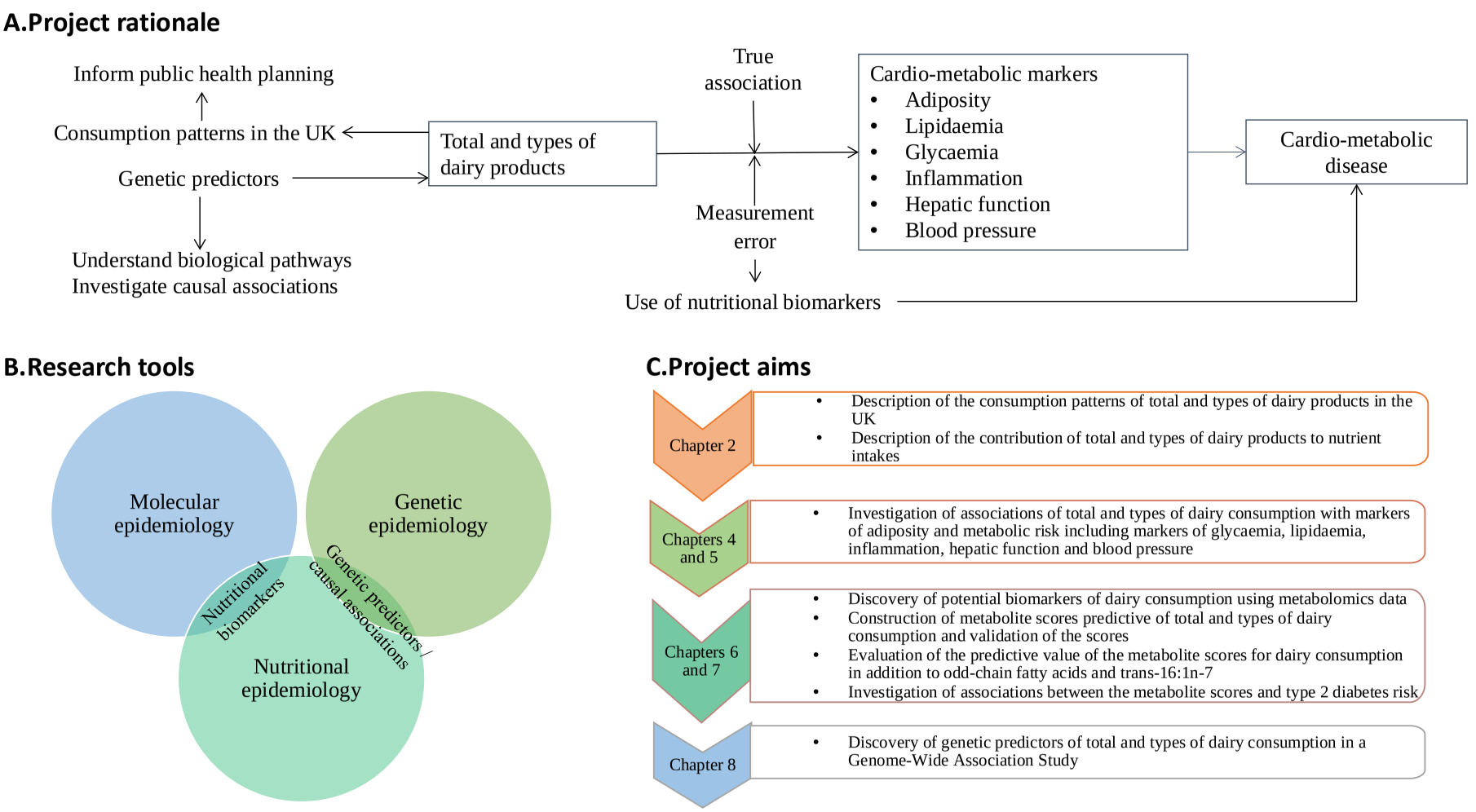


Fig. 1.7 Mindmap of the PhD including the project rationale (A), the areas of epidemiology used as research tools and their overlap (B) and the project sub-aims (C)

Chapter 2

Description of dairy consumption in the UK

Summary

Background and aims: Monitoring of consumption patterns is important to inform public health policies. In the UK, individual level consumption data from a representative UK sample are available only from the National Diet and Nutrition Survey (NDNS). The dairy information available from the NDNS reports includes milk and its fat alternatives, cheese categorised into cheddar, cottage and other cheese, one group for yoghurt, fromage frais and other dairy desserts and butter, but consumption levels do not include content of these dairy types in composite foods. The aims of this study were to describe dairy consumption patterns over time and their contribution to nutrient intakes also accounting for their consumption from composite foods and recipes.

Methods: We evaluated data of adults from the old surveys of NDNS (1994/1995 for elderly, 2000/2001 for adults <65 years) and the rolling programme years (2008/2009 to 2015/2016), which constitute random and representative samples of the UK population. Diet was assessed with weighed 4-day or 7-day (2000/2001) food diaries in the old surveys and estimated 4-day food diaries in the rolling programme. We disaggregated composite foods into food ingredients to derive more precise estimates. Time trends of weighted consumption (g/day) of total, high-fat and low-fat dairy products (milk, yoghurt, cheese, butter) were reported among dairy consumers for adults <65 years and elderly participants $\geq$65 years. Dairy contribution to nutrients across the different survey years was estimated.

Results: A range of 420-597 adults <65 years were included in each rolling programme year, while 1,723 adults were included in year 2000/2001 (mean age range: 40.1-41.3 years; 38.1-45.7% women). For elderly adults, a range of 87-184 participants were included in each rolling programme year, and 1,733 participants were included in the 1994/1995 survey (73.2-76.6 years; 35.7-49.1% women). Total dairy consumption did not significantly change among adults <65 years (242.6-245.0 g/day) or among the elderly

adults (302.8-285.9 g/day) over the period of 20008/2009 to 2015/2016. Low-fat dairy products were consumed consistently in greater amounts (almost 3-fold higher) than high-fat dairy products. Elderly adults had significantly lower consumption of high-fat dairy products over time from 119.1±16.5 g/day in 2008/2009 to 53.1±6.2 g/day in 2015/2016 (p=0.002). The consumption of low-fat dairy products did not significantly change for either age group over time.

Milk was consistently the largest contributor to total dairy consumption for both groups. While consumption of total and low-fat milk did not change over time, consumption of full-fat milk decreased by 8.6 g/day for adults <65 years and by 81.1 g/day for elderly people over the 8-year period. Yoghurt consumption did not significantly change (mean range: 37.8-45.4 g/day for adults <65 years and 39.4-64 g/day for elderly participants), but the percentage of elderly yoghurt consumers increased from 19.5% in 1994/1995 to 46.4% in 2015/2016. No large changes were observed for mean cheese consumption (24.1-27.5 g/day for adults <65 years; 19.8-25.1 g/day for elderly people) or butter consumption (6.4-8.2 g/day for adults <65 years; 7.9-11.2 g/day for elderly people).

Total dairy consumption contributed 13.5-14.6% to total energy intake. For macronutrients, total dairy consumption contributed 24.4-25.8% to total fat, 40.2-42.2% to saturated fat, 16.2-17.7% to cis-monounsaturated fat, 51-62.1% to trans fat, 5.1-5.7% to carbohydrate, and 17.7-19.1% to protein. For micronutrients, total dairy consumption contributed 24.3-32.1% to vitamin A, 11.4-13.3% to vitamin D, 39.4-41.8% to vitamin B_{12}, 48.4-50.5% to calcium, 14.5-15.8% to potassium, 12.3-13.7% to magnesium, 25.7-27.8% to phosphorus and 19.8-21.5% to zinc. Of the subtypes, high-fat cheese was the highest dairy contributor to total and types of fat, and vitamin A, while low-fat milk was the highest contributor to the other nutrients.

Conclusion: In this study of NDNS data, we reported updated consumption levels for total dairy products (low- and high-fat), milk (low- and full-fat), cheese (low- and high-fat) and butter, and their contribution to nutrient intakes further accounting for their consumption in the context of composite foods and recipes. We additionally reported consumption of low- and full-fat yoghurt, which was previously reported as part of the group "yoghurt and dairy desserts". The importance and variability of the dairy food matrix is evident when considering the high contribution of dairy products to intakes of the healthful vitamins and minerals, as well as to intakes of saturated and trans fat. Research into the mechanisms of action of the dairy food matrix on health in combination with close monitoring of consumption levels in the population will inform policy-related decisions.

What is already known

- Monitoring consumption patterns is important to inform public health policies. In the UK, individual level consumption data from a representative sample are available only from the National Diet and Nutrition Survey (NDNS).
- Current NDNS reports include consumption of total dairy products (low- and high-fat), milk (low- and full-fat), cheese (cheddar, cottage, other) and butter. Yoghurt consumption is reported as part of the group "yoghurt, fromage frais and other dairy desserts". These consumption levels do not account for the dairy content in composite foods and recipes.

What this research adds

- After accounting for dairy consumption as part of composite foods, total and low-fat dairy consumption in the UK did not significantly change from 2008 to 2016, while high-fat dairy consumption decreased from 119.1 g/day to 55.1 g/day among elderly people.
- Of the high-fat dairy products, cheese and butter consumption did not change over the 8 years, but full-fat milk consumption decreased by 8.6 g/day among adults <65 years and 81.1 g/day among elderly people.
- Yoghurt consumption did not change, but the percentage of elderly yoghurt consumers increased from 19.5% in 1994/1995 to 46.4% in 2015/2016.
- Total dairy consumption contributed by approximately 15% to total energy, monounsaturated fat and potassium, 20% to protein and zinc, 25% to total fat, vitamin A and phosphorus, 40% to saturated fat and vitamin B_{12}, 50% to calcium and 55% to trans fat intake.

Publication

Trichia E, Amoutzopoulos B, Imamura F, Forouhi N G. Dairy consumption patterns and their contribution to nutrient intakes: findings from the National Diet and Nutrition Survey 2008-2016 (Manuscript under preparation)

From a public health perspective, it is very important to monitor the food intakes of a population in order to identify the main contributors to dietary intake and ensure that the intakes are within appropriate ranges, which have been shown to promote health. This is especially of interest for foods like dairy products, which contain diverse components, with known benefit for bone health, and also related to cardio-metabolic health (sections 1.1 and 1.2). Reporting of trends of dairy consumption over time can inform public health agencies and policymakers on potential deviations from dietary recommendations, so that they can implement appropriate public health interventions. The implementation of the intervention and the choice of the target population will be based on the profile and the characteristics of the consumers and non-consumers in a demographic, socioeconomic and lifestyle context. For example, according to a study with data from a national food and health survey in Australia in 2010-2011, 1.2% of the participants reported avoiding dairy consumption due to a diagnosed disease (e.g.coeliac disease), 15.3% avoided dairy consumption because of unpleasant symptoms without any medical diagnosis and 5.8% avoided dairy consumption without reporting any symptoms[36]. This behaviour was associated with a participant profile characterised by younger age, being a woman, higher worry of illness and higher receptiveness to alternative medicine[36]. This paradigm suggests the need to establish a clear rationale and structure of dietary guidelines and investigate the consumption patterns within a population.

2.1 Data sources of dairy consumption in the UK

The food balance database from the United Nations Food and Agricultural Organisation (FAO; http://www.fao.org/faostat/en/#data/FBS), which includes the per capita food availability, had been the main source of global dietary data until about 10 years ago. Although these data can give rough estimates of food consumption patterns and trends across time and different countries, they have several limitations, as they do not provide individual food consumption data, which are necessary for more precise estimates. To fill these gaps, the Global Dietary Database was launched as part of the data collected for the 2010 Global Burden of Diseases (GBD) project[37]. This database provides individual level dietary data for 11 food groups in 1990, 2005 and 2010 across 193 countries worldwide from publicly available data, data collected for the 2010 GBD project or data collected within the scope of the Global Dietary Database (https://www.globaldietarydatabase.org/).

Concerning dairy consumption in the UK, this database includes consumption levels only for milk as aggregated from 24 data sources. As seen in **Figure 2.1**, according to these data, milk contributes the most to the total amount of food consumed compared with other food groups and it has slightly decreased over the 20-year period.

While this database is useful, it includes information only for milk and not for yoghurt, cheese and butter or total dairy consumption. As a proxy for consumption, there are several

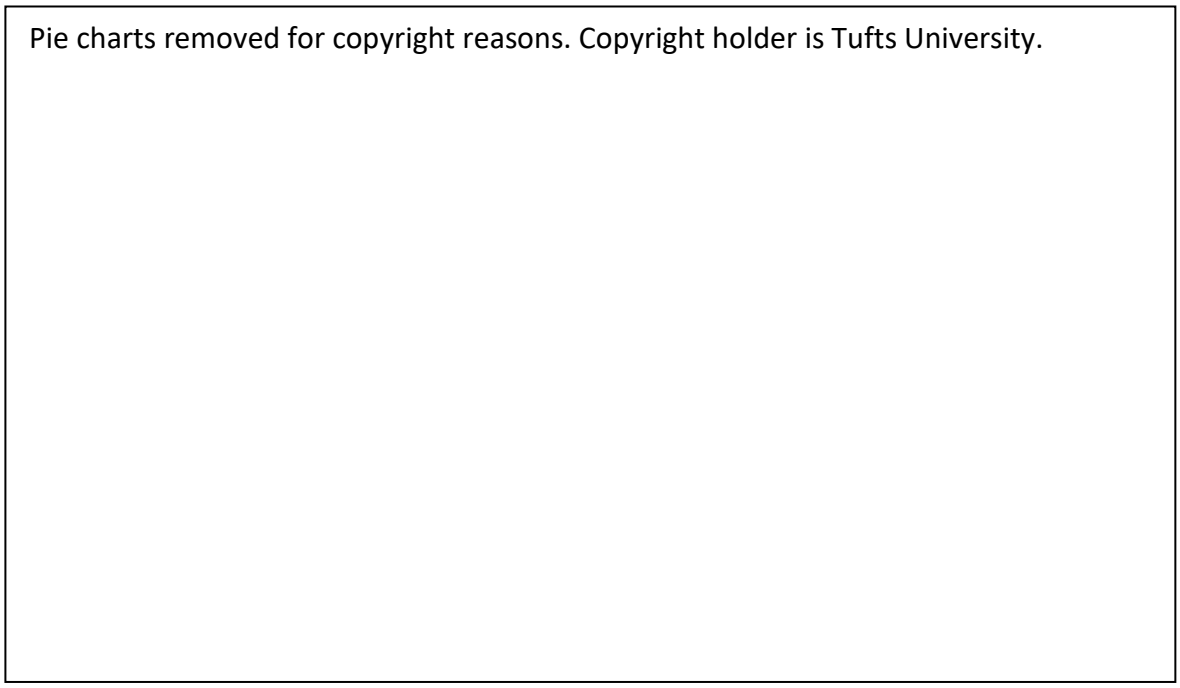


Fig. 2.1 Milk contribution to the total amount of food consumed in the UK in 1990 and 2010, Global Dietary Database (https://www.globaldietarydatabase.org/country-comparisons. html, date of access: 9 June 2018)

data sources related to dairy products in the UK. For example, the Dairy Division of the Agriculture and Horticulture Development Board (AHDB) provides data related to dairy farms, which concern milk yield, supply and flow, dairy products production, dairy trade and household level consumption (https://dairy.ahdb.org.uk/). Furthermore, the Department of Environment, Food and Rural Affairs releases statistics on the production and supply of dairy products, which at least partly overlaps with the data from AHDB (https://www.gov.uk/government/organisations/department-for-environment-food-rural-affairs/about/statistics). A combination of data from these sources and WHO/FAO statistics was used by Hobbs et al. to obtain a picture of the trends of dairy consumption over decades[38]. According to these data, full-fat milk consumption decreased from approximately 140 l/capita/year in 1970 to approximately 20 l/capita/year in 2010. On the contrary, low-fat milk appeared in 1980s and started increasing to reach a plateau of approximately 60 l/capita/year in mid 1990s[38].

For detailed individual level dietary information on consumption levels of any food in a representative UK sample, the National Diet and Nutrition Survey (NDNS) is the most suitable source. However, the dietary data processing has not been done to such an extent yet so as to get the detail that is provided by the 4-day diaries used in the survey. As a result, there are limitations in the survey reports to date. NDNS has reported specifically consumption of milk (whole, semi-skimmed, 1% fat, skimmed), cheese (cheddar, cottage,

other), butter (including ghee and spreadable butter) and yoghurt included in the group "yoghurt, fromage frais and other dairy desserts". Thus yoghurt consumption is not discriminated from consumption of dairy desserts and not categorised into low- and full-fat and cheese is categorised into cheddar and other cheese and not into low- and high-fat. It is of note that the reported consumption levels do not include disaggregated data of composite foods and recipes.

2.2 Profiles of dairy consumers

Several studies have investigated potential correlates of dairy consumption[39–51]. Substantial differences in dairy consumption by ethnicity and country of origin have been reported, which makes of interest the ethnicity and country-specific description and investigation of dairy consumption. Ethnicity is an important factor due to the higher prevalence of lactose intolerance in certain ethnic groups compared to others, as will be elaborated in Chapter 8. Country of origin is also an important factor as total dairy consumption is lower in developing countries, even though it has been increasing at a higher rate[45]. Also among developed countries in Europe, country-specific differences in dairy consumption have been reported, as well as country-specific interactions in associations of socio-demographic factors with dairy consumption[43].

Evidence on specific types of dairy products highlights variation in consumption by several characteristics. For example, men reported higher milk consumption than women in two studies of African American populations[48] and in one study of a UK population[42], while women reported higher low-fat milk consumption than men in a study in Norway[39]. Higher milk consumption has also been associated with lower socioeconomic position in a Finnish study[41], while higher consumption of low-fat milk has been related to a higher[40] or a lower[44] socioeconomic position based on occupation, higher income[40] or higher educational level[49]. Higher consumption of full-fat milk has also been related to lower[40, 46] or higher[44] socioeconomic position based on occupation or income. Higher yoghurt consumption has been consistently reported among women[43, 47, 50], people with higher educational level[47] and income[50] and people who overall adopt healthy lifestyle behaviours including higher compliance with a healthy dietary pattern characterised by higher consumption of fruit, vegetables[51, 52], legumes, nuts[52], lean meat and whole grains and lower alcohol consumption[51], higher physical activity levels[47, 51], no smoking[50, 51] and better sleep quality[51]. Cheese consumption has also been correlated with a higher socioeconomic status based on occupation[40, 41, 46], higher educational level[40, 42, 47] and higher income[40], whereas results for sex are more inconsistent, with some studies having reported higher consumption among men[48] and others among women[39]. Finally, for butter, some studies reported higher consumption among men[43, 44] and there was heterogeneity in

butter consumption and socioeconomic position with both a positive[44] and an inverse association[41] previously reported.

2.3 Aims

So far, individual data on consumption of milk (low- or full-fat), yoghurt (low- or full-fat), cheese (low- or high-fat), and butter and their contribution to nutrient intakes, also accounting for the dairy types consumed as part of composite foods and stratified by demographic factors are not available.

The aims of this study were

1. To describe consumption levels of the main total, low- and high-fat types of dairy products i.e. milk, yoghurt, cheese and butter in a representative UK sample over time and report any trends observed.

2. To describe consumption of total and types of dairy products stratified by age groups and sex in the UK.

3. To describe the contribution of total and types of dairy products to relevant macro- and micronutrient intakes over time in the UK.

2.4 Methods

2.4.1 Study design and population

The UK NDNS was launched in 1992 by the Ministry of Agriculture, Fisheries, and Food and the Department of Health. In 2000, the Food Standards Agency (FSA) and the Department of Health continued the survey with the help of the Social Survey Division of the Office for National Statistics and the Medical Research Council (MRC) Elsie Widdowson Laboratory formely known as MRC Human Nutrition Research. This initiative includes the old surveys started in 1992 with children and continued in 1994/1995 with elderly people (65 years and older) and 2000/2001 with adults and the rolling programme years, which started in 2008 and are repeated every year in participants older than 1.5 years[53]. The NDNS rolling programme is now co-funded by Public Health England of the Department of Health and FSA and conducted by NatCen Social Research, Elsie Widdowson Laboratory and University College London Medical School (only for years 1-4). From 2019, the MRC Epidemiology Unit will be responsible for the survey.

A pilot study was conducted in 1994 for the survey of the year 1994/1995, in 1999 for the survey of 2000/20001 and in 2007 for the rolling programme. The survey was approved by the Multi-centre Research Ethics Committee (years 1-4 of the rolling programme), the Cambridge South NRES Committee (years 5-8 of the rolling programme) and the National

Health Service Local Research Ethics Committees of the areas included in the sample (all years). All participants or their proxies provided written informed consent. A timeline of the survey is shown in **Figure 2.2**.

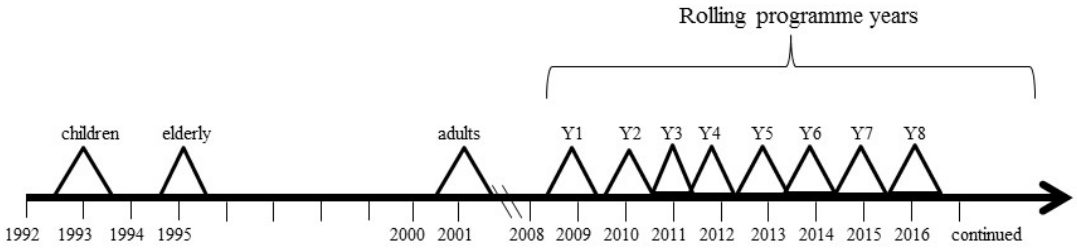


Fig. 2.2 National Diet and Nutrition Survey timeline, || denotes a time gap of more than one year

Households were selected with multi-stage random probability sampling. The first stage was the primary sampling units derived from the postal sectors and the second stage was the Government Office Region. Invitations were sent to the households selected randomly from the Postcode Address File within primary sampling units and interviewers also approached the invited households that did not respond. Eligible individuals were identified from each household who accepted the invitation and only one adult was selected per household or one adult and one child in the rolling programme, also accounting for the target age and sex groups, which were needed to complete a representative sample as defined from census data. Pregnant or breastfeeding women were excluded. If a postal code belonged to an institution, it was excluded from the study, but for the years 1994/1995, a small sub-sample of the elderly sample was recruited from institutions too, which were randomly selected.

The target sample size of the rolling programme was approximately 1,000 participants (500 children and 500 adults) each year. The response rates were 85% in 1994/1995, 61% (47% for the 7-day diary) in 2000/20001, 64% for years 1-4, 63% for years 5-6 and 60% for years 7-8 of the rolling programme. Weights were generated based on sampling probabilities of region, age and sex groups to compensate for any divergence from sample representativeness due to the selection of households and people within household and non-response. Participants were interviewed over four periods of three months across the whole year, to account for seasonality.

2.4.2 Dietary assessment

Diet was assessed in the NDNS with multiple-day food diaries. In 1994/1995, participants used a 4-day weighed diary or a 4-day estimated diary if weighing was a reason for dropout from the study. Four-day weighed diaries showed no substantial differences in intakes compared to 7-day weighed diaries, but had higher participant compliance in the 1994

pilot study. In 2000/20001, participants completed 7-day weighed food diaries. In case of eating out, participants were asked to keep a descriptive diary and then interviewers would purchase the foods described in the reported amounts and weigh them. Institutionalised people in 1994/1995 were further assisted by the interviewers who visited them once daily to weigh one meal (different meal each day) and the rest of the meals were reported in a descriptive diary. Relevant information was also obtained from the care providers.

In the rolling programme, participants completed 4-day estimated food diaries, based on better performance of the 4-day estimated diary compared to a 24-hour recall on four non-consecutive days in the pilot study. Participants were instructed to complete the diary during two weekdays and two weekend days in the year 1 of the rolling programme. The instructions changed for the rest of the years, so that all days of the week are equally represented. In order to generate estimates that are as comparable as possible, the data from the year 2000/2001 were adjusted to include four out of the seven days of reporting, because although the mean intakes should not substantially change, the variation and the percentage of consumers of certain foods are expected to be different and thus not comparable. The four days were selected so that each day of the week is included equally in the derived sample, the days are consecutive and where possible the first four days are selected from the seven days reported. Based on these criteria the number of days for each day of the week re-allocated is shown in **Table 2.1**.

Table 2.1 Re-allocation of the start day of diary reporting for a subset of four days out of the seven completed by adults in the National Diet and Nutrition Survey (NDNS) 2000/01. Table adapted from NDNS report.

Original start day	Re-allocated start day							
	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Total
Sunday	56							56
Monday		121						121
Tuesday			246		66	67		379
Wednesday				246	23	26	71	366
Thursday	190				158			348
Friday		125				154		279
Saturday							175	175
Total	246	246	246	246	247	247	246	1724

Composite foods were disaggregated into food ingredients, which are contained in the FSA standard recipes database[54]. Nutrient intakes were estimated using year-specific Nutrient Databanks, which were based on McCance and Widdowson's Food Composition Tables[35], but were updated from the survey data each year adding information from new food analyses, food manufacturers, food labels and homemade recipes. In the rolling programme, the DINO (Diet-In-Nutrients-Out) system and a Microsoft Access analysis

system were used for the dietary data processing. Nutrient intakes were corrected for the water loss during cooking and/or processing.

2.4.3 Assessment of other factors

Socio-demographic factors (age, sex, ethnicity, occupation) were assessed with an interviewer-administered questionnaire. BMI was calculated by dividing the weight (Soehnle Quantratronic digital scale 7300, 7306 in the old surveys; Soehnle, Seca 850, Seca 870, Tanita THD-305 in the rolling programme) by the square of height (kg/m^2).

2.4.4 Statistical analysis

We reported consumption of milk (low- and full-fat), yoghurt (low- and full-fat), cheese (low- and high-fat), butter, cream, fermented dairy products (sum of yoghurt and cheese), low-fat dairy products (sum of low-fat milk, total yoghurt, low-fat cheese and low-fat cream), high-fat dairy products (sum of full-fat milk, high-fat cheese, high-fat cream and butter) and total dairy products.

Day-level data were aggregated using weights, which were calculated so that weekend days are given a lower weight than weekdays depending on the total number of diary days completed. So if someone completed one day, they were given the weight 1*5/7=0.71 if it was a weekday and the weight 1*2/7=0.29 if it was a weekend day. If they completed two days, they were given 2*5/7=1.43 for weekdays or 2*2/7=0.57 for weekends. For three completed days, they were given a weight of 3*5/7=2.14 for weekdays or 3*2/7=0.86 for weekend days and finally for 4 completed days, they were given a weight of 4*5/7=2.86 for weekdays or 4*2/7=1.14 for weekends.

Due to the high number of non-consumers for some dairy types, statistics were estimated within consumers and the percentage of consumers was reported. Due to the skewness of the consumption of total and types of dairy products, the median and inter-quartile range in addition to the mean and its standard error were estimated. Dairy consumption was reported in g/day.

Consumption levels are presented stratified by age with 65 years as the cut-off point for all the survey years, because the old surveys assessed adults <65 years and elderly people with age $\geq$65 years in different years. In addition, combined consumption levels for all the years stratified by sex are presented and combined consumption levels for the rolling programme years only stratified by BMI are presented.

Statistics were calculated after the data were set as survey data using the sampling weights and accounting for the two-stage sampling design. Significance of time trends was derived from linear regression models including dairy consumption within consumers and the survey year and logistic regression models including dairy consumption (yes or no) and the survey year. Due to the different methods used for the dietary assessment in

the old surveys (weighed food diaries) and the rolling programme years (estimated food diaries), significance of time trends is reported only for the rolling programme years for the consumption within consumers, but for all the years for the percentage of consumers, which is expected to be independent of the dietary assessment method. Years are presented throughout the results as survey years e.g. 2008/2009, 2009/2010 and not as chronological years e.g. 2008, 2009, because each survey year covered all seasons within two different chronological years.

The percent contribution of the different types of dairy products for each year of the rolling programme was estimated for nutrients including: intake of total energy, total fat, saturated fat, cis-monounsaturated fat, trans fat, carbohydrates, total sugars, protein, vitamin A, vitamin D, vitamin B_{12}, calcium, potassium, magnesium, phosphorus and zinc.

All analyses were conducted using Stata 14.2 (College Station, TX: StataCorp LP, 2015).

2.5 Results

2.5.1 Basic characteristics

Basic characteristics of adults and elderly participants for each survey year are presented in **Table 2.2**. By design, the old survey years included more participants than the rolling programme years. The rolling programme years included 420-597 adults, whereas year 2000/2001 included 1,723 adults <65 years. For elderly participants, the rolling programme years included 87-184 participants and year 1994/1995 included 1,733 participants.

The mean age was consistent across years ranging from 40.1 years in 2013/2014 to 41.3 years in 2011/2012 for adults, and from 73.2 years in 2012/2013 to 76.6 years in 1994/1995 for the elderly. Likewise, the percentage of women ranged from 38.1% in 2013/2014 to 45.7% in 2010/2011 for adults, and from 35.7% in 2015/2016 to 49.1% in 1994/1995 for elderly participants.

The percentage of people in the high BMI category (>30 kg/m^2) increased overall over time starting from 21.1% in 2000/2001 and reaching 24.5% in 2015/2016 after peaking at 30.3% in 2011/2012 for adults. Among the elderly group, the prevalence of high BMI ($\geq$30 kg/m^2) increased from 14% in 1994/1995 to 24.5% in 2015/2016 after peaking at 32% in 2013/2014.

Most of the participants completed all the four days of the food diaries with percentages ranging from 81.7% in 2000/2001 to 99.2% in 2008/2009 for adults, and from 96% in 2015/2016 to 98.4% in 2008/2009 for elderly people.

Table 2.2 Basic characteristics of participants <65 years and ≥65 years in each survey year of the National Diet and Nutrition Survey (NDNS)

	Survey year	Old survey †		2008/2009		2009/2010		2010/2011		2011/2012		2012/2013		2013/2014		2014/2015		2015/2016	
<65 years	N	1,723		488		420		444		512		535		546		556		597	
Age (years)		40.9	0.4	40.4	0.7	40.6	0.8	40.2	0.8	41.3	0.7	41.1	0.9	40.1	0.9	40.3	0.9	40.9	0.9
Sex (ref. Men)	Women	44.4		42		45		45.7		40		40.9		38.1		44.8		40	
BMI (kg/m^2; ref. < 25)	25-30	35.8		32.4		32.1		29.5		30.3		33.8		33.9		32.9		33.2	
	>30	21.1		24.4		24		24.1		30.3		21.3		25.3		25		24.5	
Number of diary days completed (ref.1)	2	7.3		0		0		0		0		0		0		0		0	
	3	11		0.8		2.4		2.3		2.7		1.9		1.1		1.8		2.5	
	4	81.7		99.2		97.6		97.7		97.3		98.1		98.9		98.2		97.5	
≥65 years	N	1,733		124		116		87		131		160		178		184		151	
Age (years)		76.6	0.2	74.3	0.7	74.1	0.6	75.1	0.7	73.3	0.8	73.2	0.5	73.5	0.6	74.3	0.6	74.4	0.7
Sex (ref. Men)	Women	49.1		41.9		41.4		43.7		48.1		40.6		42.1		47.3		35.8	
BMI (kg/m^2; ref. < 25)	25-30	32.5		34.7		40.5		37.9		38.9		35		34.8		35.3		32.5	
	>30	14		28.2		25		23		20.6		24.4		32		24.5		24.5	
Number of diary days completed (ref.1)	2	0.6		0		0		0		0		0		0		0		0	
	3	0.9		1.6		2.6		2.3		3.1		2.5		2.8		2.7		4	
	4	97.3		98.4		97.4		97.7		96.9		97.5		97.2		97.3		96	

†Old survey is year 2000/2001 for the sub-group of <65 years and 1994/1995 for the sub-group of ≥65 years

2.5.2 Trends in dairy consumption

UK time trends of the percentage of consumers and average consumption levels within consumers for total and types of dairy products are presented from 2000/2001 to 2015/2016 for adults younger than 65 years, and from 1994/1995 to 2015/2016 for elderly adults older than 65 years in **Figures 2.3- 2.12**. The corresponding statistics (mean, standard error, median, interquartile range) are presented in **Tables A.1- A.2**.

Total, low- and high-fat dairy consumption

The percentage of adult consumers or elderly consumers of total and high-fat dairy products did not significantly change from 2000/2001 or 1994/1995 respectively to 2015/2016 with levels higher than 99.8% throughout the whole period (p>0.05; Figures 2.3, 2.4). Low-fat dairy consumers <65 years significantly decreased from 99.3% in 2000/2001 to 98.2% in 2015/2016 (p=0.04; Figure 2.3), but the change for elderly participants was not significant (Figure 2.4). A decrease was observed in consumption of total and high-fat dairy products from 2000/2001 (287.8±5.3 and 85.1±3.2 g/day respectively) to 2008/2009 (242.6±9 and 73.2±5.9 g/day respectively), but no significant change was observed from 2008/2009 to 2015/2016 (p>0.05). Low-fat dairy consumption was consistently higher compared to high-fat dairy consumption with 205.3±5.2 g/day in 2000/2001 to 181.8±9.9 g/day in 2015/2016 for low-fat dairy products and 85.1±3.2 g/day in 2000/2001 to 66.7±4.9 g/day in 2015/2016 for high-fat dairy products (Figure 2.3). For elderly consumers, although total and low-fat dairy consumption did not significantly change over the 8-year rolling programme period, high-fat dairy consumption significantly decreased from 167.6±5.5 g/day in 1994/1995 and 119.1±16.5 g/day in 2008/2009 to 53.1±6.2 g/day in 2015/2016 (p=0.002; Figure 2.4).

Milk consumption

Milk was consistently the largest contributor to total dairy consumption across the years, even though the percentage of total and low-fat milk consumers significantly decreased among adults (p=0.01 and p=0.0004 respectively), maintaining levels above 98.5% for total milk and 97% for low-fat milk (Figure 2.5). On the contrary, the percentage of full-fat milk consumers did not change significantly (p>0.05), but fluctuated within a range of 72 to 80% (Figure 2.5). Within adult consumers, average total and low-fat milk consumption did not significantly change (p>0.05; Figure 2.5), but full-fat milk consumption significantly decreased from 68.8±3.8 g/day in 2000/2001 and 55.3±7.2 g/day in 2008/2009 to 46.7±5.8 g/day in 2015/2016 (p=0.01; Figure 2.5). Among elderly people, the percentage of full-fat milk consumers significantly decreased from 87.9% in 1994/1995 to 73.5% in 2015/2016 (p<0.0001), while no substantial trends were observed for percentages of total or low-fat milk consumers (Figure 2.6). Within elderly consumers, consumption of total and low-fat milk did not significantly change over the 8-year period (p>0.05), but full-fat milk consumption significantly decreased from 165.6±6 g/day in

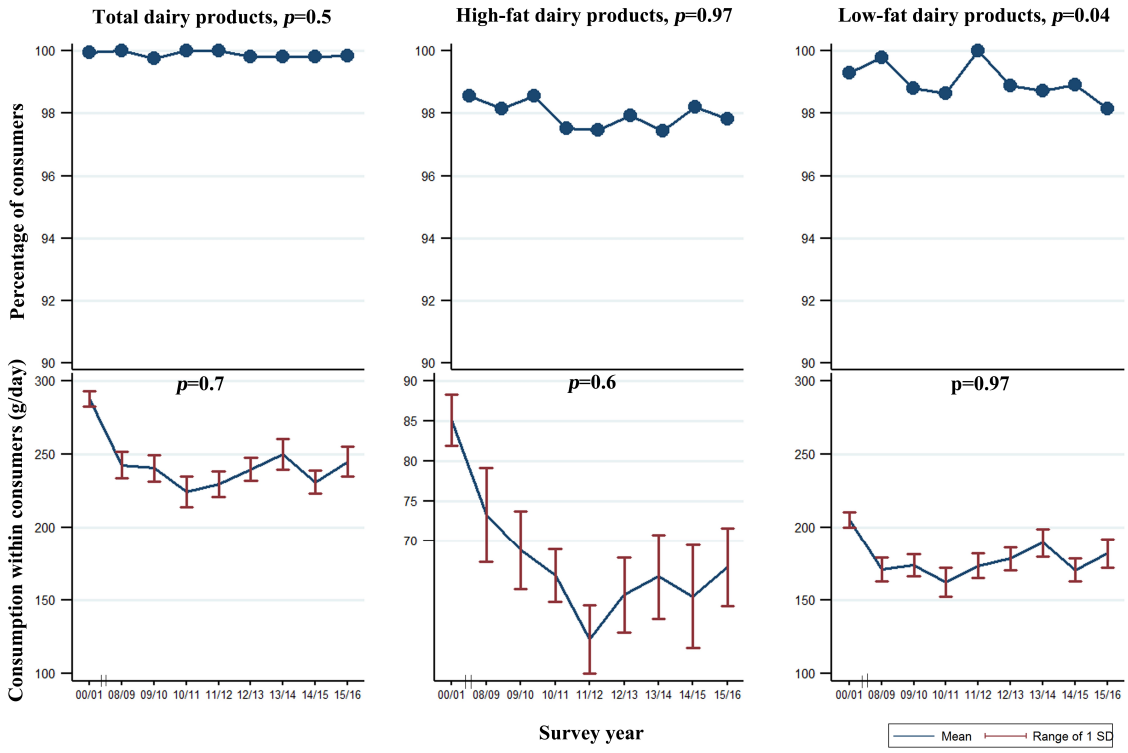


Fig. 2.3 Time trends of the percentage of consumers and the consumption levels within consumers of **total, high- and low-fat dairy products** among **adult participants <65 years** in the old survey (2000/2001) and the rolling programme years (2008-2016) of the National Diet and Nutrition Survey (NDNS) in the UK. *P* for trend is estimated for years 2008/2009-2015/2016 only for consumption within consumers, but for all the years for the percentage of consumers. ‖ denotes a time gap of more than one year between the old surveys and the rolling programme.

1994/1995 and 114.2±19.4 g/day in 2008/2009 to 33.1±7.8 g/day in 2015/2016 (p=0.02; Figure 2.6).

Yoghurt consumption

Yoghurt consumption patterns did not significantly change for adults <65 years (Figure 2.7). In the most recent survey year, 2015/2016, 48.1% of adults <65 years consumed yoghurt (average consumption: 45.4±3.7 g/day). While consumption within consumers did not change also for participants ≥65 years, the percentage of consumers of total, full-fat and low-fat yoghurt significantly increased from 19.5% to 46.4%, 13.8% to 29.8% and 13.5% to 34.4% respectively over the 12-year period (Figure 2.8).

Cheese consumption

The percentage of adult total and low-fat cheese consumers did not significantly change, whereas the percentage of high-fat cheese consumers increased from 78.4% in 2000/2001 to 80.9% in 2015/2016 (Figure 2.9). Within adult consumers, total cheese

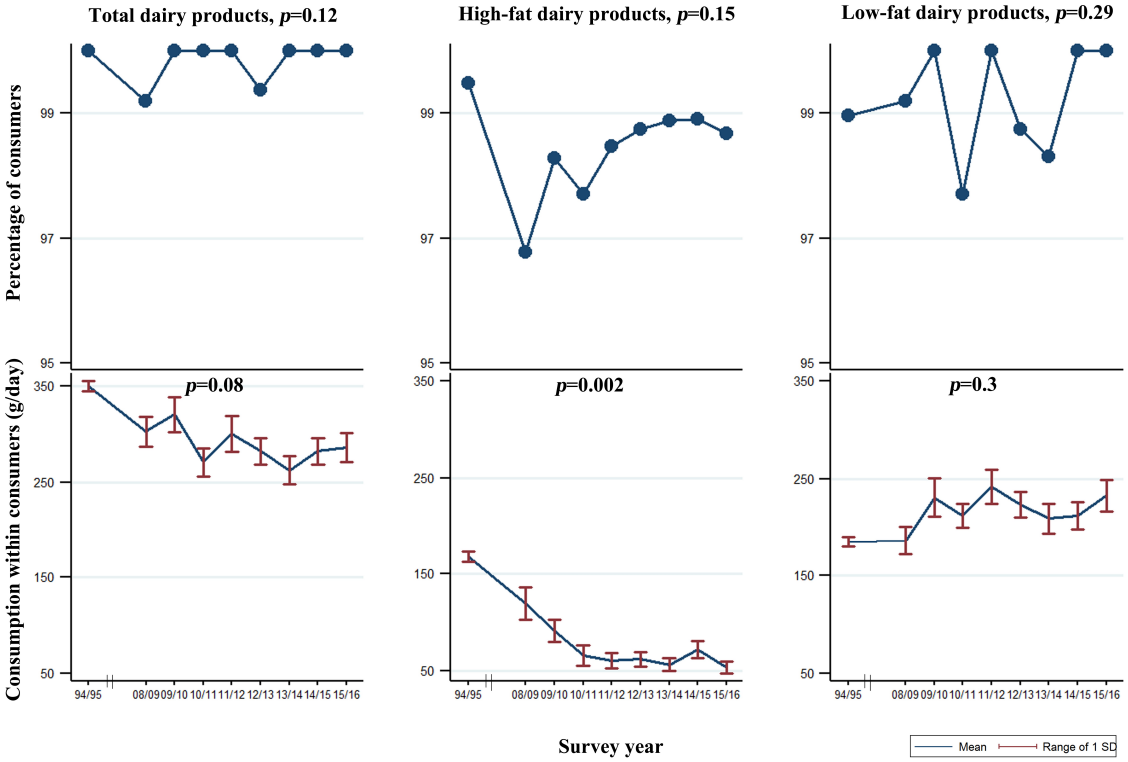


Fig. 2.4 Time trends of the percentage of consumers and the consumption levels within consumers of **total, high- and low-fat dairy products** among **elderly participants ≥65 years** in the old survey (1994/1995) and the rolling programme years (2008-2016) of the National Diet and Nutrition Survey (NDNS) in the UK. *P* for trend is estimated for years 2008/2009-2015/2016 only for consumption within consumers, but for all the years for the percentage of consumers. ‖ denotes a time gap of more than one year between the old surveys and the rolling programme.

consumption significantly slightly decreased from 27.5±1.5 in 2008/2009 to 24.1±1.4 g/day in 2015/2016, but no significant change was observed for high- or low-fat cheese consumption separately (Figure 2.9). The percentage of elderly consumers of total cheese and high-fat cheese significantly increased from 71.3% to 76.8% and 69.6% to 75.5% respectively from 1994/1995 to 2015/2016, but the percentage of consumers of low-fat cheese did not significantly change (Figure 2.10). No significant 8-year trend was observed for cheese consumption within elderly people ≥65 years (Figure 2.10).

Butter consumption

Consumption patterns for butter both among adults <65 years and elderly people ≥65 years did not significantly change over the 8-year period. No significant changes were observed within consumers. In the most recent survey year of 2015/2016, butter consumption was 8.1±0.5 g/day and 7.9±0.8 for adults <65 years and elderly people respectively (Figures 2.11, 2.12).

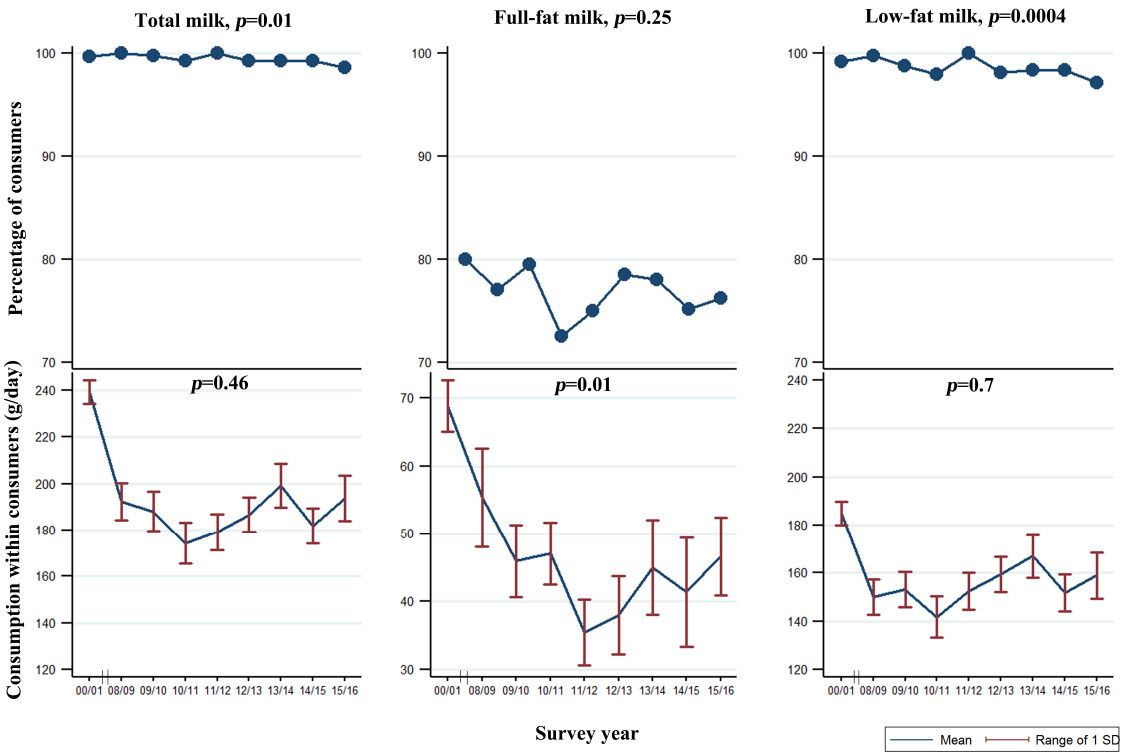


Fig. 2.5 Time trends of the percentage of consumers and the consumption levels within consumers of **total, full- and low-fat milk** among **adult participants <65 years** in the old survey (2000/2001) and the rolling programme years (2008-2016) of the National Diet and Nutrition Survey (NDNS) in the UK. *P* for trend is estimated for years 2008/2009-2015/2016 only for consumption within consumers, but for all the years for the percentage of consumers. || denotes a time gap of more than one year between the old surveys and the rolling programme.

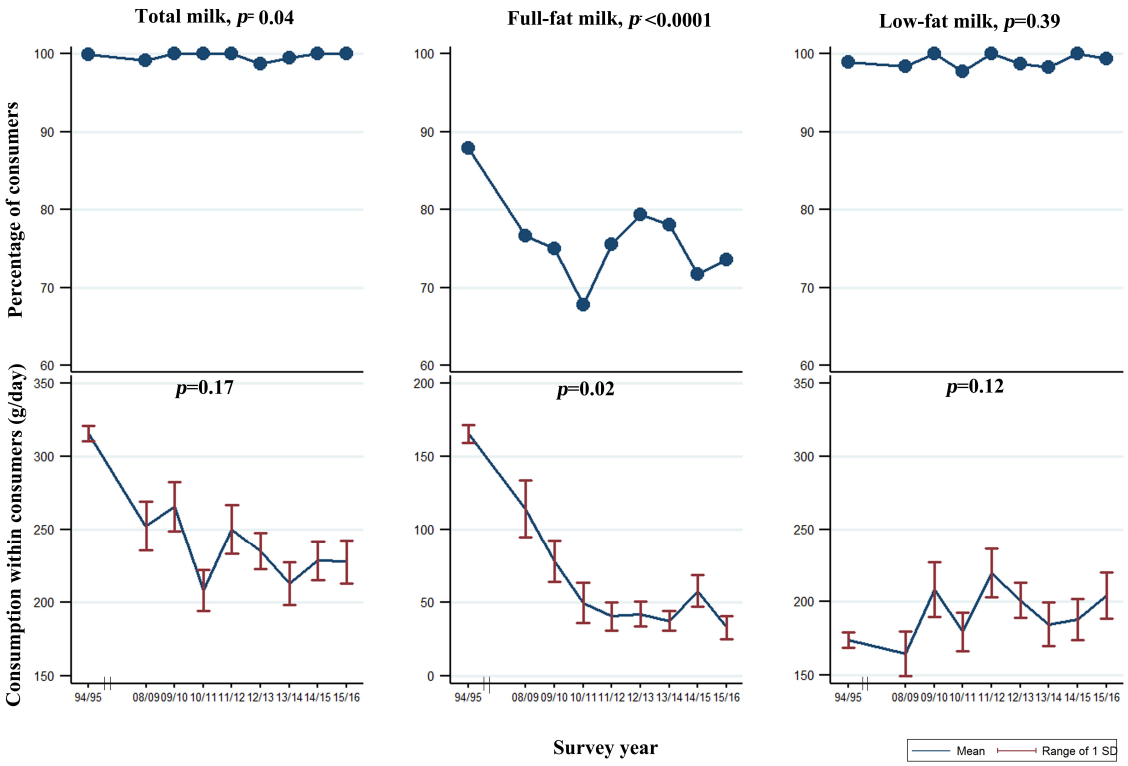


Fig. 2.6 Time trends of the percentage of consumers and the consumption levels within consumers of **total, full- and low-fat milk** among **elderly participants ≥65 years** in the old survey (1994/1995) and the rolling programme years (2008-2016) of the National Diet and Nutrition Survey (NDNS) in the UK. *P* for trend is estimated for years 2008/2009-2015/2016 only for consumption within consumers, but for all the years for the percentage of consumers. ‖ denotes a time gap of more than one year between the old surveys and the rolling programme.

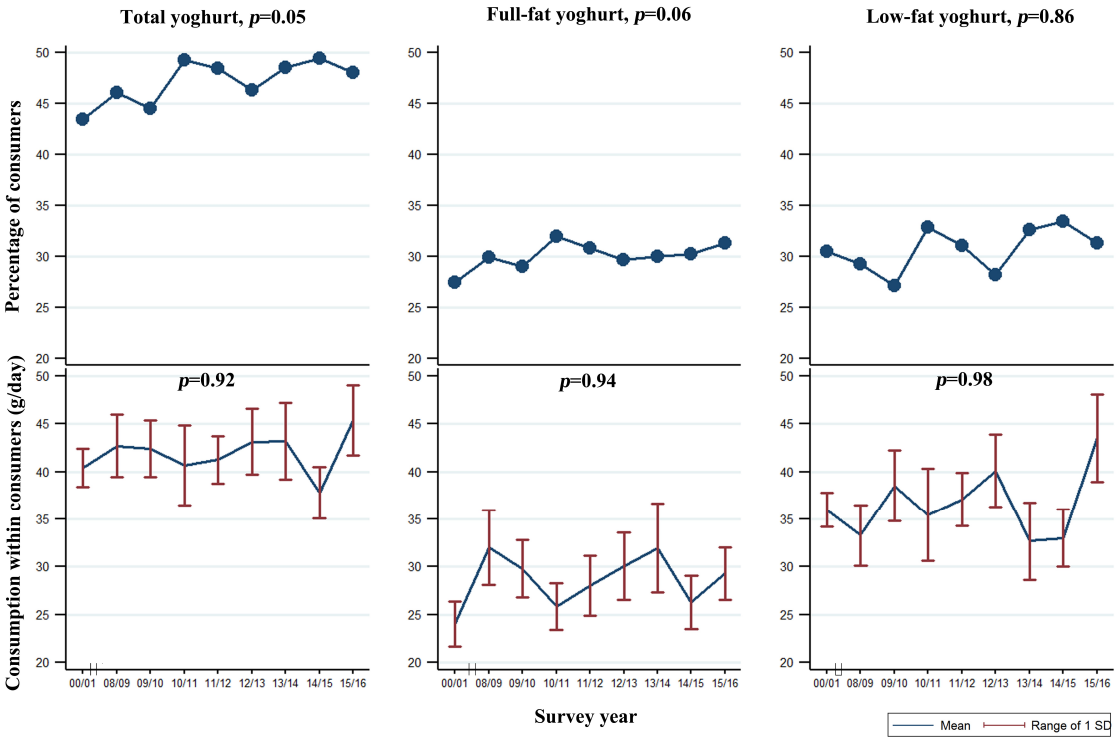


Fig. 2.7 Time trends of the percentage of consumers and the consumption levels within consumers of **total, full- and low-fat yoghurt** among **adult participants <65 years** in the old survey (2000/2001) and the rolling programme years (2008-2016) of the National Diet and Nutrition Survey (NDNS) in the UK. *P* for trend is estimated for years 2008/2009-2015/2016 only for consumption within consumers, but for all the years for the percentage of consumers. || denotes a time gap of more than one year between the old surveys and the rolling programme.

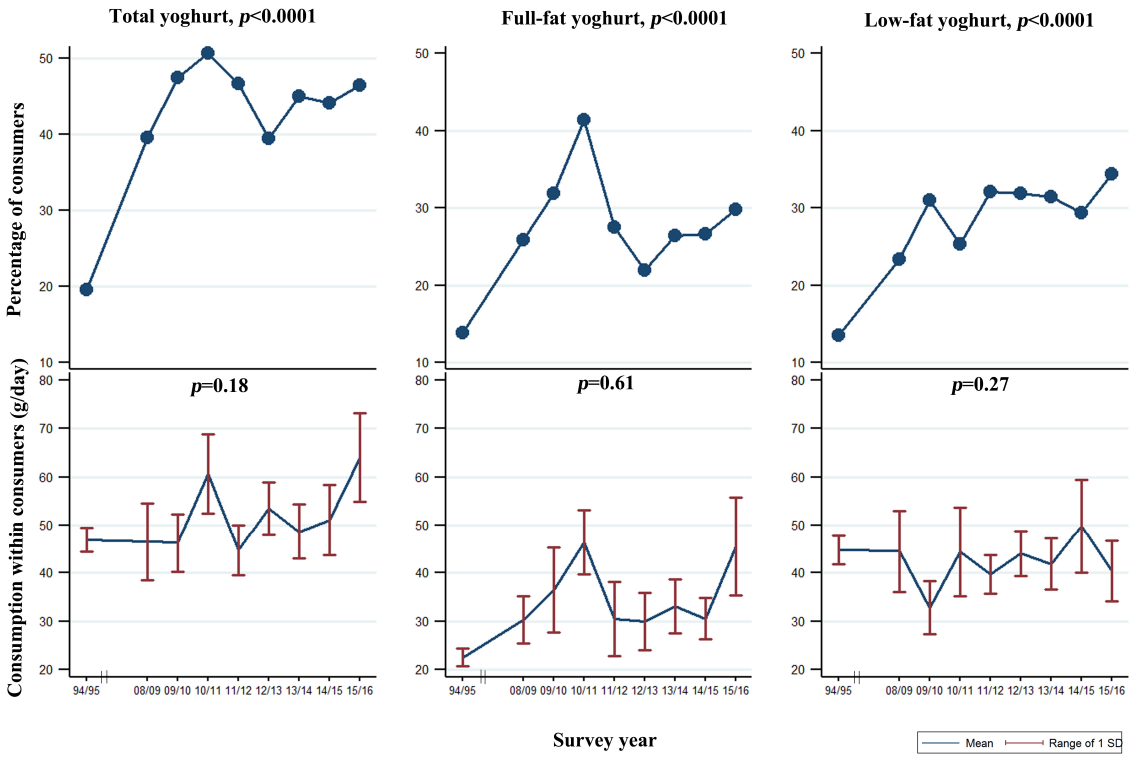


Fig. 2.8 Time trends of the percentage of consumers and the consumption levels within consumers of **total, full- and low-fat yoghurt** among **elderly participants ≥65 years** in the old survey (1994/1995) and the rolling programme years (2008-2016) of the National Diet and Nutrition Survey (NDNS) in the UK. *P* for trend is estimated for years 2008/2009-2015/2016 only for consumption within consumers, but for all the years for the percentage of consumers. ‖ denotes a time gap of more than one year between the old surveys and the rolling programme.

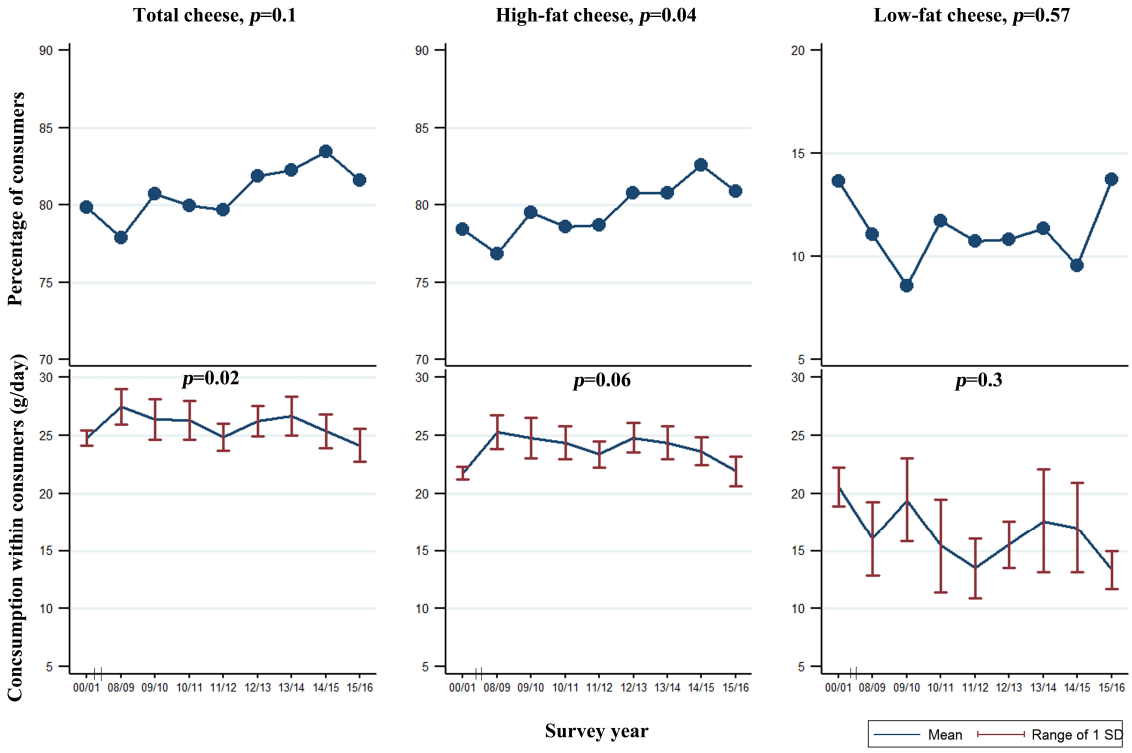


Fig. 2.9 Time trends of the percentage of consumers and the consumption levels within consumers of **total, high- and low-fat cheese** among **adult participants <65 years** in the old survey (2000/2001) and the rolling programme years (2008-2016) of the National Diet and Nutrition Survey (NDNS) in the UK. *P* for trend is estimated for years 2008/2009-2015/2016 only for consumption within consumers, but for all the years for the percentage of consumers. || denotes a time gap of more than one year between the old surveys and the rolling programme.

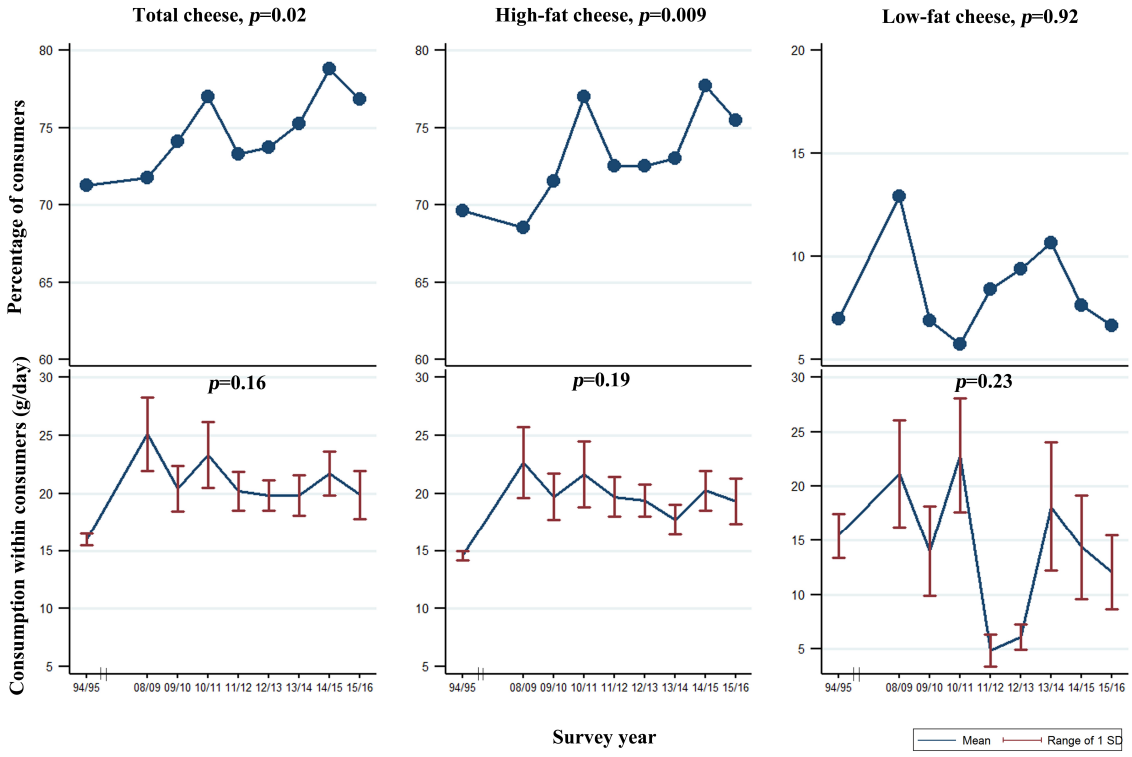


Fig. 2.10 Time trends of the percentage of consumers and the consumption levels within consumers of **total, high- and low-fat cheese** among **elderly participants ≥65 years** in the old survey (1994/1995) and the rolling programme years (2008-2016) of the National Diet and Nutrition Survey (NDNS) in the UK. *P* for trend is estimated for years 2008/2009-2015/2016 only for consumption within consumers, but for all the years for the percentage of consumers. ‖ denotes a time gap of more than one year between the old surveys and the rolling programme.

Cream consumption

The percentage of cream consumers significantly decreased for both age groups. No significant changes were observed within consumers. In 2015/2016, cream consumption was 5.7±0.4 and 7.4±1.1 g/day for adults <65 years and elderly people respectively (Figures 2.11, 2.12).

Fermented dairy consumption

Fermented dairy consumers ≥65 years significantly increased from 74.6% in 1994/1995 to 82.8% in 2015/2016 (Figure 2.12). No significant changes were observed within consumers. In 2015/2016, fermented dairy consumption was 47±2.5 and 52.6±5.9 for adults <65 years and elderly people respectively (Figures 2.11, 2.12).

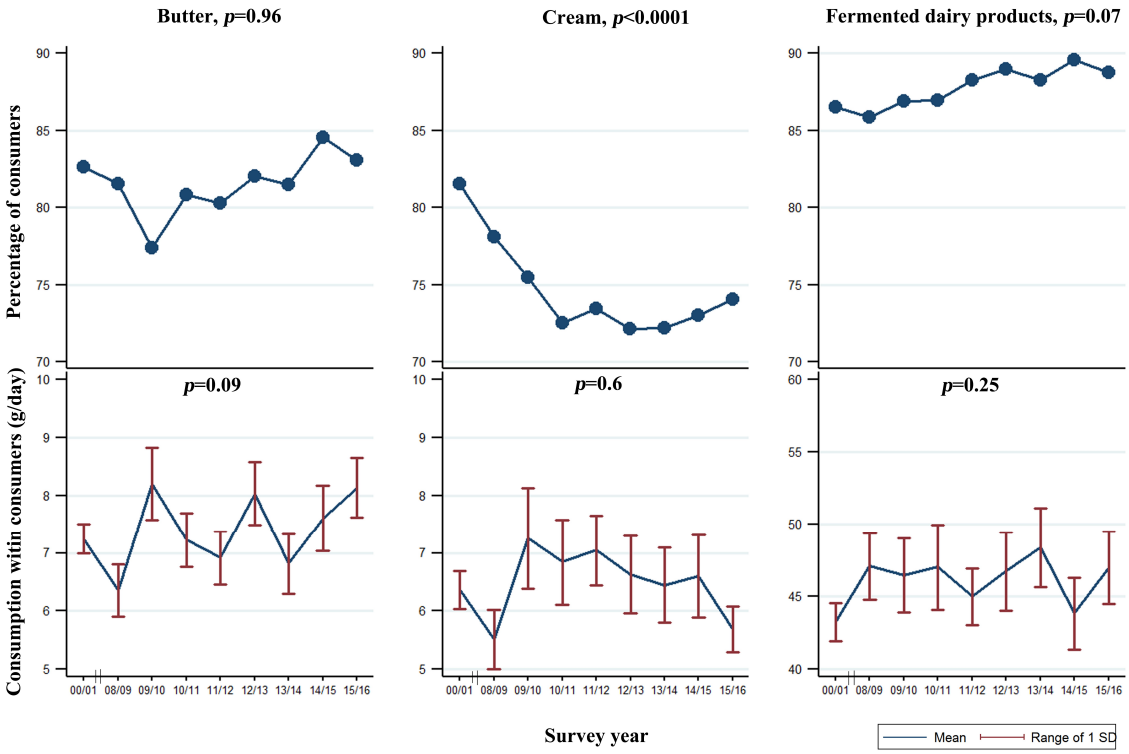


Fig. 2.11 Time trends of the percentage of consumers and the consumption levels within consumers of **butter, cream and fermented dairy products** (sum of yoghurt and cheese) among **adult participants <65 years** in the old survey (2000/2001) and the rolling programme years (2008-2016) of the National Diet and Nutrition Survey (NDNS) in the UK. *P* for trend is estimated for years 2008/2009-2015/2016 only for consumption within consumers, but for all the years for the percentage of consumers. ‖ denotes a time gap of more than one year between the old surveys and the rolling programme.

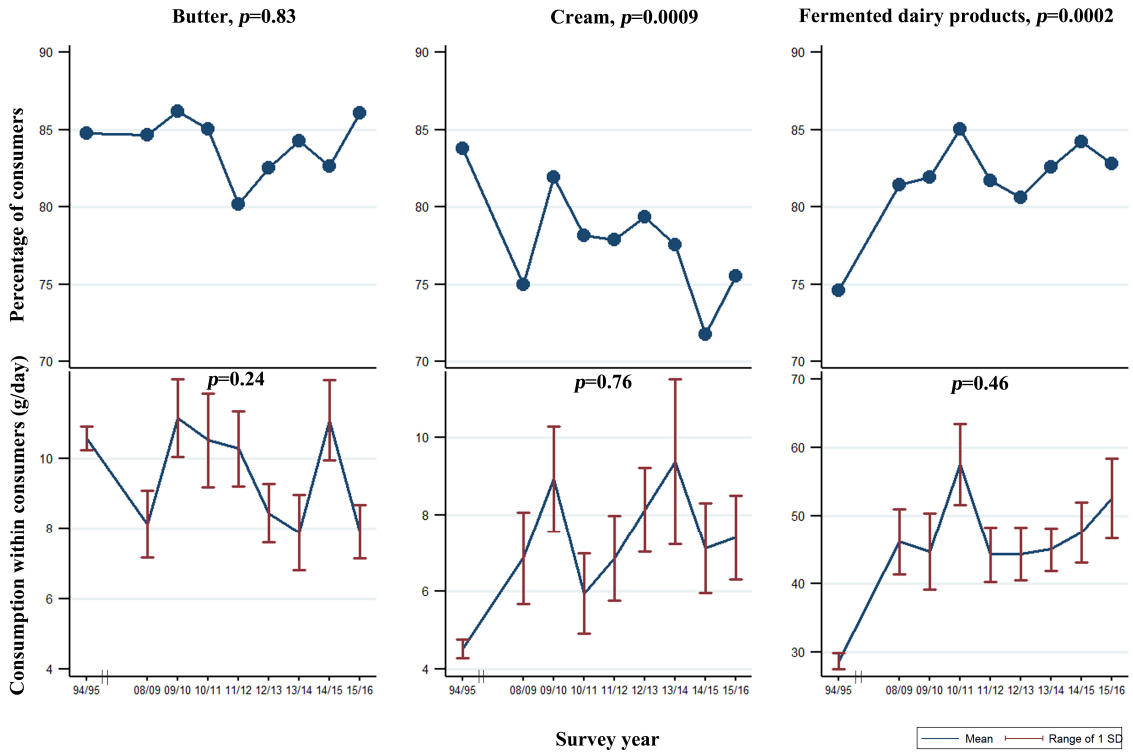


Fig. 2.12 Time trends of the percentage of consumers and the consumption levels within consumers of **butter, cream and fermented dairy products** (sum of yoghurt and cheese) among **elderly participants ≥65 years** in the old survey (1994/1995) and the rolling programme years (2008-2016) of the National Diet and Nutrition Survey (NDNS) in the UK. *P* for trend is estimated for years 2008/2009-2015/2016 only for consumption within consumers, but for all the years for the percentage of consumers. ‖ denotes a time gap of more than one year between the old surveys and the rolling programme.

2.5.3 Dairy consumption by sex

Overall, men consumed more total dairy products (264.3±4.6 g/day for <65 years and 321.4±6 g/day for ≥65 years) than women (234.1±3.2 for <65 years and 319.6±4 g/day for ≥65 years) over the 8-year period from 2008/2009 to 2015/2016 (**Table 2.3**). This difference was attributed mainly to a higher consumption of milk compared to women. Conversely, women <65 years consumed more low-fat yoghurt (38.1±1.4 g/day) than men (34.5±1.8 g/day) of the same age category. Women ≥65 years also consumed more total yoghurt (both low- and full-fat; 50.2±1.9 g/day) than elderly men (42.2±2.7 g/day; **Table 2.3**). Men <65 years consumed slightly higher amounts of total and high-fat cheese than women of the same age group, whereas women consumed higher amounts of low-fat cheese. For butter and cream, the amounts consumed were similar for both age groups.

Table 2.3 Means, SE, median and IQR of total and types of dairy products within consumers and percentage of consumers among men and women <65 years and ≥65 years in the old survey years and the first 8 years of the rolling programme of the National Diet and Nutrition Survey (NDNS) (2008-2016)

		Men					Women					Total				
<65 years †	**N**	2,482					3,339					5,821				
Dairy products (g/d)		**% ‡**	**Mean**	**SE**	**Median**	**IQR**	**% ‡**	**Mean**	**SE**	**Median**	**IQR**	**% ‡**	**Mean**	**SE**	**Median**	**IQR**
Milk	Full-fat	76.1	57.3	3.1	2.3	0-25.8	78.6	42.8	2.4	2	0.1-17	77.5	49.8	2	2.1	0.1-20
	Low-fat	98.5	171.8	4	124.8	40.5-250.1	98.9	151.1	2.9	117.5	38-218.1	98.7	161.3	2.5	120.4	39.1-232.6
	Total	99.4	213.8	4.3	173.7	89.9-292.6	99.6	183.8	3	155.6	78.1-255.6	99.5	198.6	2.6	164.4	82.8-274.3
Yoghurt	Full-fat	29.2	29	1.8	0	0-3.3	29.8	27.4	1.2	0	0-3.8	29.5	28.2	1.1	0	0-3.6
	Low-fat	27.1	34.5	1.8	0	0-0.4	33.4	38.1	1.4	0	0-2.7	30.7	36.5	1.1	0	0-1.1
	Total	42.7	41.4	1.7	0	0-17.4	49.2	42.1	1.3	0	0-29.3	46.4	41.8	1	0	0-25.9
Cheese	High-fat	78.9	26.1	0.7	13.8	2.1-31.9	79.9	21.1	0.5	11.3	2-25.7	79.5	23.5	0.4	12.5	2-28.8
	Low-fat	9.3	14.7	1.5	0	0-0	13.6	18.5	1.2	0	0-0	11.8	17	1	0	0-0
	Total	79.9	27.5	0.7	15	2.6-33.9	81.3	23.8	0.5	13.3	2.5-29.3	80.7	25.6	0.4	14.3	2.5-31.4
Cream		74	6.5	0.3	0.7	0-5	77.6	6.4	0.2	1	0-5.6	76.1	6.4	0.2	0.9	0-5.3
Butter		82.6	7.8	0.2	3.3	0.6-8.9	81.3	7	0.2	2.8	0.5-7.9	81.9	7.4	0.2	3	0.5-8.5
Fermented dairy products §		86.5	46.3	1.2	26.3	7.1-56.7	88.4	45.2	1	27.4	7.9-58.8	87.6	45.7	0.8	27.2	7.5-57.8
Total dairy products	High-fat ‖	98.1	77	2.5	40.3	16.7-85.8	98.1	62.6	2	33.1	15.2-66	98.1	69.7	1.6	36.1	16-76
	Low-fat #	99	190.1	4.2	146.7	56.2-268.7	99.2	174	3.1	139.9	55.3-251.9	99.1	182	2.7	142.2	56-261.1
	Total *	99.8	264.2	4.6	228.4	128-348.8	99.9	234.1	3.2	205.9	118.2-319	99.9	249	2.9	215.5	123.8-333.9
≥65 years †	**N**	1,333					1,531					2,864				
		% ‡	**Mean**	**SE**	**Median**	**IQR**	**% ‡**	**Mean**	**SE**	**Median**	**IQR**	**% ‡**	**Mean**	**SE**	**Median**	**IQR**
Milk	Full-fat	82.1	110.1	5.6	8.2	0.1-111.1	82.8	115.6	4.3	7.5	0.3-138.3	83.4	119.3	5.9	7.9	0.2-128.5
	Low-fat	99.2	192.3	5.9	156.4	36.9-294.4	99	183.8	3.8	137.6	32.6-269.3	98.8	178	5	146.4	34.7-281.4
	Total	99.8	279	5.9	251.8	149.4-367.7	99.8	276.2	4	236.8	146.7-366.4	99.9	274.2	5.3	245.5	148.8-366.6
Yoghurt	Full-fat	15.2	28.1	2.8	0	0-0	19.4	31.2	1.9	0	0-0	23.1	32.9	2.4	0	0-0
	Low-fat	16.2	36.3	3.1	0	0-0	20.1	43.2	1.9	0	0-0	23.5	47.2	2.3	0	0-0
	Total	23.9	42.2	2.7	0	0-2.1	29.4	50.2	1.9	0	0-28.8	34.1	54.7	2.4	0	0-20.6
Cheese	High-fat	72.8	20.6	0.7	10.3	0-22.9	71.2	17.3	0.4	6.5	0-15.6	69.9	15	0.5	7.6	0-18.6
	Low-fat	6	14.5	2.2	0	0-0	7.6	14.8	1.3	0	0-0	9.1	14.9	1.7	0	0-0
	Total	73.8	21.5	0.7	10.7	0.7-23.6	72.8	18.6	0.5	7.2	0-17.3	72	16.5	0.6	8.3	0-20
Cream		81	6.7	0.4	0.7	0-6.2	81	6	0.3	0.9	0-4.9	81.1	5.6	0.3	0.8	0-5.5
Butter		84	10.3	0.4	3.8	0.8-12.2	84.4	10	0.3	4.1	0.9-12.6	84.7	9.7	0.3	4	0.8-12.5
Fermented dairy products §		77.8	36	1.5	16.3	3-40.7	77.7	38.4	1.1	14	1.8-46.8	77.7	40.1	1.5	15	2-43.6
Total dairy products	High-fat ‖	99.5	117.3	4.7	48.8	21.3-149	99.1	120.5	3.7	42.5	18.5-168.3	98.7	122.7	5.2	45.5	19.5-156.6
	Low-fat #	99.2	205.8	6	173.6	51-313.9	99.1	201.8	3.9	162.4	44.5-296	99	199	5.2	167.4	47-304.6
	Total *	99.9	321.4	6	300.2	185.1-423.3	99.9	319.6	4	290	191.4-409.3	99.9	318.3	5.3	292.4	190-416.6

†The survey year 2000/2001 was used for adults <65 years and 1994/1995 was used for adults ≥65 years; ‡Percentage of consumers; §Sum of yoghurt and cheese; ‖Sum of full-fat milk, high-fat cheese, total cream and butter; #Sum of low-fat milk, total yoghurt, and low-fat cheese; *Sum of total milk, total yoghurt, total cheese, total cream and butter

2.5.4 Dairy contribution to nutrient intakes

Total energy and macronutrient intakes

Table 2.4 shows the contribution of total and types of dairy products to macronutrient intakes in the total sample and within dairy consumers for each of the years of the rolling programme. The contribution of total dairy consumption to total energy intake ranged from 13.5% in 2011/2012 to 14.6% in 2009/2010 with most of it coming from high-fat dairy products (from 8.3% in 2011/2012 to 9.4% in 2009/2010). Of the subtypes, low-fat milk was the highest contributor to total energy intake overall (from 3.9% to 4.4%) and high-fat cheese was the highest contributor within consumers (from 4.3% to 4.9%).

Total dairy products contributed by 24.4%-25.8% to total fat intake, 40.2%-42.2% to saturated fat intake, 16.2%-17.7% to cis-monounsaturated fat intake, 51%-62.1% to trans fat intake, 5.1%-5.7% to carbohydrate intake, 12.7%-13.8% to total sugar intake and 17.7%-19.1% to protein intake. Specifically for the contribution to trans fat intake, it seems that there is a linear increase from 51% in 2008/2009 to 62.1% in 2015/2016. Of dairy subtypes, the highest contributor to total, saturated, cis-monounsaturated and trans fat was high-fat cheese both overall (7.6-8.6%, 12.8-14.8%, 5.2-5.9% and 16-22.2% respectively) and among consumers only (9.4-10.9%, 15.7-18.9%, 6.3-7.5% and 21.2-27.9% respectively). The highest dairy contributor to carbohydrate, total sugar and protein intake was low-fat milk both overall (3.4-4%, 8.5-9.8% and 7.7-9% respectively) and among consumers only (3.5-4%, 8.5-9.9% and 7.8-9% respectively).

Table 2.4 Percentage of the contribution of total and types of dairy products to **total energy and macronutrient intakes** in the first 8 years of the National Diet and Nutrition (NDNS) rolling programme (2008-2016)

		2008/2009			2009/2010			2010/2011			2011/2012			2012/2013			2013/2014			2014/2015			2015/2016		
N			612			536			531			643			695			724			740			748		
Dairy products (g/d)		**% cons †**	**% within cons ‡**	**% total §**	**% cons †**	**% within cons ‡**	**% total §**	**% cons †**	**% within cons ‡**	**% total §**	**% cons †**	**% within cons ‡**	**% total §**	**% cons †**	**% within cons ‡**	**% total §**	**% cons †**	**% within cons ‡**	**% total §**	**% cons †**	**% within cons ‡**	**% total §**	**% cons †**	**% within cons ‡**	**% total §**
Energy (kcal)																									
Milk	Full-fat	77	2.5	1.9	78.5	2.2	1.7	71.8	2.1	1.5	75.1	1.7	1.2	78.7	1.7	1.4	78	2	1.5	74.3	2	1.5	75.7	1.9	1.5
	Low-fat	99.5	4	4	99.1	4.4	4.4	97.9	4	3.9	100	4.4	4.4	98.3	4.4	4.3	98.3	4.5	4.4	98.8	4.3	4.2	97.6	4.4	4.3
	Total	99.8	5.9	5.9	99.8	6.1	6.1	99.4	5.4	5.4	100	5.7	5.7	99.1	5.7	5.7	99.3	6	6	99.5	5.7	5.7	98.9	5.9	5.8
Yoghurt	Full-fat	29.1	1.5	0.4	29.7	1.3	0.4	33.5	1.4	0.5	30.2	1.3	0.4	27.9	1.3	0.4	29.1	1.5	0.5	29.3	1.2	0.3	31	1.7	0.6
	Low-fat	28.1	1.3	0.4	28	1.2	0.3	31.6	1.1	0.4	31.3	1.3	0.4	29.1	1.5	0.4	32.3	1.2	0.4	32.4	1.2	0.4	32	1.6	0.4
	Total	44.8	1.8	0.8	45.1	1.6	0.7	49.5	1.7	0.8	48.1	1.7	0.8	44.7	1.8	0.8	47.7	1.8	0.8	48.1	1.6	0.7	47.7	2.2	1.1
Cheese	High-fat	75.2	4.9	3.7	77.8	4.8	3.8	78.3	4.9	3.8	77.4	4.6	3.6	78.8	4.8	3.9	78.9	4.6	3.6	81.4	4.6	3.8	79.8	4.3	3.5
	Low-fat	11.4	1.1	0.1	8.2	1	0.1	10.7	1	0.1	10.3	0.7	0.1	10.5	0.9	0.1	11.2	1.1	0.1	9.1	1.1	0.1	12.3	0.8	0.1
	Total	76.6	5	3.8	79.3	4.8	3.9	79.5	5	4	78.4	4.6	3.6	80	4.9	4	80.5	4.7	3.8	82.3	4.7	3.9	80.6	4.4	3.6
Cream		77.5	1.2	0.9	76.9	1.3	1	73.4	1.2	0.9	74.3	1.3	1	73.8	1.3	0.9	73.5	1.3	1	72.7	1.1	0.8	74.3	1.2	0.9
Butter		82.2	2.7	2.2	79.3	3.6	2.9	81.5	3.1	2.6	80.2	3	2.5	82.2	3.3	2.7	82.2	3	2.5	84.1	3.4	2.8	83.7	3.4	2.8
Fermented dairy products		85	5.4	4.6	85.8	5.3	4.6	86.6	5.5	4.8	86.9	5.1	4.4	87.1	5.4	4.8	86.9	5.3	4.6	88.2	5.3	4.6	87.6	5.2	4.7
Total dairy products	High-fat	97.9	8.9	8.8	98.5	9.6	9.4	97.6	9.1	8.8	97.7	8.5	8.3	98.1	9.1	8.9	97.8	8.7	8.6	98.4	9	8.9	98	8.8	8.7
	Low-fat	99.7	5	5	99.1	5.3	5.2	98.5	4.9	4.8	100	5.3	5.3	98.8	5.3	5.2	98.6	5.5	5.4	99.2	5.1	5.1	98.5	5.6	5.5
	Total	99.8	13.7	13.7	99.8	14.7	14.6	100	13.6	13.6	100	13.5	13.5	99.7	14.2	14.1	99.9	14	14	99.9	13.9	13.9	99.9	14.2	14.2
Total fat (g)																									
Milk	Full-fat	77	3.9	3	78.5	3.5	2.7	71.8	3.3	2.3	75.1	2.7	2	78.7	2.8	2.2	78	3.2	2.5	74.3	3.1	2.3	75.7	3	2.3
	Low-fat	99.5	3.5	3.5	99.1	3.5	3.4	97.9	3.2	3.1	100	3.8	3.8	98.3	3.7	3.6	98.3	3.8	3.7	98.8	3.7	3.7	97.6	3.6	3.6
	Total	99.8	6.5	6.5	99.8	6.1	6.1	99.4	5.5	5.5	100	5.8	5.8	99.1	5.9	5.8	99.3	6.3	6.2	99.5	6	6	98.9	5.9	5.9
Yoghurt	Full-fat	29.1	2	0.6	29.7	1.9	0.6	33.5	1.9	0.7	30.2	2	0.6	27.9	1.9	0.6	29.1	2.3	0.7	29.3	1.8	0.5	31	2.6	0.9
	Low-fat	28.1	0.6	0.2	28	0.6	0.2	31.6	0.5	0.2	31.3	0.6	0.2	29.1	0.8	0.2	32.3	0.6	0.2	32.4	0.5	0.2	32	0.7	0.2
	Total	44.8	1.7	0.8	45.1	1.7	0.8	49.5	1.7	0.8	48.1	1.7	0.8	44.7	1.7	0.8	47.7	1.9	0.9	48.1	1.4	0.7	47.7	2.3	1.1
Cheese	High-fat	75.2	10.9	8.3	77.8	10.4	8.2	78.3	10.9	8.5	77.4	10.3	8	78.8	10.7	8.6	78.9	10.1	8	81.4	10.1	8.3	79.8	9.4	7.6
	Low-fat	11.4	1.4	0.2	8.2	1.3	0.1	10.7	1.7	0.2	10.3	0.9	0.1	10.5	1.2	0.1	11.2	1.3	0.2	9.1	1.3	0.1	12.3	1.2	0.2
	Total	76.6	10.9	8.4	79.3	10.3	8.3	79.5	11.1	8.7	78.4	10.3	8.1	80	10.7	8.7	80.5	10.2	8.1	82.3	10.1	8.4	80.6	9.5	7.8
Cream		77.5	3.1	2.4	76.9	3.5	2.6	73.4	3.3	2.5	74.3	3.5	2.6	73.8	3.5	2.6	73.5	3.6	2.6	72.7	3	2.2	74.3	3.1	2.3
Butter		82.2	7.7	6.3	79.3	9.9	7.9	81.5	8.8	7.2	80.2	8.7	7.1	82.2	9.4	7.9	82.2	8.3	6.9	84.1	9.5	7.8	83.7	9.3	7.6
Fermented dairy products		85	10.7	9.2	85.8	10.5	9.1	86.6	11.1	9.5	86.9	10.2	8.9	87.1	10.8	9.5	86.9	10.4	9	88.2	10.3	9.1	87.6	10	8.9
Total dairy products	High-fat	97.9	20.4	20	98.5	21.8	21.5	97.6	21.1	20.5	97.7	20.2	19.7	98.1	21.6	21.3	97.8	20.3	20	98.4	21	20.6	98	20.2	20
	Low-fat	99.7	4.4	4.4	99.1	4.3	4.3	98.5	4.2	4.2	100	4.7	4.7	98.8	4.6	4.5	98.6	4.9	4.8	99.2	4.5	4.4	98.5	4.9	4.8
	Total	99.8	24.4	24.4	99.8	25.8	25.7	100	24.7	24.7	100	24.4	24.4	99.7	25.9	25.8	99.9	24.8	24.8	99.9	25.1	25	99.9	24.8	24.8
Saturated fat (g)																									
Milk	Full-fat	77	6.3	4.9	78.5	5.5	4.3	71.8	5.4	3.9	75.1	4.4	3.3	78.7	4.6	3.7	78	5.3	4.1	74.3	5.1	3.8	75.7	5	3.9
	Low-fat	99.5	6	6	99.1	5.9	5.9	97.9	5.6	5.5	100	6.6	6.6	98.3	6.3	6.1	98.3	6.7	6.5	98.8	6.4	6.3	97.6	6.4	6.3
	Total	99.8	10.8	10.8	99.8	10.2	10.1	99.4	9.4	9.3	100	9.9	9.9	99.1	9.9	9.8	99.3	10.7	10.6	99.5	10.1	10.1	98.9	10.3	10.2
Yoghurt	Full-fat	29.1	3.5	1.1	29.7	3.3	1	33.5	3.3	1.1	30.2	3.5	1.1	27.9	3.3	1	29.1	4.1	1.3	29.3	3.1	0.8	31	4.4	1.6
	Low-fat	28.1	1.1	0.3	28	1	0.3	31.6	0.9	0.3	31.3	1	0.3	29.1	1.4	0.4	32.3	0.9	0.3	32.4	0.9	0.3	32	1.4	0.4
	Total	44.8	3.1	1.4	45.1	2.8	1.3	49.5	2.8	1.4	48.1	2.9	1.4	44.7	3	1.4	47.7	3.4	1.6	48.1	2.4	1.1	47.7	4	1.9
Cheese	High-fat	75.2	18.3	13.9	77.8	17.1	13.5	78.3	18.9	14.8	77.4	17.4	13.6	78.8	17.6	14.2	78.9	17.1	13.5	81.4	17.2	14.1	79.8	15.7	12.8
	Low-fat	11.4	2.6	0.3	8.2	2.2	0.2	10.7	3	0.3	10.3	1.8	0.2	10.5	2.2	0.2	11.2	2.2	0.3	9.1	2.3	0.2	12.3	2	0.3
	Total	76.6	18.4	14.2	79.3	17	13.7	79.5	19.2	15.1	78.4	17.5	13.7	80	17.6	14.4	80.5	17.2	13.7	82.3	17.3	14.4	80.6	15.9	13.1
Cream		77.5	5.1	3.9	76.9	5.5	4.2	73.4	5.4	3.9	74.3	5.7	4.3	73.8	5.6	4.1	73.5	5.9	4.3	72.7	5.1	3.6	74.3	5.1	3.9
Butter		82.2	12.1	9.9	79.3	15.5	12.4	81.5	13.9	11.4	80.2	14.1	11.5	82.2	14.9	12.6	82.2	13.6	11.2	84.1	15	12.4	83.7	15	12.3
Fermented dairy products		85	18.2	15.6	85.8	17.4	15	86.6	19.1	16.5	86.9	17.4	15.1	87.1	17.8	15.8	86.9	17.7	15.3	88.2	17.6	15.5	87.6	16.8	15
Total dairy products	High-fat	97.9	33.3	32.6	98.5	35	34.4	97.6	34.9	34	97.7	33.5	32.7	98.1	35.1	34.5	97.8	33.6	33	98.4	34.6	34	98	33.3	33
	Low-fat	99.7	7.7	7.6	99.1	7.4	7.3	98.5	7.3	7.2	100	8.2	8.2	98.8	7.8	7.7	98.6	8.6	8.4	99.2	7.7	7.6	98.5	8.6	8.5
	Total	99.8	40.3	40.2	99.8	41.8	41.7	100	41.1	41.1	100	40.9	40.9	99.7	42.4	42.2	99.9	41.5	41.5	99.9	41.7	41.6	99.9	41.4	41.4

Table 2.4 (continued)

		2008/2009			2009/2010			2010/2011			2011/2012			2012/2013			2013/2014			2014/2015			2015/2016		
N			612			536			531			643			695			724			740			748	
Dairy products (g/d)		**% cons †**	**% within cons ‡**	**% total §**	**% cons †**	**% within cons ‡**	**% total §**	**% cons †**	**% within cons ‡**	**% total §**	**% cons †**	**% within cons ‡**	**% total §**	**% cons †**	**% within cons ‡**	**% total §**	**% cons †**	**% within cons ‡**	**% total §**	**% cons †**	**% within cons ‡**	**% total §**	**% cons †**	**% within cons ‡**	**% total §**
cis-Monounsaturated fat (g)																									
Milk	Full-fat	77	2.7	2.1	78.5	2.5	1.9	71.8	2.2	1.6	75.1	1.8	1.3	78.7	1.8	1.5	78	2.2	1.7	74.3	2.1	1.6	75.7	2	1.5
	Low-fat	99.5	2.4	2.4	99.1	2.4	2.3	97.9	2.1	2.1	100	2.5	2.5	98.3	2.4	2.4	98.3	2.5	2.5	98.8	2.5	2.4	97.6	2.4	2.3
	Total	99.8	4.5	4.5	99.8	4.2	4.2	99.4	3.7	3.7	100	3.8	3.8	99.1	3.9	3.8	99.3	4.2	4.1	99.5	4	4	98.9	3.9	3.8
Yoghurt	Full-fat	29.1	1.4	0.4	29.7	1.4	0.4	33.5	1.3	0.5	30.2	1.4	0.4	27.9	1.3	0.4	29.1	1.6	0.5	29.3	1.3	0.3	31	1.8	0.6
	Low-fat	28.1	0.5	0.1	28	0.5	0.1	31.6	0.4	0.1	31.3	0.4	0.1	29.1	0.6	0.2	32.3	0.4	0.1	32.4	0.4	0.1	32	0.5	0.2
	Total	44.8	1.2	0.6	45.1	1.2	0.5	49.5	1.2	0.6	48.1	1.2	0.6	44.7	1.2	0.5	47.7	1.3	0.6	48.1	1	0.5	47.7	1.6	0.8
Cheese	High-fat	75.2	7.5	5.7	77.8	7.2	5.7	78.3	7.2	5.6	77.4	6.8	5.3	78.8	7.3	5.9	78.9	6.8	5.4	81.4	6.7	5.5	79.8	6.3	5.2
	Low-fat	11.4	1.1	0.1	8.2	0.8	0.1	10.7	1.1	0.1	10.3	0.6	0.1	10.5	0.9	0.1	11.2	0.9	0.1	9.1	0.9	0.1	12.3	0.8	0.1
	Total	76.6	7.5	5.8	79.3	7.1	5.7	79.5	7.3	5.8	78.4	6.8	5.4	80	7.3	6	80.5	6.9	5.5	82.3	6.8	5.6	80.6	6.4	5.3
Cream		77.5	2.1	1.6	76.9	2.3	1.7	73.4	2.2	1.6	74.3	2.3	1.7	73.8	2.3	1.7	73.5	2.4	1.7	72.7	1.9	1.4	74.3	2	1.5
Butter		82.2	5.3	4.3	79.3	6.8	5.5	81.5	6	4.9	80.2	5.8	4.7	82.2	6.3	5.3	82.2	5.5	4.6	84.1	6.9	5.7	83.7	6.5	5.4
Fermented dairy products		85	7.4	6.4	85.8	7.3	6.3	86.6	7.3	6.3	86.9	6.8	5.9	87.1	7.4	6.5	86.9	7.1	6.1	88.2	6.9	6.1	87.6	6.8	6
Total dairy products	High-fat	97.9	14	13.8	98.5	15	14.8	97.6	14.1	13.7	97.7	13.4	13.1	98.1	14.5	14.3	97.8	13.6	13.4	98.4	14.4	14.2	98	13.7	13.6
	Low-fat	99.7	3.1	3.1	99.1	3	3	98.5	2.8	2.8	100	3.1	3.1	98.8	3	3	98.6	3.3	3.2	99.2	3	3	98.5	3.2	3.2
	Total	99.8	16.8	16.8	99.8	17.8	17.7	100	16.5	16.5	100	16.2	16.2	99.7	17.4	17.3	99.9	16.6	16.6	99.9	17.2	17.1	99.9	16.8	16.8
Trans fat (g)																									
Milk	Full-fat	77	8.1	6.2	78.5	5.9	4.6	71.8	6.7	4.8	75.1	5.6	4.2	78.7	6.1	4.8	78	7.1	5.4	74.3	6.1	4.5	75.7	6.7	5.2
	Low-fat	99.5	7.3	7.3	99.1	7.7	7.6	97.9	8.4	8.3	100	11.6	11.6	98.3	10.7	10.5	98.3	11.3	11	98.8	11.2	11.1	97.6	11.5	11.2
	Total	99.8	13.5	13.5	99.8	12.2	12.2	99.4	13.2	13.1	100	15.8	15.8	99.1	15.4	15.3	99.3	16.6	16.5	99.5	15.7	15.6	98.9	16.6	16.4
Yoghurt	Full-fat	29.1	3	0.9	29.7	3.1	1	33.5	3.7	1.3	30.2	4.7	1.5	27.9	4.5	1.4	29.1	4.6	1.4	29.3	4.2	1.1	31	5.3	1.9
	Low-fat	28.1	0.5	0.1	28	0.7	0.2	31.6	0.8	0.3	31.3	1.1	0.3	29.1	1.8	0.5	32.3	0.8	0.2	32.4	1.1	0.3	32	1.3	0.4
	Total	44.8	2.4	1.1	45.1	2.6	1.2	49.5	3.1	1.5	48.1	3.7	1.8	44.7	4.1	1.9	47.7	3.6	1.7	48.1	3.1	1.4	47.7	4.6	2.2
Cheese	High-fat	75.2	21.1	16	77.8	21.5	17	78.3	27.3	21.3	77.4	27.4	21.4	78.8	27.1	21.8	78.9	27.9	22	81.4	26.9	22.2	79.8	25.7	20.9
	Low-fat	11.4	2.8	0.3	8.2	3	0.2	10.7	2.4	0.3	10.3	2.6	0.3	10.5	2.4	0.2	11.2	3.9	0.5	9.1	4.3	0.4	12.3	2.8	0.4
	Total	76.6	21.2	16.4	79.3	21.4	17.2	79.5	27.4	21.6	78.4	27.5	21.6	80	27	22.1	80.5	28.1	22.5	82.3	27.2	22.6	80.6	25.9	21.3
Cream		77.5	8.6	6.6	76.9	8.8	6.7	73.4	7.6	5.6	74.3	8.7	6.5	73.8	8.1	5.9	73.5	8.6	6.3	72.7	7.3	5.1	74.3	7.4	5.6
Butter		82.2	16.4	13.4	79.3	18.6	14.9	81.5	17.9	14.6	80.2	19.5	15.9	82.2	19.7	16.6	82.2	18.1	15	84.1	20.5	17	83.7	20.1	16.6
Fermented dairy products		85	20.3	17.4	85.8	21.4	18.4	86.6	26.8	23.1	86.9	27	23.4	87.1	27	23.9	86.9	27.9	24.1	88.2	27.4	24	87.6	26.3	23.5
Total dairy products	High-fat	97.9	43.2	42.3	98.5	43.9	43.2	97.6	47.6	46.3	97.7	49.2	47.9	98.1	50	49.2	97.8	49.5	48.7	98.4	49.6	48.8	98	48.9	48.3
	Low-fat	99.7	8.7	8.6	99.1	9.1	9	98.5	10.2	10	100	13.6	13.6	98.8	12.6	12.5	98.6	13.4	13.2	99.2	13.1	12.9	98.5	14	13.8
	Total	99.8	51	51	99.8	52.3	52.2	100	56.4	56.4	100	61.5	61.5	99.7	61.9	61.7	99.9	61.9	61.9	99.9	61.8	61.7	99.9	62.2	62.1
Carbohydrates (g)																									
Milk	Full-fat	77	1.5	1.1	78.5	1.3	1	71.8	1.2	0.9	75.1	1	0.7	78.7	1	0.8	78	1.2	0.9	74.3	1.2	0.9	75.7	1.1	0.9
	Low-fat	99.5	3.6	3.6	99.1	4	4	97.9	3.5	3.4	100	3.8	3.8	98.3	3.8	3.8	98.3	3.9	3.8	98.8	3.7	3.7	97.6	3.9	3.8
	Total	99.8	4.7	4.7	99.8	5	5	99.4	4.4	4.3	100	4.6	4.6	99.1	4.6	4.5	99.3	4.8	4.7	99.5	4.6	4.6	98.9	4.8	4.7
Yoghurt	Full-fat	29.1	0.9	0.3	29.7	0.8	0.3	33.5	0.8	0.3	30.2	0.8	0.2	27.9	0.9	0.3	29.1	0.9	0.3	29.3	0.7	0.2	31	1	0.4
	Low-fat	28.1	1.5	0.4	28	1.3	0.4	31.6	1.2	0.4	31.3	1.4	0.4	29.1	1.6	0.5	32.3	1.2	0.4	32.4	1.4	0.4	32	1.6	0.5
	Total	44.8	1.6	0.7	45.1	1.3	0.6	49.5	1.4	0.7	48.1	1.4	0.7	44.7	1.6	0.7	47.7	1.4	0.7	48.1	1.3	0.6	47.7	1.7	0.8
Cheese	High-fat	75.2	0	0	77.8	0	0	78.3	0	0	77.4	0	0	78.8	0	0	78.9	0	0	81.4	0	0	79.8	0	0
	Low-fat	11.4	0.3	0	8.2	0.3	0	10.7	0.3	0	10.3	0.2	0	10.5	0.2	0	11.2	0.3	0	9.1	0.3	0	12.3	0.2	0
	Total	76.6	0.1	0.1	79.3	0.1	0	79.5	0.1	0.1	78.4	0.1	0	80	0.1	0.1	80.5	0.1	0.1	82.3	0.1	0.1	80.6	0.1	0.1
Cream		77.5	0.1	0.1	76.9	0.1	0.1	73.4	0.1	0.1	74.3	0.1	0.1	73.8	0.1	0.1	73.5	0.1	0.1	72.7	0.1	0.1	74.3	0.1	0.1
Butter		82.2	0	0	79.3	0	0	81.5	0	0	80.2	0	0	82.2	0	0	82.2	0	0	84.1	0	0	83.7	0	0
Fermented dairy products		85	0.9	0.8	85.8	0.8	0.7	86.6	0.8	0.7	86.9	0.8	0.7	87.1	0.9	0.8	86.9	0.8	0.7	88.2	0.8	0.7	87.6	1	0.9
Total dairy products	High-fat	97.9	1.3	1.2	98.5	1.1	1.1	97.6	1	1	97.7	0.8	0.8	98.1	0.9	0.9	97.8	1	1	98.4	1	1	98	1	1
	Low-fat	99.7	4.3	4.3	99.1	4.7	4.6	98.5	4.2	4.1	100	4.5	4.5	98.8	4.6	4.5	98.6	4.6	4.5	99.2	4.4	4.4	98.5	4.8	4.7
	Total	99.8	5.6	5.6	99.8	5.8	5.7	100	5.1	5.1	100	5.3	5.3	99.7	5.4	5.4	99.9	5.5	5.5	99.9	5.3	5.3	99.9	5.7	5.7

Table 2.4 (continued)

		2008/2009			2009/2010			2010/2011			2011/2012			2012/2013			2013/2014			2014/2015			2015/2016		
N			612			536			531			643			695			724			740			748	
Dairy products (g/d)		**% cons †**	**% within cons ‡**	**% total §**	**% cons †**	**% within cons ‡**	**% total §**	**% cons †**	**% within cons ‡**	**% total §**	**% cons †**	**% within cons ‡**	**% total §**	**% cons †**	**% within cons ‡**	**% total §**	**% cons †**	**% within cons ‡**	**% total §**	**% cons †**	**% within cons ‡**	**% total §**	**% cons †**	**% within cons ‡**	**% total §**
Total sugars (g)																									
Milk	Full-fat	77	3.3	2.6	78.5	3.1	2.4	71.8	3.2	2.3	75.1	2.3	1.8	78.7	2.4	1.9	78	2.9	2.2	74.3	2.9	2.1	75.7	2.9	2.3
	Low-fat	99.5	8.5	8.5	99.1	9.9	9.8	97.9	8.7	8.5	100	9.4	9.4	98.3	9.4	9.2	98.3	9.5	9.3	98.8	9.5	9.4	97.6	9.5	9.3
	Total	99.8	11.1	11	99.8	12.2	12.1	99.4	10.9	10.8	100	11.1	11.1	99.1	11.2	11.1	99.3	11.7	11.5	99.5	11.6	11.6	98.9	11.7	11.6
Yoghurt	Full-fat	29.1	2.2	0.6	29.7	1.8	0.6	33.5	2	0.7	30.2	1.8	0.6	27.9	2	0.6	29.1	2.1	0.7	29.3	1.8	0.5	31	2.3	0.8
	Low-fat	28.1	3.5	1	28	3.1	0.8	31.6	2.8	0.9	31.3	3.2	1	29.1	3.4	1	32.3	2.7	0.9	32.4	3.2	1	32	3.8	1.1
	Total	44.8	3.7	1.6	45.1	3	1.4	49.5	3.1	1.5	48.1	3.2	1.5	44.7	3.4	1.6	47.7	3.2	1.5	48.1	3.2	1.5	47.7	3.9	1.9
Cheese	High-fat	75.2	0.1	0.1	77.8	0.1	0.1	78.3	0.1	0.1	77.4	0.1	0.1	78.8	0.1	0.1	78.9	0.1	0.1	81.4	0.1	0.1	79.8	0.1	0.1
	Low-fat	11.4	0.7	0.1	8.2	0.6	0	10.7	0.6	0.1	10.3	0.3	0	10.5	0.5	0	11.2	0.8	0.1	9.1	0.6	0.1	12.3	0.4	0.1
	Total	76.6	0.2	0.1	79.3	0.1	0.1	79.5	0.2	0.1	78.4	0.1	0.1	80	0.2	0.1	80.5	0.2	0.2	82.3	0.2	0.1	80.6	0.2	0.1
Cream		77.5	0.2	0.1	76.9	0.2	0.1	73.4	0.2	0.1	74.3	0.2	0.1	73.8	0.2	0.1	73.5	0.2	0.1	72.7	0.2	0.1	74.3	0.2	0.1
Butter		82.2	0.1	0.1	79.3	0.1	0.1	81.5	0.1	0.1	80.2	0.1	0.1	82.2	0.1	0.1	82.2	0.1	0.1	84.1	0.1	0.1	83.7	0.1	0.1
Fermented dairy products		85	2.1	1.8	85.8	1.7	1.5	86.6	2	1.7	86.9	1.9	1.6	87.1	1.9	1.7	86.9	1.9	1.7	88.2	1.8	1.6	87.6	2.3	2
Total dairy products	High-fat	97.9	2.9	2.8	98.5	2.7	2.6	97.6	2.6	2.6	97.7	2.1	2	98.1	2.3	2.2	97.8	2.5	2.5	98.4	2.5	2.4	98	2.6	2.6
	Low-fat	99.7	10.2	10.2	99.1	11.3	11.2	98.5	10.2	10.1	100	10.9	10.9	98.8	10.9	10.8	98.6	11.2	10.9	99.2	11	10.9	98.5	11.4	11.3
	Total	99.8	13	13	99.8	13.9	13.8	100	12.7	12.7	100	13	13	99.7	13.1	13	99.9	13.4	13.4	99.9	13.4	13.4	99.9	13.8	13.8
Protein (g)																									
Milk	Full-fat	77	3.5	2.7	78.5	3.1	2.4	71.8	2.9	2.1	75.1	2.2	1.7	78.7	2.3	1.8	78	2.7	2	74.3	2.7	2	75.7	2.6	2
	Low-fat	99.5	8.1	8.1	99.1	9	8.9	97.9	7.8	7.7	100	9	9	98.3	8.9	8.7	98.3	9	8.8	98.8	8.4	8.3	97.6	8.7	8.5
	Total	99.8	10.8	10.8	99.8	11.3	11.3	99.4	9.8	9.8	100	10.6	10.6	99.1	10.6	10.5	99.3	10.9	10.8	99.5	10.4	10.3	98.9	10.6	10.5
Yoghurt	Full-fat	29.1	2.1	0.6	29.7	1.9	0.6	33.5	1.9	0.6	30.2	1.7	0.5	27.9	1.8	0.5	29.1	2.1	0.6	29.3	1.6	0.4	31	2.2	0.8
	Low-fat	28.1	2.6	0.7	28	2.3	0.6	31.6	2.4	0.8	31.3	2.6	0.8	29.1	2.9	0.8	32.3	2.3	0.7	32.4	2.6	0.8	32	3.2	0.9
	Total	44.8	3	1.3	45.1	2.7	1.2	49.5	2.8	1.4	48.1	2.7	1.3	44.7	2.9	1.3	47.7	2.9	1.4	48.1	2.7	1.2	47.7	3.5	1.7
Cheese	High-fat	75.2	7.9	6	77.8	7.7	6.1	78.3	7.8	6.1	77.4	7.4	5.7	78.8	7.7	6.2	78.9	7.5	5.9	81.4	7.4	6.1	79.8	6.9	5.7
	Low-fat	11.4	2.8	0.3	8.2	2.9	0.2	10.7	2.2	0.2	10.3	2	0.2	10.5	2.5	0.2	11.2	2.7	0.3	9.1	3.1	0.3	12.3	1.8	0.2
	Total	76.6	8.2	6.4	79.3	7.8	6.3	79.5	8.1	6.4	78.4	7.6	5.9	80	7.8	6.4	80.5	7.8	6.2	82.3	7.6	6.3	80.6	7.1	5.9
Cream		77.5	0.2	0.2	76.9	0.2	0.2	73.4	0.2	0.2	74.3	0.2	0.2	73.8	0.2	0.2	73.5	0.3	0.2	72.7	0.2	0.2	74.3	0.2	0.2
Butter		82.2	0.1	0.1	79.3	0.1	0.1	81.5	0.1	0.1	80.2	0.1	0.1	82.2	0.1	0.1	82.2	0.1	0.1	84.1	0.1	0.1	83.7	0.1	0.1
Fermented dairy products		85	9	7.7	85.8	8.7	7.5	86.6	9	7.8	86.9	8.4	7.3	87.1	8.7	7.7	86.9	8.8	7.6	88.2	8.6	7.6	87.6	8.5	7.6
Total dairy products	High-fat	97.9	9.1	8.9	98.5	8.9	8.7	97.6	8.7	8.4	97.7	7.8	7.6	98.1	8.3	8.2	97.8	8.3	8.2	98.4	8.4	8.3	98	8	7.9
	Low-fat	99.7	9.8	9.7	99.1	10.5	10.4	98.5	9.4	9.3	100	10.5	10.5	98.8	10.4	10.3	98.6	10.7	10.5	99.2	9.9	9.8	98.5	10.6	10.5
	Total	99.8	18.7	18.7	99.8	19.1	19.1	100	17.7	17.7	100	18.1	18.1	99.7	18.6	18.5	99.9	18.7	18.7	99.9	18.2	18.1	99.9	18.4	18.3

†Percentage of consumers
‡Percent contribution of dairy products to the nutrient intake within dairy consumers
§Percent contribution of dairy products to the nutrient intake in the total sample

Micronutrient intakes

Table 2.5 presents the contribution of total and types of dairy products to micronutrient intakes overall and within dairy consumers for each survey year of the rolling programme. Total dairy consumption contributed to intakes of vitamin A by 24.3-32.1%, vitamin D by 11.4-13.3%, vitamin B_{12} by 39.4-41.8%, calcium by 48.4-50.5%, potassium by 14.5-15.8%, magnesium by 12.3-13.7%, phosphorus by 25.7-27.8% and zinc by 19.8-21.5%.

The highest dairy contributor to vitamin A intakes both overall (7.8-11%) and among consumers only (11-19.8%) was high-fat cheese. For vitamin D, vitamin B_{12}, calcium, potassium, magnesium, phosphorus and zinc, low-fat milk contributed the most both overall and among consumers only. Specifically, among all (overall) survey participants including consumers and non-consumers, low-fat milk contributed to micronutrient intakes as follows: vitamin D, 3.8-5.2%; vitamin B_{12}, 21.4-25.6%; calcium, 23.5-26.3%; potassium, 9.3-10.4%; magnesium, 6.9-7.9%; phosphorus, 12.9-14.8% and; zinc 8-9.4% among consumers and non-consumers. Similarly, low-fat milk also contributed the most to micronutrients as follows: vitamin D, 3.8-5.6%; vitamin B_{12}, 22.9-26.8%; calcium, 25.2-27.4%; potassium, 9.9-10.9%; magnesium, 7.4-8.3%; phosphorus, 13.8-15.4%; and zinc, 8.6-9.7% among consumers only.

Table 2.5 Percentage of the contribution of total and types of dairy products to micronutrient intakes in the first 8 years of the National Diet and Nutrition (NDNS) rolling programme (2008-2016)

		2008/2009			2009/2010			2010/2011			2011/2012			2012/2013			2013/2014			2014/2015			2015/2016		
N			612			536			531			643			695			724			740			748		
Dairy products (g/d)		% cons †	% within cons ‡	% total §	% cons †	% within cons ‡	% total §	% cons †	% within cons ‡	% total §	% cons †	% within cons ‡	% total §	% cons †	% within cons ‡	% total §	% cons †	% within cons ‡	% total §	% cons †	% within cons ‡	% total §	% cons †	% within cons ‡	% total §
Vitamin A (μg)																									
Milk	Full-fat	77	5.3	2.7	78.5	4.3	2.2	71.8	3.9	1.8	75.1	3.4	1.6	78.7	2.6	2.1	3.4	2.6	2.6	74.3	2.8	2.1	75.7	3.2	2.5
	Low-fat	99.5	4.4	4.1	99.1	4.6	4.4	97.9	4.8	4.5	100	5.4	5.2	98.3	5.1	5	5.9	5.7	5.7	98.8	6.2	6.2	97.6	5.4	5.3
	Total	99.8	7	6.8	99.8	6.7	6.6	99.4	6.5	6.3	100	6.9	6.7	99.1	7.1	7	8.4	8.4	8.4	99.5	8.3	8.3	98.9	7.8	7.8
Yoghurt	Full-fat	29.1	3.3	0.4	29.7	3.2	0.5	33.5	3.8	0.6	30.2	3.8	0.5	27.9	1.9	0.6	2.7	0.9	0.9	29.3	2.2	0.6	31	3.1	1.1
	Low-fat	28.1	1.3	0.2	28	1	0.1	31.6	1.1	0.2	31.3	1	0.2	29.1	0.7	0.2	0.7	0.2	0.2	32.4	0.8	0.2	32	0.7	0.2
	Total	44.8	2.7	0.6	45.1	2.6	0.6	49.5	2.9	0.7	48.1	2.8	0.7	44.7	1.7	0.8	2.3	1.1	1.1	48.1	1.8	0.8	47.7	2.6	1.3
Cheese	High-fat	75.2	17.7	7.8	77.8	19.3	8.5	78.3	19.8	8.7	77.4	19	8.6	78.8	12	9.6	13.9	11	11	81.4	12.9	10.6	79.8	11.5	9.4
	Low-fat	11.4	3	0.1	8.2	6.5	0.2	10.7	3.7	0.1	10.3	6.4	0.2	10.5	1.5	0.1	2	0.2	0.2	9.1	2.5	0.2	12.3	2.2	0.3
	Total	76.6	17.3	7.9	79.3	19.1	8.7	79.5	19.5	8.9	78.4	19.1	8.8	80	12	9.8	14	11.2	11.2	82.3	13	10.8	80.6	11.7	9.6
Cream		77.5	6.2	2.7	76.9	7.5	3.1	73.4	6.4	2.7	74.3	7.4	3.1	73.8	4.1	3	4.9	3.6	3.6	72.7	4.6	3.3	74.3	3.9	3
Butter		82.2	12	6	79.3	14.2	7.4	81.5	13.5	7	80.2	12.5	6.3	82.2	9	7.6	8.5	7	7	84.1	10.7	8.9	83.7	10.2	8.4
Fermented dairy products		85	14.8	8.5	85.8	16.5	9.3	86.6	16.5	9.6	86.9	16.3	9.5	87.1	11.9	10.5	14.2	12.3	12.3	88.2	13.3	11.7	87.6	12.2	10.9
Total dairy products	High-fat	97.9	22	19.2	98.5	24.4	21.3	97.6	23.6	20.2	97.7	22.6	19.5	98.1	22.6	22.3	24.6	24.2	24.2	98.4	25.3	24.8	98	23.5	23.2
	Low-fat	99.7	5.1	4.8	99.1	5.4	5.2	98.5	5.7	5.3	100	6.3	6.1	98.8	5.9	5.9	7.2	7.1	7.1	99.2	7.3	7.3	98.5	7	6.9
	Total	99.8	24.3	24	99.8	26.6	26.4	100	26	25.6	100	25.8	25.6	99.7	28.2	28.2	31.3	31.3	31.3	99.9	32.1	32.1	99.9	30.1	30.1
Vitamin D (mg)																									
Milk	Full-fat	77	0.8	0.4	78.5	1.4	0.7	71.8	0.8	0.4	75.1	0.8	0.4	78.7	0.7	0.5	0.3	0.2	0.2	74.3	0.5	0.4	75.7	0.9	0.7
	Low-fat	99.5	4.8	4.5	99.1	4.4	4.2	97.9	5.6	5.2	100	4.8	4.6	98.3	4.7	4.6	3.9	3.8	3.8	98.8	4.8	4.7	97.6	4.2	4.1
	Total	99.8	5	4.8	99.8	4.9	4.9	99.4	5.8	5.6	100	5.1	5	99.1	5.1	5.1	4.1	4	4	99.5	5.1	5.1	98.9	4.9	4.9
Yoghurt	Full-fat	29.1	2.5	0.3	29.7	2.4	0.4	33.5	2	0.3	30.2	1.7	0.2	27.9	0.8	0.2	1.5	0.5	0.5	29.3	0.9	0.2	31	1.2	0.4
	Low-fat	28.1	1.3	0.2	28	1.2	0.2	31.6	1.1	0.2	31.3	3.8	0.6	29.1	2.1	0.6	1.4	0.5	0.5	32.4	1.7	0.5	32	1.8	0.5
	Total	44.8	2.2	0.5	45.1	2.2	0.5	49.5	1.9	0.5	48.1	3.3	0.8	44.7	1.8	0.8	2	0.9	0.9	48.1	1.6	0.8	47.7	1.9	0.9
Cheese	High-fat	75.2	6.1	2.7	77.8	6.8	3	78.3	7.3	3.2	77.4	6.5	2.9	78.8	4	3.3	4.2	3.3	3.3	81.4	4	3.3	79.8	3.4	2.8
	Low-fat	11.4	0.5	0	8.2	1	0	10.7	1.3	0	10.3	2.5	0.1	10.5	0.4	0	0.2	0	0	9.1	0.8	0.1	12.3	0.8	0.1
	Total	76.6	5.9	2.7	79.3	6.6	3	79.5	7.2	3.3	78.4	6.5	3	80	4	3.3	4.2	3.4	3.4	82.3	4	3.4	80.6	3.5	2.9
Cream		77.5	1.8	0.8	76.9	1.9	0.8	73.4	1.8	0.8	74.3	2.2	0.9	73.8	1.1	0.8	1.2	0.9	0.9	72.7	1.2	0.8	74.3	0.8	0.6
Butter		82.2	6.7	3.3	79.3	7	3.6	81.5	6.2	3.2	80.2	5.7	2.9	82.2	3.6	3	2.7	2.2	2.2	84.1	3.7	3.1	83.7	3.3	2.7
Fermented dairy products		85	5.5	3.2	85.8	6.3	3.5	86.6	6.4	3.7	86.9	6.6	3.8	87.1	4.6	4.1	5	4.3	4.3	88.2	4.7	4.1	87.6	4.3	3.8
Total dairy products	High-fat	97.9	8.2	7.2	98.5	9.3	8.1	97.6	8.8	7.5	97.7	8.2	7.1	98.1	7.7	7.6	6.8	6.7	6.7	98.4	7.6	7.5	98	6.9	6.8
	Low-fat	99.7	5.2	4.9	99.1	4.9	4.7	98.5	6.1	5.7	100	5.7	5.5	98.8	5.5	5.4	4.8	4.8	4.8	99.2	5.6	5.5	98.5	5.2	5.2
	Total	99.8	12.3	12.1	99.8	12.9	12.8	100	13.5	13.3	100	12.7	12.6	99.7	13.1	13	11.4	11.4	11.4	99.9	13.1	13.1	99.9	12	12
Vitamin B_{12} (μg)																									
Milk	Full-fat	77	14	7.1	78.5	11.9	6.1	71.8	12.6	5.9	75.1	10.7	4.9	78.7	7.2	5.7	8	6.1	6.1	74.3	8.2	6.1	75.7	8.2	6.4
	Low-fat	99.5	22.9	21.4	99.1	24.8	23.7	97.9	24.6	22.9	100	26.8	25.6	98.3	25	24.4	25.2	24.7	24.7	98.8	24.7	24.5	97.6	25.1	24.5
	Total	99.8	29.4	28.5	99.8	30.2	29.7	99.4	30	28.9	100	31.4	30.6	99.1	30.4	30.1	31.1	30.8	30.8	99.5	30.7	30.6	98.9	31.2	30.9
Yoghurt	Full-fat	29.1	4	0.5	29.7	5.2	0.8	33.5	4.5	0.7	30.2	4.5	0.6	27.9	1.7	0.5	2.5	0.8	0.8	29.3	1.5	0.4	31	2.2	0.8
	Low-fat	28.1	6.4	0.8	28	5.7	0.7	31.6	6	0.9	31.3	6.2	0.9	29.1	3	0.8	2.4	0.8	0.8	32.4	3.1	1	32	3.6	1
	Total	44.8	6.1	1.3	45.1	6.6	1.5	49.5	6.2	1.6	48.1	6.4	1.6	44.7	3	1.4	3.2	1.5	1.5	48.1	3	1.4	47.7	3.7	1.8
Cheese	High-fat	75.2	19.6	8.6	77.8	20.4	9	78.3	20.6	9.1	77.4	18.6	8.4	78.8	10.6	8.6	10.2	8.1	8.1	81.4	10.5	8.7	79.8	9.9	8.1
	Low-fat	11.4	7.8	0.3	8.2	6.7	0.2	10.7	4.6	0.2	10.3	6.1	0.2	10.5	2	0.2	2.4	0.3	0.3	9.1	3.9	0.4	12.3	1.6	0.2
	Total	76.6	19.5	8.9	79.3	20.2	9.2	79.5	20.3	9.3	78.4	18.6	8.6	80	10.7	8.8	10.4	8.3	8.3	82.3	10.9	9	80.6	10.1	8.3
Cream		77.5	1	0.4	76.9	1.2	0.5	73.4	1.1	0.4	74.3	1.1	0.5	73.8	0.5	0.4	0.7	0.5	0.5	72.7	0.5	0.3	74.3	0.5	0.4
Butter		82.2	0.5	0.2	79.3	0.5	0.3	81.5	0.5	0.3	80.2	0.6	0.3	82.2	0.4	0.3	0.3	0.3	0.3	84.1	0.5	0.4	83.7	0.5	0.4
Fermented dairy products		85	17.8	10.2	85.8	19	10.7	86.6	18.6	10.8	86.9	17.5	10.2	87.1	11.4	10.1	11.4	9.9	9.9	88.2	11.8	10.4	87.6	11.3	10.1
Total dairy products	High-fat	97.9	18.8	16.4	98.5	18.1	15.8	97.6	18.4	15.7	97.7	16.3	14.1	98.1	15.3	15	15.2	14.9	14.9	98.4	15.7	15.4	98	15.4	15.3
	Low-fat	99.7	24.2	23	99.1	26.4	25.4	98.5	26.1	24.7	100	28.4	27.4	98.8	26.2	25.9	27	26.5	26.5	99.2	26.4	26.3	98.5	26.9	26.6
	Total	99.8	39.9	39.4	99.8	41.5	41.2	100	41.1	40.4	100	42	41.5	99.7	41.1	41	41.4	41.4	41.4	99.9	41.7	41.7	99.9	41.8	41.8

Table 2.5 (continued)

		2008/2009			2009/2010			2010/2011			2011/2012			2012/2013			2013/2014			2014/2015			2015/2016		
N			612			536			531			643			695			724			740			748	
Dairy products (g/d)		% cons †	% within cons ‡	% total §	% cons †	% within cons ‡	% total §	% cons †	% within cons ‡	% total §	% cons †	% within cons ‡	% total §	% cons †	% within cons ‡	% total §	% cons †	% within cons ‡	% total §	% cons †	% within cons ‡	% total §	% cons †	% within cons ‡	% total §
Calcium (mg)																									
Milk	Full-fat	77	14.7	7.5	78.5	13.4	6.8	71.8	13.2	6.2	75.1	11.2	5.2	78.7	7.4	5.8	78	8.3	6.4	74.3	8.3	6.1	75.7	8.1	6.3
	Low-fat	99.5	25.4	23.7	99.1	26.7	25.5	97.9	25.2	23.5	100	27.4	26.3	98.3	26.4	25.8	98.3	26.5	25.9	98.8	25.5	25.3	97.6	25.8	25.2
	Total	99.8	32.2	31.2	99.8	32.9	32.4	99.4	30.9	29.7	100	32.3	31.5	99.1	31.9	31.6	99.3	32.6	32.3	99.5	31.5	31.4	98.9	31.9	31.6
Yoghurt	Full-fat	29.1	12.4	1.6	29.7	11	1.6	33.5	12.2	1.9	30.2	11.3	1.6	27.9	5.4	1.6	29.1	6	1.9	29.3	4.7	1.3	31	6.2	2.2
	Low-fat	28.1	18.6	2.3	28	16.3	2.1	31.6	16.3	2.5	31.3	15.8	2.4	29.1	8.7	2.4	32.3	7	2.2	32.4	7.9	2.4	32	9	2.6
	Total	44.8	18.2	3.9	45.1	16.4	3.7	49.5	17	4.3	48.1	16.2	4	44.7	8.9	4.1	47.7	8.6	4.1	48.1	8.1	3.7	47.7	9.8	4.7
Cheese	High-fat	75.2	30.4	13.4	77.8	30.6	13.4	78.3	30.7	13.5	77.4	28.7	12.9	78.8	17	13.7	78.9	16.2	12.8	81.4	16.3	13.4	79.8	15.2	12.4
	Low-fat	11.4	11.5	0.5	8.2	9.8	0.3	10.7	6.1	0.2	10.3	8	0.3	10.5	2.7	0.2	11.2	2.8	0.3	9.1	4.5	0.4	12.3	2.5	0.3
	Total	76.6	30.3	13.9	79.3	30.2	13.7	79.5	30.2	13.8	78.4	28.6	13.1	80	17.1	13.9	80.5	16.4	13.1	82.3	16.7	13.9	80.6	15.4	12.7
Cream		77.5	1	0.4	76.9	1.1	0.5	73.4	1	0.4	74.3	1.1	0.4	73.8	0.6	0.4	73.5	0.6	0.5	72.7	0.6	0.4	74.3	0.6	0.4
Butter		82.2	0.3	0.1	79.3	0.3	0.2	81.5	0.3	0.2	80.2	0.3	0.2	82.2	0.2	0.2	82.2	0.2	0.1	84.1	0.2	0.2	83.7	0.2	0.2
Fermented dairy products		85	30.9	17.7	85.8	31.1	17.5	86.6	31	18.1	86.9	29.6	17.2	87.1	20.3	18	86.9	19.9	17.2	88.2	20	17.6	87.6	19.5	17.4
Total dairy products	High-fat	97.9	24.5	21.4	98.5	24	20.9	97.6	23.7	20.4	97.7	21.6	18.7	98.1	20.5	20.1	97.8	20.1	19.8	98.4	20.5	20.1	98	19.5	19.3
	Low-fat	99.7	29.5	28.1	99.1	30.8	29.6	98.5	29.7	28	100	31.6	30.6	98.8	30.4	30.1	98.6	30.9	30.3	99.2	29.6	29.4	98.5	30.7	30.3
	Total	99.8	50	49.5	99.8	50.8	50.5	100	49.1	48.4	100	49.7	49.2	99.7	50.4	50.2	99.9	50.1	50.1	99.9	49.6	49.5	99.9	49.6	49.6
Potassium (mg)																									
Milk	Full-fat	77	5.8	3	78.5	5.5	2.8	71.8	5.2	2.5	75.1	4.4	2	78.7	2.7	2.1	78	3.3	2.5	74.3	3.3	2.4	75.7	3.1	2.4
	Low-fat	99.5	9.9	9.3	99.1	10.9	10.4	97.9	10	9.3	100	10.7	10.3	98.3	10.6	10.4	98.3	10.6	10.4	98.8	10	9.9	97.6	10.2	9.9
	Total	99.8	12.6	12.2	99.8	13.5	13.3	99.4	12.2	11.8	100	12.6	12.3	99.1	12.6	12.5	99.3	13	12.9	99.5	12.4	12.4	98.9	12.5	12.3
Yoghurt	Full-fat	29.1	4.9	0.6	29.7	4.8	0.7	33.5	5.3	0.8	30.2	4.3	0.6	27.9	2.2	0.7	29.1	2.4	0.7	29.3	1.8	0.5	31	2.4	0.9
	Low-fat	28.1	7.9	1	28	6.9	0.9	31.6	6.9	1.1	31.3	6.8	1	29.1	3.7	1	32.3	2.9	0.9	32.4	3.3	1	32	3.7	1.1
	Total	44.8	7.5	1.6	45.1	7	1.6	49.5	7.3	1.9	48.1	6.6	1.6	44.7	3.7	1.7	47.7	3.5	1.7	48.1	3.3	1.5	47.7	3.9	1.9
Cheese	High-fat	75.2	1.2	0.5	77.8	1.3	0.6	78.3	1.3	0.6	77.4	1.2	0.5	78.8	0.7	0.6	78.9	0.7	0.5	81.4	0.7	0.6	79.8	0.6	0.5
	Low-fat	11.4	2.3	0.1	8.2	2.7	0.1	10.7	1.6	0.1	10.3	1.7	0.1	10.5	0.7	0.1	11.2	0.9	0.1	9.1	0.8	0.1	12.3	0.6	0.1
	Total	76.6	1.3	0.6	79.3	1.5	0.7	79.5	1.4	0.6	78.4	1.3	0.6	80	0.8	0.6	80.5	0.8	0.6	82.3	0.8	0.6	80.6	0.7	0.6
Cream		77.5	0.4	0.2	76.9	0.5	0.2	73.4	0.4	0.2	74.3	0.4	0.2	73.8	0.2	0.2	73.5	0.3	0.2	72.7	0.2	0.2	74.3	0.2	0.2
Butter		82.2	0.1	0.1	79.3	0.1	0.1	81.5	0.1	0.1	80.2	0.1	0.1	82.2	0.1	0.1	82.2	0.1	0.1	84.1	0.1	0.1	83.7	0.1	0.1
Fermented dairy products		85	3.9	2.2	85.8	4	2.3	86.6	4.3	2.5	86.9	3.8	2.2	87.1	2.7	2.4	86.9	2.7	2.3	88.2	2.4	2.1	87.6	2.8	2.5
Total dairy products	High-fat	97.9	4.3	3.7	98.5	4.2	3.7	97.6	3.8	3.3	97.7	3.2	2.8	98.1	3	3	97.8	3.4	3.3	98.4	3.3	3.2	98	3.2	3.2
	Low-fat	99.7	11.6	11	99.1	12.6	12.1	98.5	11.9	11.2	100	12.4	12	98.8	12.3	12.1	98.6	12.4	12.1	99.2	11.6	11.5	98.5	12.1	11.9
	Total	99.8	14.9	14.7	99.8	15.9	15.8	100	14.7	14.5	100	14.9	14.7	99.7	15.1	15.1	99.9	15.4	15.4	99.9	14.7	14.7	99.9	15.1	15.1
Magnesium (mg)																									
Milk	Full-fat	77	4.3	2.2	78.5	4	2	71.8	3.7	1.8	75.1	3	1.4	78.7	1.8	1.5	78	2.2	1.7	74.3	2.2	1.7	75.7	2.1	1.7
	Low-fat	99.5	7.5	7	99.1	8.3	7.9	97.9	7.4	6.9	100	8.1	7.7	98.3	7.9	7.7	98.3	7.9	7.7	98.8	7.4	7.4	97.6	7.6	7.4
	Total	99.8	9.5	9.2	99.8	10.1	9.9	99.4	9	8.7	100	9.4	9.1	99.1	9.2	9.2	99.3	9.5	9.4	99.5	9.1	9	98.9	9.2	9.1
Yoghurt	Full-fat	29.1	3.9	0.5	29.7	3.7	0.5	33.5	4.2	0.6	30.2	3.3	0.5	27.9	1.6	0.5	29.1	1.9	0.6	29.3	1.4	0.4	31	1.8	0.7
	Low-fat	28.1	6	0.8	28	5.4	0.7	31.6	5.2	0.8	31.3	5.1	0.8	29.1	2.9	0.8	32.3	2.2	0.7	32.4	2.5	0.8	32	2.9	0.8
	Total	44.8	5.8	1.2	45.1	5.4	1.2	49.5	5.6	1.4	48.1	5	1.3	44.7	2.8	1.3	47.7	2.7	1.3	48.1	2.5	1.1	47.7	3	1.5
Cheese	High-fat	75.2	4.4	1.9	77.8	4.8	2.1	78.3	4.4	2	77.4	4.2	1.9	78.8	2.5	2	78.9	2.3	1.8	81.4	2.4	2	79.8	2.2	1.8
	Low-fat	11.4	2.6	0.1	8.2	2.6	0.1	10.7	1.7	0.1	10.3	1.8	0.1	10.5	0.7	0.1	11.2	0.9	0.1	9.1	0.8	0.1	12.3	0.6	0.1
	Total	76.6	4.4	2	79.3	4.8	2.2	79.5	4.4	2	78.4	4.2	1.9	80	2.6	2.1	80.5	2.4	1.9	82.3	2.4	2	80.6	2.2	1.8
Cream		77.5	0.3	0.1	76.9	0.4	0.2	73.4	0.3	0.1	74.3	0.3	0.1	73.8	0.2	0.1	73.5	0.2	0.1	72.7	0.2	0.1	74.3	0.2	0.1
Butter		82.2	0.1	0	79.3	0.1	0.1	81.5	0.1	0	80.2	0.1	0	82.2	0	0	82.2	0	0	84.1	0	0	83.7	0	0
Fermented dairy products		85	5.7	3.3	85.8	6.1	3.4	86.6	5.9	3.4	86.9	5.5	3.2	87.1	3.8	3.4	86.9	3.7	3.2	88.2	3.6	3.2	87.6	3.7	3.3
Total dairy products	High-fat	97.9	4.9	4.3	98.5	5	4.3	97.6	4.5	3.9	97.7	4	3.4	98.1	3.7	3.7	97.8	3.8	3.7	98.4	3.8	3.8	98	3.6	3.6
	Low-fat	99.7	8.8	8.4	99.1	9.6	9.2	98.5	8.9	8.4	100	9.4	9	98.8	9.1	9.1	98.6	9.2	9.1	99.2	8.6	8.6	98.5	9.1	9
	Total	99.8	12.8	12.7	99.8	13.7	13.6	100	12.5	12.3	100	12.6	12.5	99.7	12.8	12.7	99.9	12.8	12.8	99.9	12.4	12.3	99.9	12.6	12.6

Table 2.5 (continued)

		2008/2009			2009/2010			2010/2011			2011/2012			2012/2013			2013/2014			2014/2015			2015/2016		
N			612			536			531			643			695			724			740			748	
Dairy products (g/d)		**% cons †**	**% within cons ‡**	**% total §**	**% cons †**	**% within cons ‡**	**% total §**	**% cons †**	**% within cons ‡**	**% total §**	**% cons †**	**% within cons ‡**	**% total §**	**% cons †**	**% within cons ‡**	**% total §**	**% cons †**	**% within cons ‡**	**% total §**	**% cons †**	**% within cons ‡**	**% total §**	**% cons †**	**% within cons ‡**	**% total §**
Phosphorus (mg)																									
Milk	Full-fat	77	7.8	4	78.5	7.1	3.6	71.8	7	3.3	75.1	5.8	2.7	78.7	3.7	2.9	78	4.3	3.3	74.3	4.4	3.2	75.7	4.2	3.3
	Low-fat	99.5	14	13.1	99.1	15.3	14.6	97.9	13.9	12.9	100	15.4	14.8	98.3	14.7	14.4	98.3	14.8	14.4	98.8	13.8	13.7	97.6	14.4	14.1
	Total	99.8	17.6	17.1	99.8	18.5	18.2	99.4	16.9	16.2	100	17.9	17.5	99.1	17.4	17.3	99.3	17.9	17.8	99.5	17	16.9	98.9	17.5	17.3
Yoghurt	Full-fat	29.1	7.4	0.9	29.7	6.9	1	33.5	7.7	1.2	30.2	6.6	0.9	27.9	3.3	1	29.1	3.6	1.1	29.3	2.9	0.8	31	3.8	1.4
	Low-fat	28.1	10.9	1.4	28	9.7	1.3	31.6	9.4	1.4	31.3	9.5	1.4	29.1	5.4	1.5	32.3	4.3	1.3	32.4	4.8	1.5	32	5.7	1.6
	Total	44.8	10.7	2.3	45.1	9.9	2.3	49.5	10.2	2.6	48.1	9.6	2.4	44.7	5.5	2.5	47.7	5.2	2.5	48.1	4.9	2.3	47.7	6.2	3
Cheese	High-fat	75.2	14.4	6.4	77.8	15	6.6	78.3	14.2	6.3	77.4	13.9	6.2	78.8	8.7	7	78.9	8.3	6.5	81.4	8.3	6.8	79.8	7.8	6.3
	Low-fat	11.4	6.8	0.3	8.2	6.9	0.2	10.7	4.4	0.2	10.3	5.6	0.2	10.5	2	0.2	11.2	2.4	0.3	9.1	2.5	0.2	12.3	1.7	0.2
	Total	76.6	14.5	6.6	79.3	15	6.8	79.5	14.2	6.4	78.4	14	6.4	80	8.8	7.2	80.5	8.5	6.8	82.3	8.5	7	80.6	8	6.6
Cream		77.5	0.7	0.3	76.9	0.8	0.3	73.4	0.7	0.3	74.3	0.7	0.3	73.8	0.4	0.3	73.5	0.4	0.3	72.7	0.4	0.3	74.3	0.4	0.3
Butter		82.2	0.2	0.1	79.3	0.3	0.2	81.5	0.3	0.1	80.2	0.2	0.1	82.2	0.2	0.1	82.2	0.1	0.1	84.1	0.2	0.1	83.7	0.2	0.1
Fermented dairy products		85	15.6	8.9	85.8	16.2	9.1	86.6	15.5	9	86.9	15.2	8.8	87.1	10.9	9.7	86.9	10.7	9.3	88.2	10.6	9.3	87.6	10.7	9.5
Total dairy products	High-fat	97.9	12.3	10.8	98.5	12.3	10.7	97.6	11.7	10	97.7	10.8	9.3	98.1	10.5	10.3	97.8	10.4	10.3	98.4	10.6	10.4	98	10.1	10
	Low-fat	99.7	16.5	15.7	99.1	17.8	17	98.5	16.6	15.7	100	17.9	17.3	98.8	17.2	17.1	98.6	17.5	17.2	99.2	16.3	16.2	98.5	17.5	17.3
	Total	99.8	26.7	26.4	99.8	28	27.8	100	26.1	25.7	100	27	26.7	99.7	27.5	27.4	99.9	27.5	27.5	99.9	26.6	26.6	99.9	27.3	27.3
Zinc (mg)																									
Milk	Full-fat	77	5.6	2.9	78.5	5.1	2.6	71.8	5.1	2.4	75.1	4.3	2	78.7	2.7	2.1	78	3.2	2.4	74.3	3.2	2.3	75.7	3	2.4
	Low-fat	99.5	8.8	8.2	99.1	9.6	9.2	97.9	8.6	8	100	9.7	9.3	98.3	9.4	9.1	98.3	9.6	9.4	98.8	9.1	9	97.6	9.4	9.2
	Total	99.8	11.4	11	99.8	11.9	11.8	99.4	10.8	10.4	100	11.6	11.3	99.1	11.3	11.2	99.3	11.9	11.8	99.5	11.4	11.3	98.9	11.6	11.5
Yoghurt	Full-fat	29.1	4.3	0.5	29.7	4.1	0.6	33.5	4.4	0.7	30.2	3.8	0.5	27.9	1.9	0.6	29.1	2.2	0.7	29.3	1.6	0.4	31	2.2	0.8
	Low-fat	28.1	6.8	0.8	28	6.3	0.8	31.6	6.1	0.9	31.3	6.2	0.9	29.1	3.4	0.9	32.3	2.7	0.9	32.4	3	0.9	32	3.5	1
	Total	44.8	6.5	1.4	45.1	6.2	1.4	49.5	6.3	1.6	48.1	6	1.5	44.7	3.3	1.5	47.7	3.3	1.5	48.1	3	1.4	47.7	3.6	1.8
Cheese	High-fat	75.2	17.1	7.5	77.8	18.1	7.9	78.3	17.1	7.5	77.4	16.6	7.5	78.8	9.8	7.9	78.9	9.4	7.4	81.4	9.2	7.5	79.8	8.7	7.1
	Low-fat	11.4	5.6	0.2	8.2	5.5	0.2	10.7	3.6	0.1	10.3	4.4	0.1	10.5	1.7	0.1	11.2	1.9	0.2	9.1	2	0.2	12.3	1.3	0.2
	Total	76.6	16.9	7.8	79.3	17.8	8.1	79.5	16.9	7.7	78.4	16.6	7.6	80	9.9	8	80.5	9.6	7.7	82.3	9.3	7.7	80.6	8.8	7.2
Cream		77.5	0.4	0.2	76.9	0.4	0.2	73.4	0.3	0.1	74.3	0.4	0.2	73.8	0.2	0.2	73.5	0.2	0.2	72.7	0.2	0.1	74.3	0.2	0.1
Butter		82.2	0	0	79.3	0	0	81.5	0	0	80.2	0	0	82.2	0	0	82.2	0	0	84.1	0	0	83.7	0	0
Fermented dairy products		85	15.9	9.1	85.8	16.9	9.5	86.6	15.9	9.3	86.9	15.7	9.1	87.1	10.8	9.6	86.9	10.6	9.2	88.2	10.4	9.1	87.6	10.1	9
Total dairy products	High-fat	97.9	12.1	10.5	98.5	12.3	10.7	97.6	11.8	10.1	97.7	11.1	9.6	98.1	10.3	10.2	97.8	10.2	10	98.4	10.2	10	98	9.7	9.6
	Low-fat	99.7	10.3	9.8	99.1	11.2	10.7	98.5	10.3	9.7	100	11.3	10.9	98.8	10.9	10.8	98.6	11.3	11.1	99.2	10.6	10.5	98.5	11.2	11.1
	Total	99.8	20.6	20.4	99.8	21.6	21.5	100	20.1	19.8	100	20.7	20.5	99.7	21	21	99.9	21.2	21.2	99.9	20.6	20.6	99.9	20.7	20.7

†Percentage of consumers
‡Percent contribution of dairy products to the nutrient intake within dairy consumers
§Percent contribution of dairy products to the nutrient intake in the total sample

2.6 Discussion

2.6.1 Summary of results and comments

Dietary consumption data at a national level in the UK with the NDNS has enabled the examination of trends in dairy consumption over an 8-year period from 2008 to 2016 and comparison with the older surveys from 1994/1995 and 2000/2001. There have been notable shifts in dairy consumption and some clear patterns have emerged.

Overall, milk was the highest contributor to total dairy consumption consistently over time constituting about 80% of total dairy products in adults <65 years and about 85% in elderly adults. It is noted that low-fat milk consumption was consistently higher than full-fat milk over time contributing approximately 80% to the total milk consumption among adults <65 years. Commensurate with this, full-fat milk consumption continued the decreasing trend observed in earlier years through food supply and availability data[38], dropping by 8.6 g/day among adults <65 years and by 81.1 g/day among elderly people over the 8-year period of 2008-2016. The percentage of elderly consumers of full-fat milk dropped by 14.4% over the 22-year period of 1994-2016. In contrast, the percentage of elderly consumers of total, full-fat and low-fat yoghurt increased by 26.9%, 16% and 20.9% respectively over the same 22-year period. Smaller increases were observed for the percentages of high-fat cheese consumers and smaller decreases for the percentages of cream consumers in both age groups. Differences of consumption levels between men and women were small.

Total dairy consumption contributed on average approximately 14% to total energy intake, 25% to total fat, 5.5% to carbohydrate, and 18% to protein. Of the macronutrient subtypes, total dairy consumption contributed 41% to saturated fat, 17% to monounsaturated fat, 55% to trans fat, and 13% to sugars.

Of the micronutrients, total dairy consumption contributed approximately 25% to vitamin A, 10% to vitamin D, 40% to vitamin B_{12}, 50% to calcium, 15% to potassium, 13% to magnesium, 27% to phosphorus, and 21% to zinc.

The highest dairy contributor to total fat and its types was cheese followed by butter. For carbohydrates, sugars and protein, the highest contributor was milk. For micronutrients, the highest dairy contributor was low-fat milk except for vitamin A, for which the highest dairy contributor was high-fat cheese.

The highest dairy contribution to macronutrient intakes was observed for saturated fat (approximately 41%) and for trans fat (approximately 50%), which are the main macronutrients of concern for cardio-metabolic risk (as appraised in section 1.3). The contribution of dairy products to saturated fat intake is in agreement with that reported from the TRANSFAIR study, which consisted of chemical analysis of approximately 100 foods from 14 European countries and also from the use of purchase data or consumption data (individual or household level) to derive food contributions to nutrient intakes[55].

Specifically for the UK, this study used household level consumption data from the National Food Survey 1997 and reported a contribution of dairy products to saturated fat intake of 38.8%[55]. However, for trans fat, dairy consumption contributed 24.7%[55], while in our analysis we observed a much greater (2-fold higher) contribution of approximately 50%, which actually increased from 51% contribution to trans fat intake in 2008/2009 to 62.1% in 2015/2016. It can be assumed that the increased contribution of dairy to trans fat intake in the recent years, as reported in this PhD, is a consequence of the voluntary decrease of the industrial trans fat content by the food industry in the UK because of the reported harmful effects of industrial trans fat on health[56], while at the same time the ruminant trans fat intake might not have considerably changed. Indeed, in the most recent report on trans fat content of foods from the Department of Health, it was observed that the industrial trans fat content of foods was significantly decreased compared to that observed in foods 20 or 30 years ago[57].

It is of note that some descriptive statistics reported in the official reports of NDNS differ from the statistics I reported here. For example, the contribution of milk and milk products, as well as butter to saturated fat intake of people 19-64 years was reported to be 27% previously[53], while this analysis reported 41% and the numbers were similar both among consumers only and in the total sample. The biggest difference was for butter, for which prior report was of a 5% contribution, while I reported a 15% contribution to saturated fat intake. Such difference is expected if we consider that we estimated dairy consumption also accounting for dairy content in composite foods and recipes. Butter is an ingredient used very often in recipes, so whether we account for recipes or not should make a large difference in the estimation of butter consumption levels.

2.6.2 Strengths and limitations

This descriptive study has several strengths. First, we used data from NDNS, which includes a representative sample from the UK, so that we can generalise the results of our study to a national level with confidence. Since the start of the rolling programme in 2008, there are annual data available, which enabled us to look at consumption trends over time. We further contributed to the precision of the consumption levels by incorporating in our analysis dairy products consumed as part of composite foods and recipes. This was an enhancement and an important contribution, since it made a big difference to nutrients such as saturated fat, for which butter is a high contributor, but also a frequent ingredient in recipes.

However, this PhD analysis has also some limitations. The rolling programme started in 2008, so we could only investigate trends over 8 years, while it is of greater interest to investigate trends over decades. Even the availability of the old survey data did not allow for the investigation of trends in consumption patterns, because of the different methodology used (weighed vs estimated food diary), but it only allowed the description of trends of

the percentage of dairy consumed, which should be equally captured irrespective of the dietary assessment method used. As a result, trends of dairy consumption for years earlier than the NDNS rolling programme in the UK are available only as an approximation from household level consumption data or national level dairy supply/availability/production data.

2.6.3 Conclusion

The healthfulness of yoghurt has been consistently supported from previous evidence, whereas the saturated fat content of full-fat milk has been used as the reason to recommend lower consumption of full-fat milk and a preference towards low-fat milk and dairy products in general[23]. These recommendations are reflected in the consumption levels and trends observed in our study from a representative UK sample with a continuous decrease over time in the consumption of full-fat milk and an increase in the consumption of yoghurt. However, evidence regarding saturated fat intake from specific food sources such as dairy products has started to change.

The importance and variability of the dairy food matrix is evident when considering the contribution of dairy products to nutrient intakes. Dairy products contribute more than a quarter to vitamin A, vitamin B_{12}, calcium, and phosphorus intakes, which are important nutrients for health, but also to saturated and trans fat intakes, which have been linked to potential harmful effects on health. The balance of effects from these nutrients is what defines the final effect of dairy products as food entities on health. Research into the mechanisms of action of the dairy food matrix on health in combination with close monitoring of consumption levels in the population will contribute to informed formulation of dietary guidelines and design of effective public health interventions.

Chapter 3

Previous evidence on dairy consumption and cardio-metabolic markers

Summary

The differential associations observed between different dairy types and cardio-metabolic disease outcomes (section 1.2) pose the question as to what the underlying pathways for such associations are, which might further explain the observed heterogeneity. Biological pathways to cardio-metabolic disease might be related to a number of factors including adiposity, lipidaemia, glycaemia, inflammation, hepatic function and blood pressure.

Evidence from randomised controlled trials (RCTs) supports favourable effects of total dairy consumption on body weight and composition under conditions of energy restriction. Specifically, total dairy consumption decreased body weight, total fat mass, waist circumference, abdominal fat and visceral adipose tissue and increased lean mass in RCTs. Results on specific dairy types are limited.

Concerning lipid markers, there is established evidence from RCTs that butter increases total, low-density lipoprotein (LDL-C) and high-density lipoprotein cholesterol (HDL-C) when compared with other types of fat, whereas null effects on blood lipids were reported for total, high-fat or low-fat dairy products. Cheese decreased LDL-C and HDL-C when compared with butter. Evidence on other dairy types is sparse.

RCTs have failed to report any effects of total, low-fat or high-fat dairy products on glycaemic markers including fasting blood glucose, insulin, haemoglobin A1c (HbA1c) or the homoestasis model assessment for insulin resistance (HOMA-IR). As for other markers, evidence on specific dairy types is sparse.

Fermented dairy consumption decreased inflammatory markers overall in RCTs, but low-fat or high-fat dairy consumption did not have an effect on C-reactive protein (CRP). The number of studies on specific dairy types is limited, and there are no studies on associations between dairy consumption and markers of hepatic function.

Finally, evidence from RCTs indicates a beneficial effect of fermented milk on systolic (SBP) and diastolic (DBP) blood pressure, but no effects for total, low-fat or high-fat dairy products. Milk, total and low-fat dairy products have been associated with a lower risk of hypertension in prospective cohort studies.

Overall, although there are many studies on associations of dairy consumption with cardio-metabolic risk markers, the evidence is not sufficient or consistent to draw conclusions on associations of different dairy types with a spectrum of cardio-metabolic markers. The inconsistency in the definition of dairy products as a food group and the limitations of the studies that have been conducted so far suggest the need for more research to elucidate the link between dairy consumption and cardio-metabolic disease.

Cardio-metabolic diseases are multifactorial and thus the investigation of biological pathways including adiposity, glycaemia, lipidaemia, inflammation, hepatic function and blood pressure could elucidate the role of dairy products in their pathogenesis. There is an extensive literature on associations of total and types of dairy consumption with several cardio-metabolic markers. However, the evidence synthesis is not very straightforward because different studies have investigated associations of different dairy types or groups with different markers resulting in sparse evidence on associations between certain dairy types and certain cardio-metabolic markers. In this chapter, evidence on associations of total and types of dairy products with cardio-metabolic markers from randomised controlled trials (RCTs) and prospective cohort studies is presented.

3.1 Markers of body weight and composition

3.1.1 Body weight

Randomised controlled trials

Meta-analyses of RCTs show differential effects of dairy products on body weight depending on whether the interventions applied energy restriction or not[58]. After I searched meta-analyses on associations of dairy consumption with cardio-metabolic markers, I summarised the results for total, low- and high-fat dairy products in **Table 3.1**.

A meta-analysis of 37 RCTs reported an overall null effect of a mean 2.6$\pm$1 servings/day of total dairy consumption on body weight during a mean 7.7$\pm$7.9 months of intervention[58]. After stratification by energy restriction status, the same meta-analysis reported an increasing effect of total dairy consumption on body weight in trials without energy restriction [b=0.36 kg (0.01, 0.7)], but a decreasing effect in trials with energy restriction of 500 kcal/day on average [-0.64 kg (-1.05, -0.24)] in the intervention groups compared to the control[58].

The distinction of trials based on the application of energy restriction or not is not sufficient to describe the independent effects of dairy consumption on body weight. Instead, it is important to know whether an experiment is isocaloric, thus whether the intervention and control foods contain the same amount of energy in order to distinguish the effect of energy from the effect of dairy consumption. During no energy restriction, it is even more important to address this issue, because if the intervention is compared with the habitual diet, then we have increased energy intake in the intervention group, so the increasing effects of the intervention on body weight might be the result of the increase in energy intake rather than a net effect of dairy products. A meta-analysis of 27 RCTs, which included only RCTs with energy restriction, identified and reported that the allocation of intervention and control foods was isocaloric, so it seems reasonable to assume that trials with energy restriction have applied isocaloric interventions to the two trial arms[59]. Of the

27 RCTs included in this meta-analysis, 15 overlapped with the more recent meta-analysis reported above[58] and of the remaining 12 RCTs, six allocated an intervention of casein or whey protein and five allocated an intervention of total (n=2) or types (n=4) of dairy products[59]. This meta-analysis also reported a decreasing effect of dairy consumption on body weight with a larger magnitude [-0.92 kg (-1.62, -0.2)][59]. The effect was even more pronounced among women [-1.16 kg (-1.66, -0.66)], but it was attenuated to null in case of mutli-component interventions that included resistance training[59]. Likewise, older meta-analyses on the effect of dairy consumption on body weight with most of the included studies overlapping with the aforementioned meta-analyses reported null effects when examining trials with and without energy restriction together[60–62]. There were weight decreasing effects when examining trials with energy restriction only[61, 62].

Under no energy restriction, both low- and high-fat dairy products increased body weight in another meta-analysis of RCTs[63]. Meta-analyses on types of dairy products are not available, because most of the interventions included total dairy products, often giving the option to participants to select the type of preference based on a specific calcium intake that they aimed to achieve and usually suggesting a minimum amount of milk to be consumed[58]. Some interventions included milk in different forms (fluid or powder), of different fat content or of different processing (whey protein or casein mixture)[58, 59].

Prospective cohort studies

Since the average duration of RCTs is limited to approximately 8 months, evidence from prospective cohort studies can inform on longer term associations between dairy consumption and body weight or BMI. The majority of the prospective cohort studies have examined body weight rather than body mass index (BMI) in real-life settings. Prospective cohort studies have reported null[64–66], positive[66] or inverse[67–69] associations between total dairy consumption and body weight after adjustment for total energy intake. It can be assumed that adjustment for energy intake in observational studies is equivalent to the application of an isocaloric experiment in RCTs to assess the association of dairy consumption with body weight independent of total energy intake. However, in observational studies, we should also accommodate the possibility of residual confounding due to the use of the subjective measure of dietary assessment.

Similar results have been reported for high-fat dairy consumption with null[67, 70], positive[66] or inverse[66, 69] associations with body weight. For low-fat dairy products null associations with body weight have been more consistently reported[66, 67, 69–71]. For the main dairy types, prospective analyses resulted in either null (milk[67, 70, 72], yoghurt[73], cheese[65, 67, 70, 71]) or inverse (milk[65, 72], yoghurt[65, 67, 70], cheese[68, 72]) associations. The heterogeneity of these results from the prospective cohort studies will be further explored in Chapter 5. After a literature search, I have summarised the magnitude and variance of these associations in **Table 3.2**.

3.1.2 Body composition

Randomised controlled trials

Some of the studies investigating associations between dairy consumption and body weight, also examined associations with measures of body composition (Table 3.1). Five meta-analyses of RCTs showed that total dairy consumption decreased body fat mass in energy restriction trials[58–62] and two of them showed null effects in trials with no energy restriction[58, 62]. For example, the meta-analysis that included the most RCTs (n=27) reported a decrease of 1.24 kg in body fat mass after an intervention with a mean of 2.7±1 dairy servings/day during 5.1±3.1 months[59].

Three meta-analyses of RCTs reported that total dairy consumption increased lean mass in energy restriction trials[58, 59, 61], although in one of them the effect was not significant[58], but dairy consumption did not change lean mass in trials with no energy restriction as reported by one of the meta-analyses[58]. For example, the meta-analysis with the highest number of RCTs (n=27), reported that an average of 2.9 dairy servings/day increased body lean mass by 0.36 kg over an average period of 4.2 months[59].

Measures of body fat and lean mass distribution are peripheral fat and appendicular lean mass. Since the association between abdominal fat and cardio-metabolic disease has been shown to be even stronger, compartments of abdominal fat including visceral adipose tissue (VAT) and subcutaneous adipose tissue (SCAT) are also measured[74]. VAT or intraperitoneal fat is the part of the adipose tissue in the abdominal area, which surrounds the organs, and SCAT is the part of the adipose tissue, which is accumulated under the skin[74]. The equipment to measure abdominal fat is often expensive (computer tomography, magnetic resonance imaging, dual energy X-ray absorptiometry), so most of the studies measure waist circumference (sometimes also the ratio of waist to hip circumference) as a proxy for fat mass distribution[74]. Meta-analyses of RCTs have reported decreasing effects of total dairy consumption on waist circumference in trials with energy restriction[58, 61], but null effects in trials without energy restriction[58]. In accordance with the effects on body fat mass reported from the meta-analyses, one 6-month and one 12-month cross-over RCT, which did not apply energy restriction, reported null effects of four low-fat dairy servings on abdominal fat when compared with one or two servings[75, 76]. Conversely, in two trials with energy restriction dairy consumption decreased abdominal fat[77, 78], whereas in one trial with energy restriction no effect was observed[79]. One of the trials, which examined VAT, also observed a decrease with higher dairy consumption[78].

Prospective cohort studies

Evidence on associations of dairy products with waist circumference from prospective cohort studies, after adjustment for total energy intake[64, 65, 71], BMI[68, 80, 81] or both[67] shows null associations for cheese[65, 67, 68, 71], low-fat[71, 80] and total dairy products[64, 67, 68], but inverse associations for yoghurt[65, 67] and high-fat dairy

products[67, 80]. In addition, a prospective observational analysis of data from the PREDIMED trial (3.2 years of follow-up) reported null associations with central adiposity for all dairy types and groups studied apart from yoghurt and its sub-types, for which inverse associations were found[82].

Table 3.1 Directions of associations reported in meta-analyses of randomised controlled trials (RCTs) or prospective cohort studies for total and types of dairy consumption and cardio-metabolic markers

Cardio-metabolic marker	First author and year	Ref.	# studies	Dairy type	Status of energy restriction	Direction of association*
Markers of body weight and composition						
Body weight (kg)	Geng, 2018	[58]	37 RCTs	Total dairy products	Total	↔
			17 RCTs		ER	↓
			19 RCTs		No ER	↑
	Stonehouse , 2016	[59]	27 RCTs	Total dairy products	ER	↓
	Booth, 2015	[60]	19 RCTs	Total dairy products	Total	↔
	Benatar, 2013	[63]	8 RCTs	Low-fat dairy products	No ER	↑
			10 RCTs	High-fat dairy products	No ER	↑
	Abargouei, 2012	[61]	14 RCTs	Total dairy products	Total	↔
			10 RCTs		ER	↓
	Chen, 2012	[62]	29 RCTs	Total dairy products	Total	↔
			16 RCTs		ER	↓
Body fat mass (kg)	Geng, 2018	[58]	25 RCTs	Total dairy products	Total	↔
			16 RCTs		ER	↓
			9 RCTs		No ER	↔
	Stonehouse , 2016	[59]	27 RCTs	Total dairy products	ER	↓
	Booth, 2015	[60]	13 RCTs	Total dairy products	ER	↓
	Abargouei, 2012	[61]	12 RCTs	Total dairy products	Total	↓
			9 RCTs		ER	↓
	Chen, 2012	[62]	29 RCTs	Total dairy products	Total	↓
			15 RCTs		ER	↓
			14 RCTs		No ER	↔
Body lean mass (kg)	Geng, 2018	[58]	12 RCTs	Total dairy products	Total	↑
			7 RCTs		ER	↔
			5 RCTs		No ER	↔
	Stonehouse , 2016	[59]	27 RCTs	Total dairy products	ER	↑
	Abargouei, 2012	[61]	6 RCTs	Total dairy products	Total	↑
			4 RCTs		ER	↑

Table 3.1 (continued)

Cardio-metabolic marker	Author, year	Ref.	# studies	Dairy type	Status of energy restriction	Direction of association
Waist circumference (cm)	Geng, 2018	[58]	37 RCTs	Total dairy products	Total	↓
			10 RCTs		ER	↓
			5 RCTs		No ER	↔
	Benatar, 2013	[63]	2 RCTs	Low-fat dairy products	No ER	↔
			4 RCTs	High-fat dairy products	No ER	↔
	Abargouei, 2012	[61]	8 RCTs	Total dairy products	Total	↓
			7 RCTs		ER	↓
Lipid markers						
Total cholesterol	de Goede, 2015	[83]	5 RCTs	Cheese (control: butter)	ER	↓
			4 RCTs	Cheese (control: tofu)	ER	↑
LDL-cholesterol	Benatar, 2013	[63]	3 RCTs	Low-fat dairy products	No ER	↔
			6 RCTs	High-fat dairy products	No ER	↔
	de Goede, 2015	[83]	5 RCTs	Cheese (control: butter)	ER	↓
			4 RCTs	Cheese (control: tofu)	ER	↑
HDL-cholesterol	Benatar, 2013	[63]	3 RCTs	Low-fat dairy products	No ER	↔
			5 RCTs	High-fat dairy products	No ER	↔
	de Goede, 2015	[83]	5 RCTs	Cheese (control: butter)	ER	↓
Triglycerides	de Goede, 2015	[83]	5 RCTs	Cheese (control: butter)	ER	↔
Glycaemic markers						
Fasting glucose	Benatar, 2013	[63]	4 RCTs	Low-fat dairy products	No ER	↔
			4 RCTs	High-fat dairy products	No ER	↔
HOMA-IR	Benatar, 2013	[63]	4 RCTs	High-fat dairy products	No ER	↔
Inflammatory markers						
CRP	Benatar, 2013	[63]	3 RCTs	Low-fat dairy products	No ER	↔
			3 RCTs	High-fat dairy products	No ER	↔

Table 3.1 (continued)

Cardio-metabolic marker	Author, year	Ref.	# studies	Dairy type	Status of energy restriction	Direction of association
Blood pressure						
SBP	Ding, 2017	[84]	8 RCTs	Total dairy products	Total	↔
	Benatar, 2013	[63]	3 RCTs	Low-fat dairy products	No ER	↔
			4 RCTs	High-fat dairy products	No ER	↔
	Dong, 2013	[85]	14 RCTs	Fermented milk	Total	↓
DBP	Benatar, 2013	[63]	3 RCTs	Low-fat dairy products	No ER	↔
			4 RCTs	High-fat dairy products	No ER	↔
	Dong, 2013	[85]	14 RCTs	Fermented milk	Total	↓
Hypertension	Soedamah-Muthu, 2012	[86]	9 prospective cohort studies	Total dairy products	Total	↓
			6 prospective cohort studies	Low-fat dairy products	Total	↓
			6 prospective cohort studies	High-fat dairy products	Total	↔
			7 prospective cohort studies	Milk	Total	↓
			5 prospective cohort studies	Yoghurt	Total	↔
			8 prospective cohort studies	Cheese	Total	↔

Abbreviations: RCTs: randomised controlled trials; ER: energy restriction

*↔: no association, ↑: positive association/increase in risk, ↓: inverse association/decrease in risk

Table 3.2 Observational studies examining the prospective associations between dairy consumption and anthropometric markers in multivariable models

Author, year	Ref.	Country	N	Follow-up years	Type of analysis	Dairy type	Direction of association*
Body weight							
Rautiainen ,2016	[69]	USA	18,438	17	Baseline dairy consumption, outcome change	Total dairy products	↓
						High-fat dairy products	↓
						Low-fat dairy products	↔
Wang ,2014	[67]	USA	3,440	13	Average of dairy consumption repeated measured, outcome change	Total dairy products	↓
						High-fat dairy products	↔
						Low-fat dairy products	↔
						Low-fat milk	↔
						Yoghurt	↓
						Cheese	↔
Martinez-Gonzalez ,2014	[73]	Spain	8,516	6.6	Baseline dairy consumption, average outcome change	Yoghurt	↔
						Full-fat yoghurt	↔
						Low-fat yoghurt	↔
Samara ,2013	[71]	France	588	5	Baseline dairy consumption, outcome change	Low-fat dairy products	↔
						Cheese	↔
Fumeron ,2011	[68]	France	3,417	9	Baseline dairy consumption, outcome change	Total dairy products	↓
						Cheese	↓
Mozaffarian ,2011	[70]	USA	120,877	4	Parallel change	High-fat dairy products	↔
						Low-fat dairy products	↔
						Full-fat milk	↔
						Low-fat milk	↔
						Yoghurt	↓
						Cheese	↔
						Butter	↑

Table 3.2 (continued)

Author, year	Ref.	Country	N	Follow-up years	Type of analysis	Dairy type	Direction of association
Snijder ,2008	[64]	Netherlands	1,124	6.4	Baseline dairy consumption, outcome change	Total dairy products	↔
Vergnaud ,2008	[65]	France	2,267	6	Baseline dairy consumption, outcome change	Total dairy products	↔
						Milk	↓
						Yoghurt	↓
						Cheese	↓
Rajpathak ,2006	[66]	USA	19,615	12	Baseline dairy consumption, outcome change	Total dairy products	↔
						High-fat dairy products	↓
						Low-fat dairy products	↔
					Parallel change	Total dairy products	↑
						High-fat dairy products	↑
						Low-fat dairy products	↔
Rosell ,2006	[72]	Sweden	19,352	9	Categories of dairy consumption change, outcome change	Full-fat milk	↓
						Low-fat milk	↔
						Cheese	↓
						Butter	↔
Body mass index							
Zong ,2014	[81]	China	2,091	6	Baseline dairy consumption, outcome change	Total dairy products	↓
Fumeron ,2011	[68]	France	3,417	9	Baseline dairy consumption, outcome change	Total dairy products	↔
						Cheese	↔
Snijder ,2008	[64]	Netherlands	1,124	6.4	Baseline dairy consumption, outcome change	Total dairy products	↔
Samara ,2013	[71]	France	588	5	Baseline dairy consumption, outcome change	Low-fat dairy products	↔
						Cheese	↔

Table 3.2 (continued)

Author, year	Ref.	Country	N	Follow-up years	Type of analysis	Dairy type	Direction of association
Waist circumference							
Zong ,2014	[81]	China	2,091	6	Baseline dairy consumption, outcome change	Total dairy products	↓
Wang ,2014	[67]	USA	3,440	13	Average of dairy consumption repeated measured, outcome change	Total dairy products	↔
						High-fat dairy products	↓
						Low-fat dairy products	↔
						Low-fat milk	↔
						Yoghurt	↓
						Cheese	↔
Samara ,2013	[71]	France	588	5	Baseline dairy consumption, outcome change	Low-fat dairy products	↔
						Cheese	↔
Fumeron ,2011	[68]	France	3,417	9	Baseline dairy consumption, outcome change	Total dairy products	↔
						Cheese	↔
Halkjaer ,2009	[80]	Denmark	42,696	5	Baseline dairy consumption, outcome change	High-fat dairy products	↓
						Low-fat dairy products	↔
						Butter	↓
Snijder ,2008	[64]	Netherlands	1,124	6.4	Baseline dairy consumption, outcome change	Total dairy products	↔
Vergnaud ,2008	[65]	France	2,267	6	Baseline dairy consumption, outcome change	Total dairy products	↔
						Milk	↓
						Yoghurt	↓
						Cheese	↔

*↔: no association, ↑: positive association/increase in risk, ↓: inverse associations/decrease in risk

3.2 Lipid markers

Randomised controlled trials of butter

There is a large number of studies on the associations between butter and blood lipids, because of the hypothesis that it increases blood cholesterol due to its high content of saturated fatty acids. Despite this large literature, it is challenging to draw conclusions as there is no meta-analysis of these RCTs so far. A simple synthesis of results would be subject to heterogeneity related to control groups, intervention duration and intensity, study design, sample size, country of origin, participants' gender, specific characteristics of the study sample e.g. postmenopausal women, hypercholesterolaemic individuals or individuals with metabolic syndrome. A meta-analysis is beyond the scope of the present PhD project, but a qualitative appraisal of the literature is presented.

Previous meta-analysis

The only meta-analysis published on this topic was by Zock and Katan in 1997 on 20 clinical trials published between 1957 and 1995. This meta-analysis examined the effects of a 10% energy substitution of hard and soft margarine with butter on total, low-density lipoprotein (LDL-), and high-density lipoprotein cholesterol (HDL-C), total-to-HDL-C ratio and triglyceride levels[87]. Butter increased total and LDL-C compared to both types of margarine and the increase was more pronounced when soft margarine was used as the control group (increase by 0.25 and 0.20 mmol/l of total and LDL-C respectively) than when hard margarine was used (increase by 0.19 and 0.11 mmol/l of total and LDL-C respectively)[87]. HDL-C and total-to-HDL-C ratio increased only when butter substituted hard margarine (increase by 0.02 mmol/l and 0.20 respectively), while no effect on triglyceride levels was observed when butter substituted either type of margarine[87].

Literature search

From a literature search in PubMed from 1980 to 2016 and extraction of relevant studies from references of other papers, 56 RCTs, which use butter either in the intervention or as a control and examine effects on blood lipids, were identified. Of those, 35 RCTs remained[88–122] after exclusion of 21 RCTs because either they used butter as part of both the intervention and the control group to compare the effects of other fat sources (n=11), or the intervention did not include solely butter (n=5), or they were included in the meta-analysis by Zock and Katan published in 1997 (n=3) or the intervention or control group was not of interest (n=2).

Description of identified studies

Among the 35 eligible RCTs, the majority of the trials are cross-over studies (n=27)[88–95, 97, 99, 101, 102, 104–115, 117, 118, 120] with a mean sample size of 43$\pm$42. Most of the studies included both men and women (n=26)[88–90, 92, 94–96, 98, 100, 102, 105–108, 110–112, 114–122], seven studies included men only[97, 99, 101, 103, 104, 109, 113] and two studies women only[91, 93]. Nine of the trials are postprandial studies[90, 93, 97,

101, 105, 106, 108, 109, 113], three had a duration of less than one month[92, 96, 103], 16 had a duration between one and three months[89, 91, 94, 95, 98, 100, 102, 104, 110–112, 114, 116, 119, 121, 122] and seven had a duration between three and six months[88, 99, 107, 115, 117, 118, 120]. Eight studies were done in Australia and New Zealand[90, 92, 102, 115–119], 10 in northern Europe[89, 93, 94, 97, 99, 100, 106, 111, 113, 121, 122], six in southern Europe[91, 95, 101, 104, 108, 109], five in the United States[105, 110, 112, 114, 120] and five elsewhere[88, 96, 98, 103, 107]. Thirteen out of the 35 studies did not include individuals from the general population, but included hypercholesterolemic individuals (n=4)[111, 116–118], individuals with metabolic syndrome (n=1)[98] or type 2 diabetes (n=3)[96, 106, 108].

High diversity was also observed concerning the food items used as comparison (control) groups in RCTs on the effect of butter on blood lipids. In order to make a qualitative summary synthesis of the results, it was considered reasonable to classify the comparison groups according to the intervention content of different fatty acids (monounsaturated fatty acids –MUFA-, polyunsaturated fatty acids -PUFA-, saturated fatty acids –SFA- and trans fatty acids –TFA-). The category of high MUFA content includes olive oil (n=12)[89, 91, 93, 95, 101, 104, 106, 109, 113, 119, 120, 122], rapeseed oil (n=3)[99, 113, 121], palm oil (n=1)[113], argan oil (n=1)[96] and soft margarine (n=3)[98, 105, 107]. The category of high PUFA content includes sunflower oil (n=4)[90, 97, 109, 113], soybean oil (n=2)[110, 120], safflower oil (n=3)[115, 118, 119], unsaturated margarine (n=4)[100, 114, 117, 121] and walnuts (n=2)[101, 104]. The category of high SFA content included coconut oil (n=3)[115, 118, 122] and cocoa butter (n=1)[120] and the category of high TFA content included hard margarine (n=5)[107, 110–112, 114] and a canola oil and PUFA blend enriched with TFAs (n=1)[116]. The comparison groups which included less SFA than butter such as cheese (covered in the next paragraph), milk (n=2)[92, 108] and the habitual diet (n=10)[89, 91, 94, 96, 98, 101–104, 116], cannot be classified in any of the aforementioned categories and constitute a separate category. Habitual diet as a comparison group is a special case, as its content might be characterised by a higher heterogeneity. Certain comparison groups such as carbohydrates (n=1)[119], and interesterified linseed and rapeseed oil that are high both in PUFA and MUFA[93] do not belong to any of the aforementioned categories and they have not been included in separate categories, as the number of relevant trials identified is insufficient.

Qualitative evidence synthesis

Based on this classification, overall, LDL-C increased when butter substituted food sources of MUFA[88, 89, 95, 99, 101, 104–107, 109, 113, 119–122] or PUFA[88, 100, 105, 110, 114, 115, 117, 118, 120, 121], or TFA[87, 107, 112, 114, 116]. Results are conflicting for lower SFA intake showing increase[88, 89, 91, 94, 102, 102, 103] or no effect[96, 98, 101, 104, 116] on LDL-C. For HDL-C, there was no effect when butter substituted

food sources of MUFA[87–89, 93, 95, 99, 101, 105, 107, 120–122], an increasing effect when it substituted hard margarine[87] and mixed effects when it substituted PUFA food sources[88, 90, 105, 110, 114, 115, 117, 118, 120, 121] or the habitual diet[91, 94, 96, 98, 101, 102, 104] in most of the studies. Total cholesterol increased when PUFA food sources[88, 100, 105, 110, 114, 115, 117, 118, 120, 121] or food items with generally lower SFA content[89, 91, 94, 102–104, 116, 120] were substituted with butter. For MUFA, results were conflicting showing no[91, 93, 95, 101] or increasing effects on total cholesterol[87–89, 95, 99, 105, 107, 120–122]. Most of the trials reported null effects on triglycerides when butter replaced any fat sources[87, 89–91, 93–99, 101, 102, 104, 105, 107–110, 113–117, 120–122].

Some characteristics of the 35 studies included in this qualitative evidence synthesis are presented in **Tables 3.3-3.6**.

It is recommended for future research that a meta-analysis of the RCTs examining the effects of butter on blood lipids, is conducted to obtain quantitative effects and to account for all sources of heterogeneity. The last meta-analysis of part of these RCTs was conducted in 1997 by Zock and Katan[87], thus an updated meta-analysis is needed.

Table 3.3 Characteristics of randomised controlled trials (RCTs) on the effects of butter consumption on **total cholesterol** stratified by the fat intake category of the comparison group

Comparison group category	Comparison group	Author, year	Ref.	Type of RCT	Country	Duration	N	Direction of association*
MUFA sources	Olive oil	Khaw, 2018	[122]	parallel	UK	1 month	91	↑
		Engel, 2015	[89]	cross-over	Denmark	2.6 months	47	↑
		Anderson-Vasquez, 2015	[91]	cross-over	Spain	1.9 months	18	↑
		Perez-Martinez, 2011	[95]	cross-over	Spain	3 months		↔
		Jimenez-Gomez, 2009	[101]	cross-over	Spain	prostprandial	20	↔
		Thomsen, 2003	[106]	cross-over	Denmark	prostprandial	12	↔
		Kris-Etherton, 1993	[120]	cross-over	USA	4 months	33	↑
	Argan oil	Ould Mohamedou, 2011	[96]	parallel	Morocco	3 weeks	86	↑
	Soft margarine	Gagliardi, 2010	[98]	parallel	Brazil	1.3 months	53	↔
		Lichtenstein, 2003	[105]	cross-over	USA	prostprandial	36	↔
		Mauger, 2003	[107]	cross-over	Canada	4 months	46	↑
	MUFA-enriched cheese	Brassard, 2017	[88]	cross-over	Canada	1 month	92	↑
PUFA sources	Sunflower oil	Dias, 2015	[90]	cross-over	Australia	prostprandial	26	↔
	Soybean oil	Han, 2002	[110]	cross-over	USA	1.3 months	19	↑
		Kris-Etherton, 1993	[120]	cross-over	USA	4 months	33	↑
	Safflower oil	Cox, 1998	[115]	cross-over	New Zealand	3.2 months	41	↔
		Cox, 1995	[118]	cross-over	New Zealand	3.2 months	28	↑
	PUFA-enriched margarine	Lecerf, 2009	[100]	parallel	UK	1.3 months	121	↑
		Judd, 1998	[114]	cross-over	USA	1.3 months	46	↑
		Chisholm, 1996	[117]	cross-over	New Zealand	4.3 months	49	↑
		Seppanen-Laakso, 1992	[121]	parallel	Finland	3 months	54	↑
	Walnuts	Jimenez-Gomez, 2009	[101]	cross-over	Spain	prostprandial	20	↔
	PUFA-enriched cheese	Brassard, 2017	[88]	cross-over	Canada	1 month	92	↑

Table 3.3 (continued)

Comparison group category	Comparison group	Author, year	Ref.	Type of RCT	Country	Duration	N	Direction of association
SFA sources	Coconut oil	Khaw, 2018	[122]	parallel	UK	1 month	91	↔
		Cox, 1998	[115]	cross-over	New Zealand	4.3 months	41	↔
		Cox, 1995	[118]	cross-over	New Zealand	4.3 months	28	↑
	Cocoa butter	Kris-Etherton, 1993	[120]	cross-over	USA	4 months	33	↑
TFA sources	Hard margarine	Mauger, 2003	[107]	cross-over	Canada	5.8 months	46	↔
		Han, 2002	[110]	cross-over	USA	3.2 months	19	↑
		Tonstad, 2001	[111]	cross-over	Norway	2.8 months	77	↑
		Denke, 2000	[112]	cross-over	USA	2.5 months	226	↑
		Judd, 1998	[114]	cross-over	USA	1.3 months	46	↑
	TFA-enriched PUFA blend	Noakes, 1998	[116]	parallel	Australia	2.8 months	38	↑
Sources of lower SFA content than butter	Habitual diet	Engel, 2015	[89]	cross-over	Denmark	2.6 months	47	↑
		Anderson-Vasquez, 2015	[91]	cross-over	Spain	1.9 months	18	↑
		Hjerpsted, 2011	[94]	cross-over	Denmark	3 months	49	↑
		Ould Mohamedou, 2011	[96]	parallel	Morocco	3 weeks	86	↔
		Gagliardi, 2010	[98]	parallel	Brazil	1.3 months	53	↔
		Jimenez-Gomez, 2009	[101]	cross-over	Spain	prostprandial	20	↔
		Gorguc, 2005	[103]	no control	Turkey	2.9 weeks	15	↑
		Nestel, 2005	[102]	cross-over	Australia	2.5 months	19	↑
		Noakes, 1998	[116]	parallel	Australia	2.8 months	38	↑

Abbreviations: MUFA: Monounsaturated fatty acids; PUFA: Polyunsaturated fatty acids; RCT: Randomised controlled trials; SFA: Saturated fatty acids; TFA: Trans fatty acids

*↔: no association, ↑: positive association/increase in risk, ↓: inverse associations/decrease in risk

Table 3.4 Characteristics of randomised controlled trials (RCTs) on the effects of butter consumption on **LDL-C** stratified by the fat intake category of the comparison group

Comparison group category	Comparison group	Author, year	Ref.	Type of RCT	Country	Duration	N	Direction of association*
MUFA sources	Olive oil	Khaw, 2018	[122]	parallel	UK	1 month	91	↑
		Engel, 2015	[89]	cross-over	Denmark	2.6 months	47	↑
		Anderson-Vasquez, 2015	[91]	cross-over	Spain	1.9 months	18	↑
		Perez-Martinez, 2011	[95]	cross-over	Spain	3 months		↑
		Jimenez-Gomez, 2009	[101]	cross-over	Spain	prostprandial	20	↔
		Sanders, 1994	[119]	parallel	Australia	1 month	30	↑
		Kris-Etherton, 1993	[120]	cross-over	USA	4 months	33	↑
	Argan oil	Ould Mohamedou, 2011	[96]	parallel	Morocco	3 weeks	86	↑
	Soft margarine	Gagliardi, 2010	[98]	parallel	Brazil	1.3 months	53	↔
		Lichtenstein, 2003	[105]	cross-over	USA	prostprandial	36	↑
		Mauger, 2003	[107]	cross-over	Canada	5.8 months	46	↑
	MUFA-enriched cheese	Brassard, 2017	[88]	cross-over	Canada	1 month	92	↑
PUFA sources	Sunflower oil	Dias, 2015	[90]	cross-over	Australia	postprandial	26	↔
	Soybean oil	Han, 2002	[110]	cross-over	USA	3.2 months	19	↑
		Kris-Etherton, 1993	[120]	cross-over	USA	4 months	33	↑
	Safflower oil	Cox, 1998	[115]	cross-over	New Zealand	4.5 months	41	↑
		Cox, 1995	[118]	cross-over	New Zealand	4.5 months	28	↑
		Sanders, 1994	[119]	parallel	Australia	1 month	30	↑
	PUFA-enriched margarine	Lecerf, 2009	[100]	parallel	UK	1.3 months	121	↑
		Judd, 1998	[114]	cross-over	USA	1.3 months	46	↑
		Chisholm, 1996	[117]	cross-over	New Zealand	4.3 months	49	↑
		Seppanen-Laakso, 1992	[121]	parallel	Finland	3 months	54	↑
	Walnuts	Jimenez-Gomez, 2009	[101]	cross-over	Spain	prostprandial	20	↑
	PUFA-enriched cheese	Brassard, 2017	[88]	cross-over	Canada	1 month	92	↑

Table 3.4 (continued)

Comparison group category	Comparison group	Author, year	Ref.	Type of RCT	Country	Duration	N	Direction of association
SFA sources	Coconut oil	Khaw, 2018	[122]	parallel	UK	1 month	91	↑
		Cox, 1998	[115]	cross-over	New Zealand	4.5 months	41	↑
		Cox, 1995	[118]	cross-over	New Zealand	4.5 months	28	↑
	Cocoa butter	Kris-Etherton, 1993	[120]	cross-over	USA	4 months	33	↔
TFA sources	Hard margarine	Mauger, 2003	[107]	cross-over	Canada	5.8 months	46	↑
		Han, 2002	[110]	cross-over	USA	3.2 months	19	↑
		Tonstad, 2001	[111]	cross-over	Norway	2.8 months	77	↑
		Denke, 2000	[112]	cross-over	USA	2.5 months	226	↑
		Judd, 1998	[114]	cross-over	USA	1.3 months	46	↑
	TFA-enriched PUFA blend	Noakes, 1998	[116]	parallel	Australia	2.8 months	38	↑
Sources of lower SFA content than butter	Habitual diet	Engel, 2015	[89]	cross-over	Denmark	2.6 months	47	↑
		Anderson-Vasquez, 2015	[91]	cross-over	Spain	1.9 months	18	↑
		Hjerpsted, 2011	[94]	cross-over	Denmark	3 months	49	↑
		Ould Mohamedou, 2011	[96]	parallel	Morocco	3 weeks	86	↔
		Gagliardi, 2010	[98]	parallel	Brazil	1.3 months	53	↔
		Jimenez-Gomez, 2009	[101]	cross-over	Spain	postprandial	20	↔
		Gorguc, 2005	[103]	no control	Turkey	2.9 weeks	15	↑
		Nestel, 2005	[102]	cross-over	Australia	2.5 months	19	↑
		Noakes, 1998	[116]	parallel	Australia	2.8 months	38	↔

Abbreviations: LDL-C: Low-density lipoprotein cholesterol; MUFA: Monounsaturated fatty acids; PUFA: Polyunsaturated fatty acids; RCT: Randomised controlled trials; SFA: Saturated fatty acids; TFA: Trans fatty acids

*↔: no association, ↑: positive association/increase in risk, ↓: inverse association/decrease in risk

Table 3.5 Characteristics of randomised controlled trials (RCTs) on the effects of butter consumption on **HDL-C** stratified by the fat intake category of the comparison group

Comparison group category	Comparison group	Author, year	Ref.	Type of RCT	Country	Duration	N	Direction of association*
MUFA sources	Olive oil	Khaw, 2018	[122]	parallel	UK	1 month	91	↔
		Engel, 2015	[89]	cross-over	Denmark	2.6 months	47	↑
		Anderson-Vasquez, 2015	[91]	cross-over	Spain	1.9 months	18	↓
		Perez-Martinez, 2011	[95]	cross-over	Spain	3 months		↔
		Jimenez-Gomez, 2009	[101]	cross-over	Spain	postprandial	20	↔
		Thomsen, 2003	[106]	cross-over	Denmark	postprandial	12	↓
		Sanders, 1994	[119]	parallel	Australia	1 month	30	↑
		Kris-Etherton, 1993	[120]	cross-over	USA	4 months	33	↔
	Argan oil	Ould Mohamedou, 2011	[96]	parallel	Morocco	3 weeks	86	↓
	Soft margarine	Gagliardi, 2010	[98]	parallel	Brazil	1.3 months	53	↔
		Lichtenstein, 2003	[105]	cross-over	USA	postprandial	36	↔
		Mauger, 2003	[107]	cross-over	Canada	5.8 months	46	↔
	MUFA-enriched cheese	Brassard, 2017	[88]	cross-over	Canada	1 month	92	↔
PUFA sources	Sunflower oil	Dias, 2015	[90]	cross-over	Australia	postprandial	26	↔
	Soybean oil	Han, 2002	[110]	cross-over	USA	3.2 months	19	↑
	Safflower oil	Cox, 1998	[115]	cross-over	New Zealand	4.5 months	41	↔
		Cox, 1995	[118]	cross-over	New Zealand	4.5 months	28	↑
		Sanders, 1994	[119]	parallel	Australia	1 month	30	↑
	PUFA-enriched margarine	Judd, 1998	[114]	cross-over	USA	1.3 months	46	↔
		Chisholm, 1996	[117]	cross-over	New Zealand	4.3 months	49	↔
		Seppanen-Laakso, 1992	[121]	parallel	Finland	3 months	54	↔
	Walnuts	Jimenez-Gomez, 2009	[101]	cross-over	Spain	postprandial	20	↔
	PUFA-enriched cheese	Brassard, 2017	[88]	cross-over	Canada	1 month	92	↔

Table 3.5 (continued)

Comparison group category	Comparison group	Author, year	Ref.	Type of RCT	Country	Duration	N	Direction of association
SFA sources	Coconut oil	Khaw, 2018	[122]	parallel	UK	1 month	91	↓
		Cox, 1998	[115]	cross-over	New Zealand	4.5 months	41	↔
		Cox, 1995	[118]	cross-over	New Zealand	4.5 months	28	↑
	Cocoa butter	Kris-Etherton, 1993	[120]	cross-over	USA	4 months	33	↔
TFA sources	Hard margarine	Mauger, 2003	[107]	cross-over	Canada	5.8 months	46	↑
		Han, 2002	[110]	cross-over	USA	3.2 months	19	↑
		Denke, 2000	[112]	cross-over	USA	2.5 months	226	↔
		Judd, 1998	[114]	cross-over	USA	1.3 months	46	↔
	TFA-enriched PUFA blend	Noakes, 1998	[116]	parallel	Australia	2.8 months	38	↔
Sources of lower SFA content than butter	Milk	Nestel, 2012	[92]	cross-over	Australia	3 weeks	21	↓
	Habitual diet	Engel, 2015	[89]	cross-over	Denmark	2.6 months	47	↑
		Anderson-Vasquez, 2015	[91]	cross-over	Spain	1.9 months	18	↔
		Hjerpsted, 2011	[94]	cross-over	Denmark	3 months	49	↔
		Ould Mohamedou, 2011	[96]	parallel	Morocco	3 weeks	86	↔
		Gagliardi, 2010	[98]	parallel	Brazil	1.3 months	53	↔
		Jimenez-Gomez, 2009	[101]	cross-over	Spain	postprandial	20	↔
		Gorguc, 2005	[103]	no control	Turkey	2.9 weeks	15	↑
		Nestel, 2005	[102]	cross-over	Australia	2.5 months	19	↔
		Noakes, 1998	[116]	parallel	Australia	2.8 months	38	↑

Abbreviations: HDL-C: High-density lipoprotein cholesterol; MUFA: Monounsaturated fatty acids; PUFA: Polyunsaturated fatty acids; RCT: Randomised controlled trials; SFA: Saturated fatty acids; TFA: Trans fatty acids

*↔: no association, ↑: positive association/increase in risk, ↓: inverse association/decrease in risk

Table 3.6 Characteristics of randomised controlled trials (RCTs) on the effects of butter consumption on **triglycerides** stratified by the fat intake category of the comparison group

Comparison group category	Comparison group	Author, year	Ref.	Type of RCT	Country	Duration	N	Direction of association*
MUFA sources	Olive oil	Khaw, 2018	[122]	parallel	UK	1 month	91	↔
		Engel, 2015	[89]	cross-over	Denmark	2.6 months	47	↔
		Anderson-Vasquez, 2015	[91]	cross-over	Spain	1.9 months	18	↔
		Svensson, 2011	[93]	cross-over	Sweden	postprandial	19	↔
		Perez-Martinez, 2011	[95]	cross-over	Spain	3 months		↔
		Jimenez-Gomez, 2009	[101]	cross-over	Spain	postprandial	20	↔
		Bellido, 2004	[104]	cross-over	Spain	1 month	80	↔
		Thomsen, 2003	[106]	cross-over	Denmark	postprandial	12	↑
		Mekki, 2002	[109]	cross-over	France	postprandial	10	smaller ↑
		Nielsen, 2000	[113]	cross-over	Denmark	postprandial	18	↔
		Kris-Etherton, 1993	[120]	cross-over	USA	4 months	33	↔
	Palm oil	Nielsen, 2000	[113]	cross-over	Denmark	postprandial	18	↔
	Argan oil	Ould Mohamedou, 2011	[96]	parallel	Morocco	3 weeks	86	↑
	Soft margarine	Gagliardi, 2010	[98]	parallel	Brazil	1.3 months	53	↔
		Lichtenstein, 2003	[105]	cross-over	USA	postprandial	36	↔
		Mauger, 2003	[107]	cross-over	Canada	5.8 months	46	↔
	MUFA-enriched cheese	Brassard, 2017	[88]	cross-over	Canada	1 month	92	↔
PUFA sources	Sunflower oil	Dias, 2015	[90]	cross-over	Australia	postprandial	26	↔
		Masson, 2011	[97]	cross-over	Netherlands	postprandial	13	↔
		Mekki, 2002	[109]	cross-over	France	postprandial	10	smaller ↑
		Nielsen, 2000	[113]	cross-over	Denmark	postprandial	18	↔
	Soybean oil	Han, 2002	[110]	cross-over	USA	3.2 months	19	↔
		Kris-Etherton, 1993	[120]	cross-over	USA	4 months	33	↑

Table 3.6 (continued)

Comparison group category	Comparison group	Author, year	Ref.	Type of RCT	Country	Duration	N	Direction of association
	Safflower oil	Cox, 1998	[115]	cross-over	New Zealand	4.5 months	41	↔
		Cox, 1995	[118]	cross-over	New Zealand	4.5 months	28	↑
	PUFA-enriched margarine	Judd, 1998	[114]	cross-over	USA	1.3 months	46	↔
		Chisholm, 1996	[117]	cross-over	New Zealand	4.3 months	49	↔
	Walnuts	Jimenez-Gomez, 2009	[101]	cross-over	Spain	postprandial	20	↔
		Bellido, 2004	[104]	cross-over	Spain	1 month	80	↔
	PUFA-enriched cheese	Brassard, 2017	[88]	cross-over	Canada	1 month	92	↔
SFA sources	Coconut oil	Khaw, 2018	[122]	parallel	UK	1 month	91	↔
		Cox, 1998	[115]	cross-over	New Zealand	4.5 months	41	↔
		Cox, 1995	[118]	cross-over	New Zealand	4.5 months	28	↑
	Cocoa butter	Kris-Etherton, 1993	[120]	cross-over	USA	4 months	33	↔
TFA sources	Hard margarine	Mauger, 2003	[107]	cross-over	Canada	5.8 months	46	↔
		Han, 2002	[110]	cross-over	USA	3.2 months	19	↔
		Denke, 2000	[112]	cross-over	USA	2.5 months	226	↑
		Judd, 1998	[114]	cross-over	USA	1.3 months	46	↔
	TFA-enriched PUFA blend	Noakes, 1998	[116]	parallel	Australia	2.8 months	38	↑

Table 3.6 (continued)

Comparison group category	Comparison group	Author, year	Ref.	Type of RCT	Country	Duration	N	Direction of association
Sources of lower SFA content than butter	Milk	Clemente, 2003	[108]	cross-over	Italy	postprandial	8	↔
	Habitual diet	Engel, 2015	[89]	cross-over	Denmark	2.6 months	47	↔
		Anderson-Vasquez, 2015	[91]	cross-over	Spain	1.9 months	18	↔
		Hjerpsted, 2011	[94]	cross-over	Denmark	3 months	49	↔
		Ould Mohamedou, 2011	[96]	parallel	Morocco	3 weeks	86	↔
		Gagliardi, 2010	[98]	parallel	Brazil	1.3 months	53	↔
		Jimenez-Gomez, 2009	[101]	cross-over	Spain	postprandial	20	↔
		Gorguc, 2005	[103]	no control	Turkey	2.9 weeks	15	↑
		Nestel, 2005	[102]	cross-over	Australia	2.5 months	19	↔
		Bellido, 2004	[104]	cross-over	Spain	1 month	80	↔
		Noakes, 1998	[116]	parallel	Australia	2.8 months	38	↑
		Seppanen-Laakso, 1992	[121]	parallel	Finland	3 months	54	↔

Abbreviations: MUFA: Monounsaturated fatty acids; PUFA: Polyunsaturated fatty acids; RCT: Randomised controlled trials; SFA: Saturated fatty acids; TFA: Trans fatty acids

*↔: no association, ↑: positive association/increase in risk, ↓: inverse association/decrease in risk

Evidence on other dairy types

Total dairy consumption did not exert any effects on total cholesterol in four RCTs[77–79, 123], but exerted an increasing effect in one RCT[124] when compared with habitual consumption under an isocaloric experiment and a decreasing effect in another RCT when compared with a diet lower in dairy products under no energy restriction[125]. The effects on total cholesterol reported by four RCTs for low-fat dairy consumption were null[75, 76, 126, 127]. Evidence on dairy types in relation to total cholesterol and the ratio of total to HDL-C from RCTs and prospective cohort studies is limited.

RCTs reported null effects of total dairy consumption on LDL-C[77–79, 123, 124] or HDL-C[77–79, 124]. A meta-analysis of nine RCTs without energy restriction indicated null effects of low- and high-fat dairy consumption on LDL-C or HDL-C[63] (Table 3.1).

Cheese consumption had a decreasing effect on total cholesterol, LDL-C and HDL-C, but no effect on triglycerides, when compared with butter consumption, even though the intervention and control group had the same intake of PUFA/SFA ratio, as reported in a meta-analysis of five RCTs[83] (Table 3.1). In the same publication, through a narrative review of RCTs on the effect of cheese on blood lipids when compared with tofu consumption (n=4) or a modified cheese with a higher PUFA/SFA ratio (n=3), cheese increased total cholesterol and LDL-C overall and had no effect on HDL-C and triglycerides[83]. Interventions of dairy types are limited.

Associations on total dairy consumption and LDL-C reported from prospective cohort studies are null[64, 81], and the same holds for most associations with HDL-C[68, 81, 82, 128].

Null associations have also been reported between cheese and HDL-C[68, 71, 82, 129], whereas evidence on associations between other dairy types and LDL-C or HDL-C from prospective cohort studies is sparse. No effect on fasting triglycerides, was reported from interventions of total[77, 78, 123, 124] or low-fat dairy products[75, 76, 126, 127] and most of the prospective cohort studies also reported null associations for total dairy consumption[64, 68, 81, 82, 128].

The number of studies on associations between specific dairy types and triglycerides is limited. Specifically for yoghurt, there have been several RCTs examining its effect on the lipid profile, but there is a high variability in the type of yoghurt used, as many of them did not involve conventional yoghurt, but yoghurt with additional microbial species[130]. A review of such studies concluded that despite this heterogeneity most studies showed beneficial effects of the different types of yoghurt on the lipid profile, which is not though very informative for the effect of conventional yoghurt[130].

Evidence on associations of total and types of dairy products consumption with apolipoprotein A1 (ApoA1), ApoB and non-esterified fatty acids (NEFA) is sparse.

3.3 Glycaemic markers

Glucose

Evidence from RCTs does not show any effect of total dairy consumption on fasting blood glucose[77–79, 123–125]. In a meta-analysis of eight RCTs, no effect on fasting blood glucose was reported from interventions of low-fat or high-fat dairy products[63] (Table 3.1). The number of interventions of dairy types also examining fasting blood glucose is limited. Results from prospective cohort studies for total and low-fat dairy consumption are conflicting, showing null[82, 128, 131, 132] or inverse[64, 71, 81, 82, 131] associations, while associations are consistently null for cheese consumption[71, 82, 131, 132]. Studies on associations of other dairy types and fasting blood glucose, as well as on dairy consumption and postprandial glucose, are sparse.

Haemoglobin A1c

For haemoglobin A1c (HbA1c), there seems to be a tendency towards a null association with total dairy consumption as reported in prospective cohort studies[81, 131, 132]. For the association with dairy types the available studies are few, and evidence from RCTs is sparse, so no meaningful conclusion can be drawn.

Insulin

Most of the RCTs reported null effects of total dairy consumption on fasting insulin[77–79, 123, 124]. The number of interventions of dairy types examining postprandial insulin are insufficient to draw conclusions. The number of prospective cohort studies for such associations is even more limited.

Indices for insulin resistance

Homeostasis Model Assessment for Insulin Resistance (HOMA-IR) is highly correlated (r=0.88) with euglycemic clamp, which is considered the gold standard for the measurement of insulin sensitivity[133]. The index is estimated as shown in equation (3.1).

$$HOMA-IR = \frac{FastingPlasmaGlucose * FastingPlasmaInsulin}{22.5} \tag{3.1}$$

Homeostasis Model Assessment for β-cell function (HOMA-%B) is a less frequently used index for the assessment of β-cell function and is estimated as shown in equation (3.2).

$$HOMA-\%B = \frac{20 * FastingPlasmaInsulin}{FastingPlasmaGlucose - 3.5} \tag{3.2}$$

RCTs reported null[79, 123] or decreasing[124, 125] effects of total dairy consumption on HOMA-IR and null effects of low-[63] or high-fat[63] dairy consumption. Evidence on dairy types and evidence from prospective cohort studies, as well as for HOMA-%B, is sparse and thus inconclusive.

3.4 Inflammatory markers

In a systematic review of 52 controlled trials and prospective cohort studies (78 results extracted), Bordoni et al. created an inflammation score from the studies identified. This score was positive for a study if the results indicated a beneficial effect/association of dairy consumption on inflammatory markers, negative if the results indicated an increasing effect/association, or zero for null effects/associations[134]. The higher the absolute value of the score, the more criteria the study met among a range of 12 criteria. These criteria related to the study design (controlled trial, randomisation), the results (effect magnitude, sustainable effects, dose-response effects) and the interpretation (biological plausibility, clinical significance)[134].The most frequently investigated inflammatory markers were C-reactive protein (CRP; n=51), interleukin-6 (n=44) and tumour necrosis factor α (n=36)[134].

The findings of this systematic review indicated statistically significant anti-inflammatory effects of dairy consumption (n=61), which were more pronounced in persons with metabolic disease (n=24)[134]. No effects were indicated among people with gastrointestinal disorders (n=8) and pro-inflammatory effects were observed among people with a reported milk allergy (n=6)[134]. When results were stratified by dairy fat content, both low-(n=20) and high-fat (n=35) dairy consumption exerted anti-inflammatory effects with more pronounced effects for low-fat dairy[134].

Stratification by fermentation status, showed that only fermented dairy products (n=16) had a significant anti-inflammatory effect[134]. Despite the significant effects/associations, it should be noted that the magnitude of the effects was low. In another meta-analysis of six RCTs, low- and high-fat dairy consumption was not associated with CRP[63] (Table 3.1), whereas results for dairy types and results from prospective cohort studies are insufficient to draw conclusions.

3.5 Markers of hepatic function

According to my literature search, there are no studies available on the association between dairy consumption and markers of hepatic function. However, there is some evidence, which supports the generation of a hypothesis that dairy consumption could be associated with hepatic function. Liver function tests i.e. gamma-glutamyl transferase (γ-GT), aspartate aminotransferase (AST), and alanine transaminase (ALT)[135, 136], and non-alcoholic fatty liver disease (NAFLD)[137] have been reported to be associated with incident type 2 diabetes. De novo lipogenesis is an underlying pathway for NAFLD[138], which may be linked to cardio-metabolic disease through liver fat accumulation[139]. Diet provides the substrates for de novo lipogenesis i.e. fat and carbohydrates, but can

also be associated with NAFLD through obesity, insulin resistance, hyperlipidaemia and inflammation[140].

3.6 Blood pressure

Dairy consumption did not have an effect on systolic blood pressure (SBP) in a meta-analysis of eight RCTs[84] (Table 3.1). Similar results were reported for low- and high-fat dairy consumption and SBP or diastolic blood pressure (DBP) in a meta-analysis of seven RCTs with no energy restriction[63] (Table 3.1). In another meta-analysis of 14 RCTs investigating the effect of probiotic fermented milk, the pooled result indicated a decrease in SBP and DBP with a more pronounced effect among hypertensive individuals[85] (Table 3.1). In observational research, two meta-analyses of prospective cohort studies concluded that total dairy products[84, 86], low fat-dairy products[86] and milk[86] were associated with a 3-4% lower risk of elevated blood pressure per 200 g/d, but null associations were reported for high-fat dairy products (n=6)[86], yoghurt (n=4)[86] or cheese (n=8)[86] (Table 3.1). Evidence on low- and high-fat dairy types is sparse.

In summary, fermented dairy products may play a beneficial role in the regulation of blood pressure, while total, low-fat dairy products and milk may exert such beneficial effects mainly within high blood pressure levels, while they do not seem to have an effect within normal levels.

3.7 Metabolic syndrome

Metabolic syndrome, which is a cluster of risk factors of type 2 diabetes (T2D) and cardiovascular disease (CVD) including high blood glucose, high blood pressure, high triglycerides, low HDL-C and central overweight/obesity[141], has been reported to highly predict mainly T2D[142], but also CVD[143]. A meta-analysis of eight prospective cohort studies with an average follow-up of 6.7 years showed a decrease in the risk of developing metabolic syndrome by 15% for higher compared to lower total dairy consumption categories[144]. For types of dairy products it is not possible to reach definite conclusions due to the limited number of available studies.

3.8 Gaps in the literature

From the literature review presented above, it becomes evident that the combinations of dairy exposures and cardio-metabolic outcomes investigated differ across studies substantially. This results in inconsistent evidence and for certain associations, especially for specific dairy types, too sparse evidence to draw any conclusions. Even in RCTs, considered the gold standard of the hierarchy of evidence, many limitations can be identified,

such as questionable compliance, impossible blinding, short duration, low sample size, diversity in comparison groups, no assessment of habitual consumption or funding by the food industry[63].

In addition, the inconsistency of the definitions of dairy groups (total, low- and high-fat), which have been used in different studies might result in inconsistent findings. Dairy products are a heterogeneous food group with a variety of macro- and micro-nutrients. Thus, a consistent definition of dairy exposures is needed for the comparability of results across studies. This inconsistency is also evident from the relevant published meta-analyses, which pool results from studies of different dairy products, as the number of studies for each dairy type separately is not sufficient to draw conclusions. Consequently, there is a need for a comprehensive set of exposures, outcomes and covariates and a consistent methodology to be used in the study of the associations of dairy products with cardio-metabolic markers to further elucidate the link to cardio-metabolic disease.

Chapter 4

Associations of dairy consumption with cardio-metabolic markers: a cross-sectional analysis in the Fenland Study, UK

Summary

Background and aims: Epidemiological evidence on associations between dairy types and cardio-metabolic risk is inconsistent. We aimed to investigate cross-sectional associations between total and types of dairy products and markers of metabolic risk and adiposity.

Methods: We included 12,065 adults (54% women) aged 30 to 65 years recruited to the Fenland study between 2005 and 2015 in Cambridgeshire UK. Diet including dairy products (milk, yoghurt, cheese, cream, butter, ice-cream) was assessed with a food frequency questionnaire. Markers of adiposity were measured with dual energy X-ray absorptiometry and ultrasonography. Blood pressure and circulating concentrations of lipid, glycaemic, and hepatic markers, C-reactive protein and adiponectin were measured. Associations between dairy consumption and these outcomes were assessed using robust regression, adjusted for socio-demographic, lifestyle, and dietary factors including total energy intake and body-mass index and corrected for false-discovery rate.

Results: The median (IQR) of milk, yoghurt, and cheese consumption were 293(146 - 439), 35.3 (8.8 – 71.8), and 14.6 (4.8 – 26.9) g/day, respectively. Low-fat dairy consumption was inversely associated with visceral-to-subcutaneous fat ratio [% difference per serving (95% CI): -2.58 (-3.91, -1.23); p=0.0002]. Habitual daily consumption per one glass of milk was associated with 0.33 kg higher lean mass (95% CI: 0.19, 0.46; $p=2.5\times10^{-6}$). High-fat dairy products were positively associated with total and low-density lipoprotein cholesterol [0.06 mmol/l (0.03, 0.09); $p=3.5\times10^{-5}$ and 0.05 mmol/l (0.03, 0.08); $p=7\times10^{-6}$

respectively], while low-fat dairy and milk consumption were inversely associated with high-density lipoprotein cholesterol [-0.02 mmol/l (-0.03, -0.01); p=1.5×10^{-6} for both]. Other associations were not significant.

Conclusion: Our novel finding of an inverse association between low-fat dairy and milk consumption and visceral-to-subcutaneous fat ratio suggests fat distribution as a potential pathway for the link between dairy consumption and metabolic risk. These findings should be confirmed in prospective and experimental studies, in other populations, and combined with research investigating mechanisms.

What is already known

- Dairy products have been associated with no or a lower risk of cardio-metabolic disease endpoints.
- Evidence from randomised controlled trials supports favourable effects of total dairy consumption on body weight and composition under conditions of energy restriction, increasing effects of butter, but more favourable effects of cheese on blood lipipds, no effects of total dairy consumption on glycaemia and no effects of total, high-fat or low-fat dairy consumption on blood pressure.
- Evidence on specific dairy types, effects on other cardio-metabolic markers such as inflammation and hepatic function and more long-term associations of habitual dairy consumption with cardio-metabolic markers is limited.

What this research adds

- Low-fat dairy consumption was associated with lower visceral-to-subcutaneous fat ratio by 2.58% per serving/day.
- Habitual daily consumption of one glass of milk was associated with 0.33 kg higher lean mass.
- The cross-sectional design of this study does not allow causal inferences, but findings are hypothesis-generating, which might be further examined in prospective cohort studies and randomised controlled trials.

Publication

Trichia E, Imamura F, Brage S, De Lucia Rolfe E, Griffin SJ, Wareham NJ, Forouhi N G. Associations of types of dairy consumption with cardio-metabolic risk and adiposity: cross-sectional findings from over 12,000 adults in the Fenland Study, UK, (Manuscript in revision at the Journal of Nutrition)

4.1 Previous evidence on cross-sectional associations

The main limitation of cross-sectional studies is the higher possibility of reverse causation due to the weakness of this study design to fulfil the temporality requirement for causality, as the exposure and the outcome are measured at the same time point. Certain phenotypes

can influence people's behaviour both in the context of actual consumption, but also in the context of reporting their consumption and it can be assumed that the more apparent this phenotype is, the higher the probability of behaviour change.

The most well-characterised phenotype related to change in consumption and reporting behaviour is body mass index (BMI), as people with higher BMI might under-report or consume lower amounts of foods perceived as unhealthy and over-report or consume higher amounts of foods perceived as healthy[145–152]. An example could be that people of higher BMI over-report or consume more low-fat dairy products and under-report or consume less high-fat dairy products. For this reason, the validity of cross-sectional studies on the associations between dairy products and body weight, BMI or obesity status could be questionable. In addition, it might be useful for cross-sectional studies to adjust for BMI when examining associations between dairy consumption and other cardio-metabolic markers. Thus, for the appraisal of previous evidence on cross-sectional associations between dairy consumption and cardio-metabolic markers, studies using body weight, BMI or obesity status as outcomes and other studies not adjusting for BMI or another measure of body mass were not considered. In future meta-analyses of cross-sectional studies, it would be of interest to investigate heterogeneity of associations by adjustment for BMI in individual studies, which is out of the scope of the present project.

Cross-sectional studies investigating associations of total and types of dairy products with markers of body composition, lipid markers and markers of hepatic function also adjusting for BMI are limited. Concerning glycaemic markers, results on the associations between total dairy consumption and fasting blood glucose are inconsistent indicating null[131, 153] or inverse[154] associations, whereas the number of studies for other glycaemic markers and specific dairy types is limited. High-fat dairy consumption was not associated with C-reactive protein (CRP)[154–156], while evidence on total and types of dairy products is sparse. Results on associations between total dairy consumption and systolic blood pressure (SBP) showed inverse[131, 154] or positive[157] associations and for diastolic blood pressure (DBP) showed inverse associations[131, 154, 157], while for dairy types results are limited and thus inconclusive.

4.2 Study aims

Based on the broader inconsistencies on associations of total and types of dairy products with cardio-metabolic markers from prospective or randomised controlled trial (RCT) evidence (Chapter 3) and that from cross-sectional studies (appraised above), this study aimed to investigate associations of total and types of dairy products with markers of metabolic risk and adiposity. Of special interest were associations between total and types of dairy products and body fat mass and lean mass distribution, for which evidence is even more limited, but they provide objective and precise markers of body composition.

4.3 Methods

4.3.1 Study design and population

The Fenland study is a prospective cohort study with complete baseline measurements (Phase 1; 2005-2015; n=12,434) and an ongoing Phase 2 (follow-up) since 2014. Eligible participants were born between 1950 and 1975, were residing in Ely, Wisbech, Cambridge or surrounding villages and were recruited through general practitioners. Exclusion criteria included known history of diabetes, psychotic or terminal illness, inability to walk unaided, pregnancy, or lactation. For the present, cross-sectional analysis, a sample of 12,065 participants was used after exclusion of participants with missing dietary data (n=17), participants in the bottom and top 1% of total energy intake (n=250) and in the top 1% of total dairy consumption (n=97) and pregnant women (n=5). The study was approved by the Cambridge Regional Ethics Committee. All participants gave written informed consent.

4.3.2 Dietary assessment

Participants' diet over the last year was assessed with a 130-item semi-quantitative food frequency questionnaire (FFQ)[158]. The FFQ consisted of two parts: the first part included a list of items and nine frequencies from "Never or less than once/month" to "6+ per day" and the second part included questions on consumption of milk, breakfast cereals, cooking fat, added salt and dietary supplements. The dietary data from the FFQ were processed with the FFQ EPIC Tool for Analysis (FETA) software[159]. Dairy products were assessed in servings/day and were categorised as previously described[160] and shown in **Table 4.1**.

Table 4.1 Definitions of dairy groups

Dairy group	Definition
Full-fat milk	Goat's milk; Channel Islands milk; Silver top full-cream milk; Evaporated milk whole diluted; Sheep's milk
Low-fat milk	Semi-skimmed milk; Skimmed milk; Skimmed milk as reconstituted dried milk
Milk	Full-fat milk, Low-fat milk
Yoghurt	Full-fat yoghurt †; Low-fat yoghurt †
Cheese	High-fat cheese ‡; Low-fat cheese §
Cream ‖	Single cream; Double cream
Low-fat fermented dairy products	Yoghurt; Low-fat cheese
Fermented dairy products	Yoghurt; Cheese
High-fat dairy products	Full-fat milk; High-fat cheese; Cream; Butter; Ice-cream
Low-fat dairy products	Low-fat milk; Yoghurt; Low-fat cheese
Total dairy products	Milk; Yoghurt; Cheese; Cream; Butter; Ice-cream

†The variables derived directly from the FFQ questions were used

‡The variable derived directly from the FFQ questions on hard cheese intake was used. The assumption made here is that high-fat cheese is equivalent to hard cheese.

§The variable derived directly from the FFQ questions on cottage and low-fat soft cheese intake was used. The assumption made here is that low-fat cheese is equivalent to cottage and low-fat, soft cheese.

‖Cream was used as a contributor to high-fat and total dairy products, but results separately for it and its types are not presented, as the very low intakes result in very unstable and imprecise estimates

Abbreviations: FFQ: Food frequency questionnaire

4.3.3 Assessment of markers of metabolic risk and adiposity

Markers of body weight, body composition, and blood pressure

Trained research staff conducted all measurements according to standardised procedures. BMI was calculated as weight divided by height squared (kg /m^2). Waist and hip circumferences were averaged from two repeated measures with a non-stretch tape. Total body fat and its distribution (peripheral fat, visceral adipose tissue -VAT-, subcutaneous adipose tissue -SCAT-) and lean mass and its distribution (appendicular lean mass) were estimated with a Dual-Energy X-ray Absorptiometry (DEXA; Lunar Prodigy Advanced fan beam scanner, GE Healthcare, Bedford, United Kingdom; GE encore software, version 14.10.022 to 16, GE Medical Systems14). Abdominal ultrasonography (LOGIQ Book ultrasound system, and Logic eGE Healthcare with a 3C MHz-RS and 2-5 MHz 3C-RC curved array transducers respectively), was also used to derive VAT and SCAT[161] and to assess hepatic fat. VAT was defined as the distance from the peritoneum boundary to the lumbar spine and SCAT as the depth from the skin boundary to the linea alba. This method has been validated against measurements of intra-abdominal and subcutaneous adipose tissue with magnetic resonance imaging[162]. A semi-quantitative hepatic fat score was derived from ultrasonography using a method previously validated against the gold standard proton magnetic resonance spectroscopy[163]. The score ranged from 3 to 12, representing the increasing degree of hepatic fat. Blood pressure was measured with an automated sphygmomanometer (Omron, 705CP-II) three times and the average of the measurements was used.

Biochemical analyses

A 47.2 ml blood sample was taken from every participant after an overnight fast. Fasting and 2-hour plasma glucose were measured during a standard (75g) oral glucose tolerance test. Haemoglobin A1c (HbA1c) was measured in whole blood. Insulin, total cholesterol, high density lipoprotein cholesterol (HDL-C), triglycerides, apolipoprotein A1 (ApoA1), apolipoprotein B (ApoB), non-esterified fatty acids (NEFA), alanine aminotransferase (ALT), γ-glutamyl transferase (GGT), C-Reactive Protein (CRP) and adiponectin were measured in serum. Low density lipoprotein cholesterol (LDL-C) was calculated from triglycerides, HDL-C and total cholesterol using the Friedewald formula (LDL-C= Total cholesterol – HDL-C – (Triglycerides/2.2)[164]. The Homeostasis Model Assessment for Insulin Resistance (HOMA-IR) and HOMA for beta cell function (HOMA-%B) were calculated using the HOMA calculator (https://www.dtu.ox.ac.uk/homacalculator/).

4.3.4 Assessment of socio-demographic and lifestyle factors

A general questionnaire was administered for information on ethnicity, occupation, educational level, income, marital status, smoking and medication use. Physical activity was objectively measured over seven days using a combined heart rate and movement sensor (Actiheart, CamNtech, Papworth, UK) and individually calibrated with a treadmill test to derive physical activity energy expenditure[165]. This method has been previously validated with doubly-labelled water[166].

4.3.5 Statistical analysis

Descriptive characteristics of socio-demographic, behavioural, clinical and dietary factors were derived for the whole sample, as well as for the top and bottom categories of milk (non-consumers and 585-732 g/d), yoghurt (non-consumers and top quartile within consumers) and cheese (non-consumers and top quartile within consumers) consumption. The mean and standard deviation (SD) for the continuous variables and the frequencies and percentages of each category for categorical variables were calculated.

As primary exposures, I considered the three main dairy types i.e. milk, yoghurt and cheese. As primary outcomes, I considered one representative marker from each of the marker clusters of anthropometry, glycaemia, lipidaemia, inflammation, blood pressure and hepatic function. Thus, I used the ratio of VAT to SCAT (VAT/SCAT), HbA1c, total-to-HDL-C ratio, CRP, SBP and DBP and hepatic fat score respectively.

Positively-skewed outcome variables, (VAT/SCAT, HOMA2-IR, HOMA2-%B, fasting glucose, 2-hour glucose, fasting insulin, triglycerides, NEFA, ALT, γ-GT, CRP, adiponectin), were log-transformed. A metabolic risk z-score was calculated[167] and used in a secondary analysis to investigate the association between dairy products and a composite index of cardio-metabolic markers, by taking averages of z-scores of blood pressure (average of the z-scores of SBP and DBP), HbA1c, and waist circumference and lipids (-1 $\times$ HDL-C and log-transformed triglycerides).

Missing covariates were imputed with multiple imputation by chained equations (MICE) under the assumption of data missing at random[168]. Multiple imputation involved linear, logistic and predictive mean matching models according to variable distribution and 5 imputation datasets were derived. Outcome variables were accounted for in the imputation, but only participants with non-missing observations for the outcome variables were included in the estimation models. Imputation diagnostics (comparison of the original and the imputed data distributions, convergence of the imputation models, fraction of missing information, Monte Carlo errors) were performed to assess the validity of the imputation[169].

To examine cross-sectional associations between different dairy types and cardio-metabolic markers, robust multiple linear regression was used, deriving multiple maximum

likelihood (MM-) estimators, which are robust against the influence of outliers[170]. Poisson regression was used for hepatic fat score, where the ordinal variable (1 to 12) was assumed to follow a Poisson distribution (better fit in preliminary analysis than ordered logistic or negative binomial model, based on Akaike information criteria)[171].

The initial probability of false positive findings was set to 5%. Because of the large number of tests, false discovery rate correction was applied accounting for correlations between tests and assuming an arbitrary correlation pattern[172]. Associations were considered significant if they passed this correction (p<0.00025).

Associations were adjusted for potential confounders based on previous knowledge and biological plausibility using four statistical models. In model 1, I adjusted for socio-demographic factors including age, sex, test-site and ethnicity, total energy intake, and dairy products other than the dairy exposure. In model 2, I further adjusted for age at completion of full-time education, pack-years of smoking, physical activity energy expenditure (all continuous), and education level, occupation, household income, marital status, smoking status, hormone-replacement therapy (HRT; for women only), lipid-lowering medication, and anti-hypertensive medication (categorical). In model 3, I further adjusted for plasma vitamin C levels (as a marker of diet quality, reflecting fruit and vegetable intake), dietary supplement use, and consumption of non-dairy dietary factors. In model 4, I further adjusted for BMI to partly account for the possibility of dietary misreporting and lifestyle confounding due to obesity status. When the outcome was lean mass, models were further adjusted for height.

Pre-specified tests for effect modification by sex and BMI for each association were investigated. As sensitivity analyses, I repeated regression analyses in the complete-case dataset and with 10 imputed datasets, to examine stability of results based on five imputed datasets in the primary analyses. To assess whether non-linear associations were present, restricted cubic spline regressions (three knots at the 10^{th}, 50^{th} and 90^{th} percentiles) were fitted in the maximally adjusted model. For the same purpose, categorical exposures were used with five categories including non-consumers and quartiles among consumers for dairy types. The categories were generated from the residuals of the regression of total energy intake against dairy products[173].

In post-hoc analyses, I examined whether the identified significant associations can be explained by nutrients contained in dairy products including calcium, potassium, magnesium, phosphorus, vitamin A, vitamin B_{12}, lactose, mono-unsaturated fat and saturated fat from the dairy exposure one-by-one.

All analyses were conducted using Stata 14.2 (College Station, TX: StataCorp LP, 2015).

4.4 Results

4.4.1 Descriptive characteristics

The analyses evaluated 12,065 adults (54% women) with a mean±SD age of 48.6±7.5 years. The median (IQR) dairy consumption was: milk 293(146 - 439) g/day; yoghurt 35.3 (8.8 – 71.8) g/day; and cheese 14.6 (4.8 – 26.9) g/day (**Table 4.2**). Almost two-thirds of high yoghurt consumers were women. Participants of non-white ethnic background were 2-4 times less likely to consume dairy products than participants of white background. Yoghurt and cheese consumption were positively correlated with higher socio-economic status, educational level and income and negatively correlated with likelihood of being current smokers.

Among dietary consumption levels, overall, low-fat dairy consumption was approximately six times higher than that of high-fat dairy (**Table 4.3**). For dairy types, low-fat milk and yoghurt consumption were higher than consumption of the full-fat alternatives, but high-fat cheese consumption was higher than low-fat cheese consumption. The consumption of low-fat milk, fruit, vegetables, whole-grain cereals and fish were higher in the highest quartile of yoghurt consumption, whereas consumption of alcohol, sugar-sweetened beverages and processed meat was lower compared to non-consumers. For the highest quartile of milk consumption, lower consumption of fruit and vegetables and higher consumption of potatoes, processed cereals, sweet snacks, coffee and tea were observed than among non-consumers of milk.

Table 4.2 Descriptive characteristics of **socio-demographic, behavioural and clinical factors**† for the top and bottom categories of milk, yoghurt and cheese consumption (g/d), as well as in the total sample of the Fenland study

			Milk ‡		Yoghurt ‡		Cheese ‡	
Mean (SD) consumption (g/d)		**Overall (%)**	293 (146 - 439)		35.3 (8.8 - 71.8)		14.6 (4.8 - 26.9)	
			0 g/d	585 - 732 g/d	0 g/d	99.5 - 1,134 g/d	0 g/d	26.8 - 284.6 g/d
Participants (N)		12,065	921	1,490	2,787	3,014	779	3,028
Socio-demographic factors								
Age (years)		48.8 (42.7 - 54.7)	49.6 (42.8 - 55.1)	48.0 (42.2 - 54.0)	48.6 (42.4 - 54.6)	50.0 (43.7 - 55.2)	48.8 (42.4 - 55.2)	48.1 (41.9 - 54.3)
Sex (ref. Men)	Women	53.8	63.7	38.7	39.5	65.7	49.0	60.2
Ethnicity § (ref.White)	Non-white	2.9	4.9	1.4	3.3	1.8	7.9	1.6
Educational level § (ref. Low)	Medium	46.2	41.8	48.0	49.0	46.7	48.1	41.9
	High	33.9	40.7	27.5	25.0	33.5	23.9	42.1
Age completing education (years) §		18.0 (16.0 - 21.0)	18.0 (16.0 - 22.0)	17.0 (16.0 - 21.0)	17.0 (16.0 - 20.0)	18.0 (16.0 - 21.0)	16.5 (16.0 - 19.0)	18.0 (16.0 - 22.0)
Socio-economic status (based on occupation) ‖(ref. Low)	Medium	19.8	19.7	14.6	16.4	23.1	21.2	19.8
	High	53.3	58.3	47.6	46.1	52.5	42.7	58.4
Income § (ref. <£20,000)	£20,000 - 40,000	35.4	32.6	37.5	37.8	37.3	38.5	34.4
	>£40,000	50.8	52.0	47.3	45.1	49.6	41.9	52.1
Marital status # (ref. Single)	Married	81.4	75.1	80.5	79.9	80.8	77.1	80.2
	Widowed / Separated	9.5	10.3	9.9	9.3	10.3	9.6	9.8
Lifestyle factors								
Smoking status § (ref. Never smoker)	Former smoker	33.3	35.6	31.8	32.0	34.0	27.6	34.3
	Current smoker	12.3	13.3	17.8	20.1	7.6	13.5	11.1
Smoking (pack-years) ‖		0 (0 - 2,376)	0 (0 - 2,696)	0 (0 - 3,701)	0 (0 - 4,362)	0 (0 - 1,414)	0 (0 - 2,725)	0 (0 - 2,192)
Physical activity energy expenditure (kJ/kg/d) §		50.7 (37.6 - 66.5)	49.9 (38.1 - 63.5)	55.4 (40.8 - 72.4)	51.2 (37.7 - 67.8)	50.5 (37.0 - 65.4)	49.4 (35.7 - 65.7)	52.0 (38.8 - 67.6)
Energy intake (kj/d)		1,851 (1,524 - 2,265)	1,600 (1,312 - 2,030)	2,140 (1,733 - 2,613)	1,768 (1,433 - 2,180)	1,932 (1,603 - 2,345)	1,681 (1,360 - 2,081)	2,067 (1,691 - 2,507)
BMI (kg/m2)		26.2 (23.6 - 29.4)	25.8 (23 - 29.1)	26.7 (24.2 - 29.9)	26.6 (23.8 - 29.6)	26.1 (23.5 - 29.3)	26.5 (23.8 - 29.6)	25.8 (23.2 - 29.1)
Medications / Supplements								
Lipid-lowering medication § (ref. No)	Yes	4.0	4.0	3.0	4.7	4.0	4.4	3.3
Anti-hypertensive medication § (ref. No)	Yes	7.4	7.4	6.3	8.2	7.1	7.8	6.9
Hormonal therapy § (ref. No for women / Men)	Yes	2.8	3.4	1.9	2.0	3.5	2.7	3.2
Dietary supplements § (ref. No)	Yes	41.2	49.5	40.5	33.5	47.7	39.5	44.5

†Continuous variables are presented as median (interquartile range) and categorical variables are presented as column percentages
‡Five categories; Milk: Non-consumers, 146 g/d, 293 g/d, 439 g/d, 585 or 732 g/d; Yogurt: Non-consumers and quartiles within consumers; Cheese: Non-consumers and quartiles within consumers
§Percentage of missing values < 5% with a total of 28.8% of missing values when accounting for non-overlapping missing values across all the variables
‖Percentage of missing values 5-15% with a total of 28.8% of missing values when accounting for non-overlapping missing values across all the variables
#Percentage of missing values 15-25% with a total of 28.8% of missing values when accounting for non-overlapping missing values across all the variables

Table 4.3 Descriptive characteristics of **dietary factors** † for the top and bottom categories of milk, yoghurt and cheese consumption (g/d), as well as in the total sample of the Fenland study

		Milk ‡		**Yoghurt** ‡		**Cheese** ‡	
	Overall (%)	293 (146 - 439)		35.3 (8.8 - 71.8)		14.6 (4.8 - 26.9)	
		0 g/d	585 - 732 g/d	0 g/d	99.5 - 1,134 g/d	0 g/d	26.8 - 284.6 g/d
Participants (N)	12,065	921	1,490	2,787	3,014	779	3,028
Dairy products (g/d)							
Full-fat milk	0.0 (0.0 - 0.0)	0.0 (0.0 - 0.0)	0.0 (0.0 - 0.0)	0.0 (0.0 - 0.0)	0.0 (0.0 - 0.0)	0.0 (0.0 - 0.0)	0.0 (0.0 - 0.0)
Low-fat milk	293.0 (146.0 - 293.0)	0.0 (0.0 - 0.0)	585.0 (585.0- 585.0)	146.0 (132.3 - 293.0)	293.0 (146.0 - 397.7)	265.5 (132.2 - 293.0)	293.0 (146.0 - 293.0)
Full-fat yoghurt	0.0 (0.0 - 8.8)	0.0 (0.0 - 8.8)	0.0 (0.0 - 8.8)	0.0 (0.0 - 0.0)	0.0 (0.0 - 17.6)	0.0 (0.0 - 8.8)	0.0 (0.0 - 8.8)
Low-fat yoghurt	17.6 (0.0 - 54.2)	8.8 (0.0 - 54.2)	17.6 (0.0 - 54.2)	0.0 (0.0 - 0.0)	99.5 (99.5 - 126.0)	8.8 (0.0 - 54.2)	54.2 (0.0 - 99.5)
High-fat cheese	14.6 (4.8 - 14.6)	14.6 (2.4 - 14.6)	14.6 (4.8 - 14.6)	14.6 (2.4 - 14.6)	14.6 (4.8 - 14.6)	0.0 (0.0 - 0.0)	26.9 (14.6 - 34.0)
Low-fat cheese	0 (0 - 4.2)	0 (0 - 4.2)	0 (0 - 4.2)	0.0 (0.0 - 0.0)	0 .0 (0.0 - 8.4)	0.0 (0.0 - 0.0)	4.2 (0.0 - 25.8)
Cream	0 (0 - 2.8)	0 (0 - 1.4)	0 (0 - 2.8)	0 (0 - 1.4)	0.0 (0.0 - 2.8)	0.0 (0.0 - 0.0)	0.0 (0.0 - 2.8)
Butter	1.4 (0.0 - 8.2)	1.3 (0.0 - 7.6)	1.4 (0.0 - 8.6)	1.4 (0.0 - 8.0)	0.9 (0.0 - 7.9)	0 (0 - 4.3)	2.8 (0.0 - 10.0)
Ice-cream	5.7 (0.0 - 5.7)	5.7 (0.0 - 5.7)	5.7 (0.0 - 11.3)	5.7 (0.0 - 5.7)	5.7 (0.0 - 11.3)	5.7 (0.0 - 5.7)	5.7 (0.0 - 11.3)
Low-fat fermented dairy products	39.5 (8.8 - 99.5)	21.8 (0.0 - 88.8)	34.9 (8.8 - 80.0)	0.0 (0.0 - 0.0)	126.0 (108.4 - 139.0)	8.8 (0.0- 54.2)	63.0 (17.6 - 125.3)
Fermented dairy products	54.2 (22.0 - 104.3)	38.2 (14.6 - 103.4)	49.9 (22.4 - 97.0)	14.6 (4.8 - 14.6)	139.4 (123.0 - 160.0)	8.8 (0.0- 54.2)	91.2 (48.2 - 146.9)
High-fat dairy products	26.0 (14.6 - 45.3)	20.3 (9.5 - 34.5)	27.4 (14.6 - 56.4)	24.6 (12.7 - 46.7)	26.0 (14.6 - 44.0)	7.9 (1.4 - 31.8)	40.4 (29.0 - 62.3)
Low-fat dairy products	301.8 (159.0 - 439.0)	21.8 (0.0 - 88.8)	639.2 (585.0 - 728.6)	154.4 (136.5 - 297.2)	401.4 (272.0 - 538.5)	293.0 (132.3 - 427.8)	318.0 (180.9 - 447.8)
Total dairy products	344.0 (220.5 - 478.1)	55.5 (23.0 - 117.2)	681.8 (627.7 - 759.6)	303.0 (162.1 - 408.7)	436.2 (306.0 - 570.5)	303.2 (160.5 - 453.8)	382.5 (259.2 - 508.4)
Other food groups (g/d)							
Fruits	203.7 (109.5 - 323.2)	208.5 (106.0 - 333.9)	195.1 (98.1 - 318.7)	139.0 (57.7 - 253.7)	271.6 (170.4 - 405.7)	176.6 (84.0 - 307.7)	238.8 (135.2 - 368.3)
Vegetables	234.8 (166.6 - 319.4)	261.7 (177.7 - 357.7)	225.1 (149.8 - 309)	198.8 (134.6 - 278.6)	262.8 (191.4 - 360.3)	209.5 (130.5 - 303.5)	263.6 (189.6 - 357)
Potatoes	80.2 (54.2 - 116.8)	71.4 (35.2 - 92.2)	89 (62.6 - 125.6)	89.0 (59.1 - 125.6)	80.2 (53.8 - 116.8)	80.2 (44.0 - 125.6)	80.2 (53.8 - 119.5)
Legumes	54.2 (31.4 - 81.1)	53.6 (28.4 - 87.6)	57.2 (34.8 - 83.7)	52.5 (28.4 - 79.7)	56.9 (33.3 - 82.7)	54.5 (27.4 - 81.3)	59.5 (35.3 - 87.4)
Processed cereals	78.9 (46.5 - 122.1)	66.6 (35.3 - 117)	82.9 (53.0 - 130.4)	74.7 (42.7 - 116.2)	76.4 (45.8 - 122.1)	64.3 (36.6 - 111.6)	89.7 (53.1 - 134.7)
Whole-grain cereals	55.7 (19.7 - 120.1)	52.9 (17.8 - 132)	58.8 (18.9 - 129.8)	30.0 (9.1 - 87.5)	71.3 (28.0 - 145.6)	39.8 (12.9 - 114.9)	72.4 (28.0 - 143.4)
Poultry and eggs	53,0 (23.1 - 56.5)	49.5 (21.5 - 56.5)	53.0 (23.1 - 71)	49.5 (23.1 - 56.5)	53.0 (23.1 - 71)	49.5 (21.5 - 56.5)	53.0 (23.1 - 71.0)
Red meat	43.2 (21.7 - 69.3)	36.2 (12.9 - 63)	49.6 (26.8 - 80.5)	46.2 (24.2 - 74.0)	39.7 (20.3 - 64.5)	42.6 (20.3 - 71.5)	39.2 (19.3 - 67.4)
Processed meat	26.5 (13.2 - 42.7)	20.4 (6.2 - 35.1)	30.7 (16.5 - 49.1)	30.1 (15.5 - 50.4)	24.4 (11.3 - 39.9)	24.6 (10.6 - 46)	24.9 (10.4 - 43.2)
Fish	35.4 (20.2 - 57.1)	35.4 (19.2 - 65.1)	36.3 (23.4 - 58.7)	30.5 (16.1 - 48.3)	39.4 (24.0 - 65.1)	32.1 (16.1 - 54.6)	36.3 (20.2 - 60.2)
Sauces	18.0 (8.7 - 30.1)	15.9 (6.6 - 27.9)	19.2 (9.5 - 30.9)	17.1 (6.6 - 29.5)	19.3 (9.7 - 31.9)	13.8 (4.5 - 27.9)	20.5 (10.8 - 33.4)
Margarine	4.3 (0.0 - 10.0)	4.3 (0.0 - 10.0)	7.1 (0.7 - 10.0)	4.3 (0.0 - 10.0)	4.3 (0.0 - 10.0)	4.3 (0.0 - 10.0)	5.0 (0.0 - 10.0)
Nuts	2.1 (0.0 - 4.2)	2.1 (0.0 - 12.9)	2.1 (0.0 - 4.2)	2.1 (0.0 - 4.2)	2.1 (0.0 - 4.2)	2.1 (0.0 - 4.2)	2.1 (0.0 - 4.2)
Sweet snacks	74.5 (43.4 - 121.2)	59.0 (32.8 - 104.3)	86.7 (50.8 - 138.2)	76.0 (43.3 - 128.5)	69.8 (40.2 - 114.1)	69.7 (34.9 - 120.4)	79.7 (46.7 - 131.1)
Sugar-sweetened beverages	14.0 (0.0 - 40.0)	2.8 (0.0 - 31.6)	17.2 (0.0 - 45.6)	16.8 (0.0 - 68.0)	5.6 (0.0 - 40.0)	14 (0.0 - 45.2)	14 (0.0 - 40)

Table 4.3 (continued)

	Overall (%)	**Milk ‡** 293 (146 - 439)		**Yoghurt ‡** 35.3 (8.8 - 71.8)		**Cheese ‡** 14.6 (4.8 - 26.9)	
		0 g/d	585 - 732 g/d	0 g/d	99.5 - 1,134 g/d	0 g/d	26.8 - 284.6 g/d
Artificially sweetened beverages	0.0 (0.0 - 28.0)	0.0 (0.0 - 28.0)	0.0 (0.0 - 28.0)	0.0 (0.0 - 28.0)	0.0 (0.0 - 86.0)	0.0 (0.0 - 28.0)	0.0 (0.0 - 28.0)
Fruit juice	16.8 (8.4 - 94.8)	16.8 (8.4 - 94.8)	16.8 (8.4 - 94.8)	16.8 (0.0 - 51.6)	51.6 (8.4 - 120.0)	16.8 (0.0 - 94.8)	51.6 (8.4 - 120.0)
Regular coffee	190.0 (13.3 - 47.05)	150.1 (0.0 - 475.0)	190.0 (13.3 - 47.05)	190.0 (0.0 - 475.0)	190.0 (0.0 - 475.0)	81.7 (0.0 - 475.0)	190.0 (13.3 - 475)
Decaffeinated coffee	0.0 (0.0 -13.3)	0.0 (0.0 -13.3)	0.0 (0.0 -13.3)	0.0 (0.0 -0.0)	0.0 (0.0 -13.3)	0.0 (0.0 -0.0)	0.0 (0.0 -13.3)
Tea	475.0 (150.1 - 855.0)	150.1 (0.0 - 475.0)	475.0 (150.1 - 855.0)	475.0 (81.7 - 855.0)	475.0 (190.0 - 475.0)	475.0(81.7 - 855.0)	475.0 (150.1 - 855.0)
Alcoholic beverages	73.9 (17.5 - 177.6)	53.8 (8.8 - 144.6)	59.4 (10.4 - 177.6)	67.1 (10.4 - 222.6)	55.4 (10.4 - 141.3)	32.4 (0.0 - 127.3)	75.3 (18.2 - 179.2)
Plasma vitamin C (-mol/l)	69.4 (56.1 - 82.2)	72.6 (59.2 - 87.1)	64.9 (49.5 - 78.5)	63.2 (47.3 - 77.8)	73.1 (61.2 - 85.2)	67.3 (53.2 - 79.7)	70.9 (58.4 - 83.4)

†Variables are presented as median (IQR)
‡Five categories; Milk: Non-consumers, 146 g/d, 293 g/d, 439 g/d, 585 or 732 g/d; Yogurt: Non-consumers and quartiles within consumers; Cheese: Non-consumers and quartiles within consumers

4.4.2 Dairy products and body composition

Habitual milk consumption was significantly associated with higher BMI with each additional glass of milk/day associated with 0.26 kg/m^2 (95% CI: 0.16, 0.36) higher BMI (**Table A.4**). Similar patterns were observed for low-fat milk and low-fat and total dairy products. Other dairy subtypes were not significantly associated with BMI.

Dairy products were not associated with the percentage of body fat (data not shown), total fat mass, or peripheral body fat mass (**Figure 4.1**, **Table A.5**), nor were any significant associations found between any dairy type and waist circumference or the ratio of waist to hip circumference, as proxies for fat mass distribution.

Low-fat dairy consumption was associated with 2.6% (-3.91, -1.23) lower VAT/SCAT ratio. A similar association was observed for VAT.

Habitual milk consumption was significantly associated with 0.33 kg (0.19, 0.46) higher lean mass per one glass/day (Figure 4.1, Table A.5). Both full- and low-fat milk showed similar associations (Figure 4.1, Table A.5). The association was partly attenuated when adjusted for height, but was still significant (0.18 kg/glass of milk/day; 0.10, 0.27).

Effect modification by BMI was suggested for the association between high-fat dairy and appendicular lean mass (p-interaction=0.0001). Among non-overweight adults (BMI<25 kg/m^2), appendicular lean mass was lower by 0.11 kg (-0.22, -0.001) per one serving of high-fat dairy, but associations were not observed among overweight (0.05 kg; 95% CI: -0.07 to 0.18 kg) or obese adults (0.02 kg; 95% CI -0.16 to 0.20 kg).

No significant association was observed for cheese, butter, ice-cream or any other dairy group and any marker of body composition.

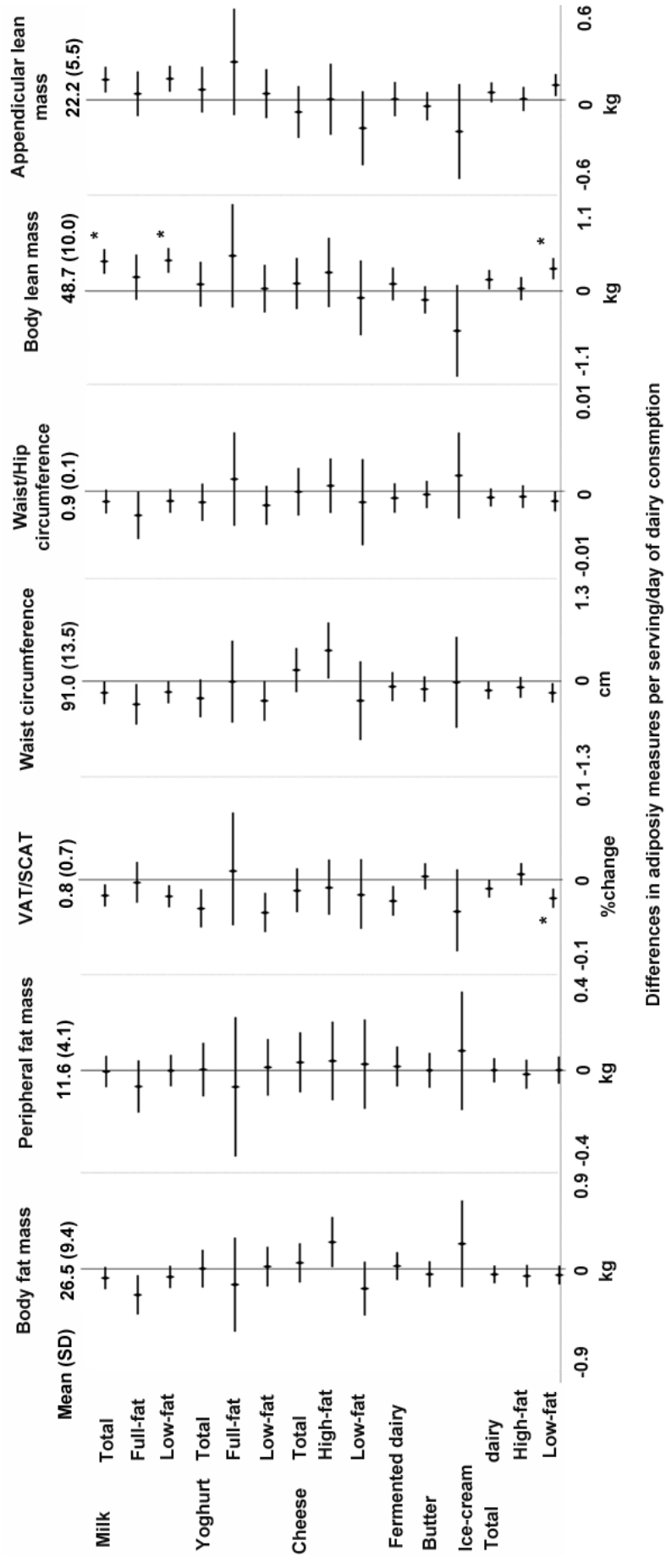


Fig. 4.1 Adjusted associations of types of dairy consumption (servings/day) with **markers of body composition**. Statistically significant associations after false discovery rate corrections are marked with an asterisk. VAT: Visceral Adipose Tissue, SCAT: Subcutaneous Adipose Tissue

4.4.3 Dairy products and metabolic markers

No association was found between any type of dairy products and the ratio of total to HDL-C (**Figure 4.2**, **Table A.6**). High-fat dairy consumption was weakly positively associated with total cholesterol [0.06 mmol/L (95% CI:0.03, 0.09) per serving/day] and with LDL-C [0.05 mmol/l (0.03, 0.08)]. Of subtypes of high-fat dairy, butter consumption showed similar associations. Slight attenuation of the association was observed after adjustment for vitamin A, total, mono-unsaturated and saturated fat (**Table A.7**). Low-fat dairy and milk consumption were weakly inversely associated with HDL-C [-0.02 mmol/l (-0.03, -0.01) for both]. None of the dairy types were significantly associated with triglycerides, apoA1, apoB or NEFA. Further adjustment for nutrients did not materially alter these findings.

Consumption of different dairy types was not significantly associated with glycaemic markers, hepatic markers, blood pressure measures, CRP, adiponectin, or the metabolic risk z-score (**Figures 4.3- 4.5**, **Tables A.8- A.10**).

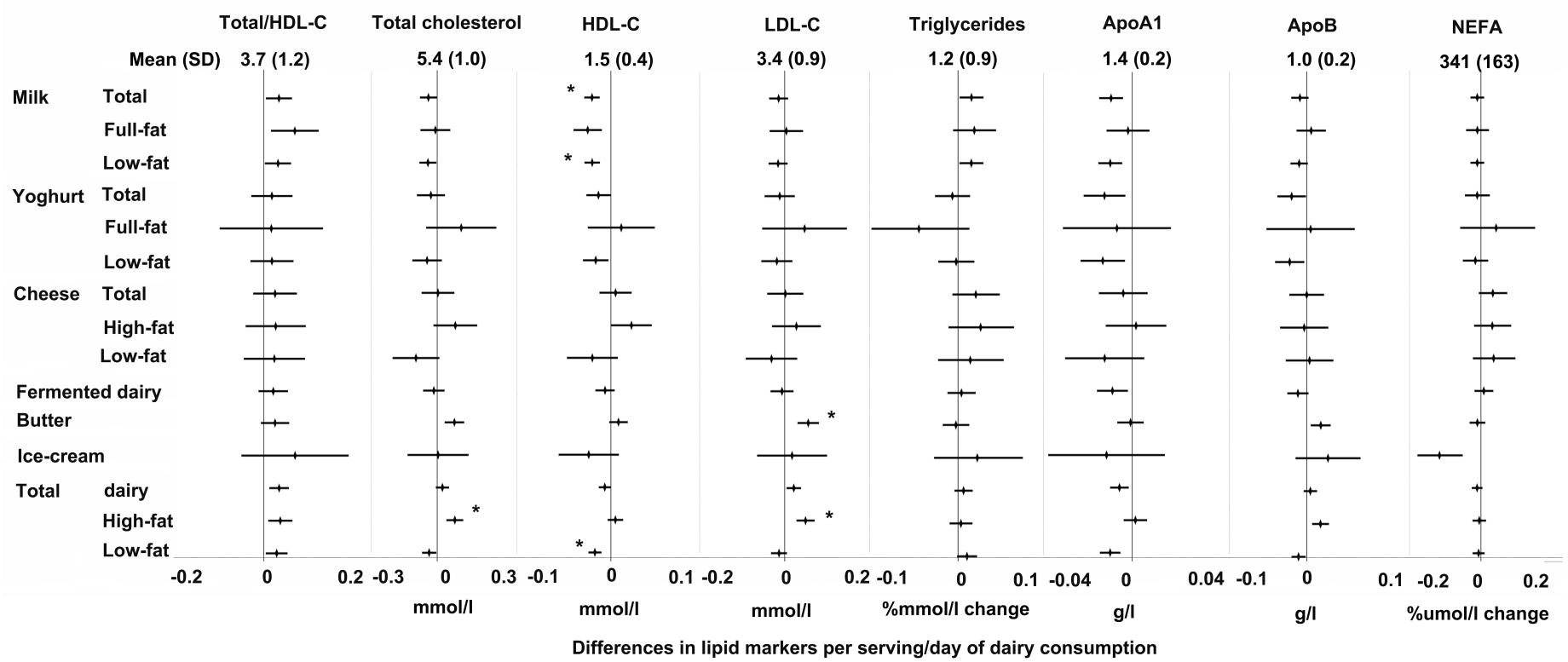


Fig. 4.2 Adjusted associations of types of dairy consumption (servings/day) with **lipid markers**. Statistically significant associations after false discovery rate corrections are marked with an asterisk. HDL: High-Density Lipoprotein, LDL: Low-Density Lipoprotein, Apo: Apolipoprotein, NEFA: Non-esterified fatty acids

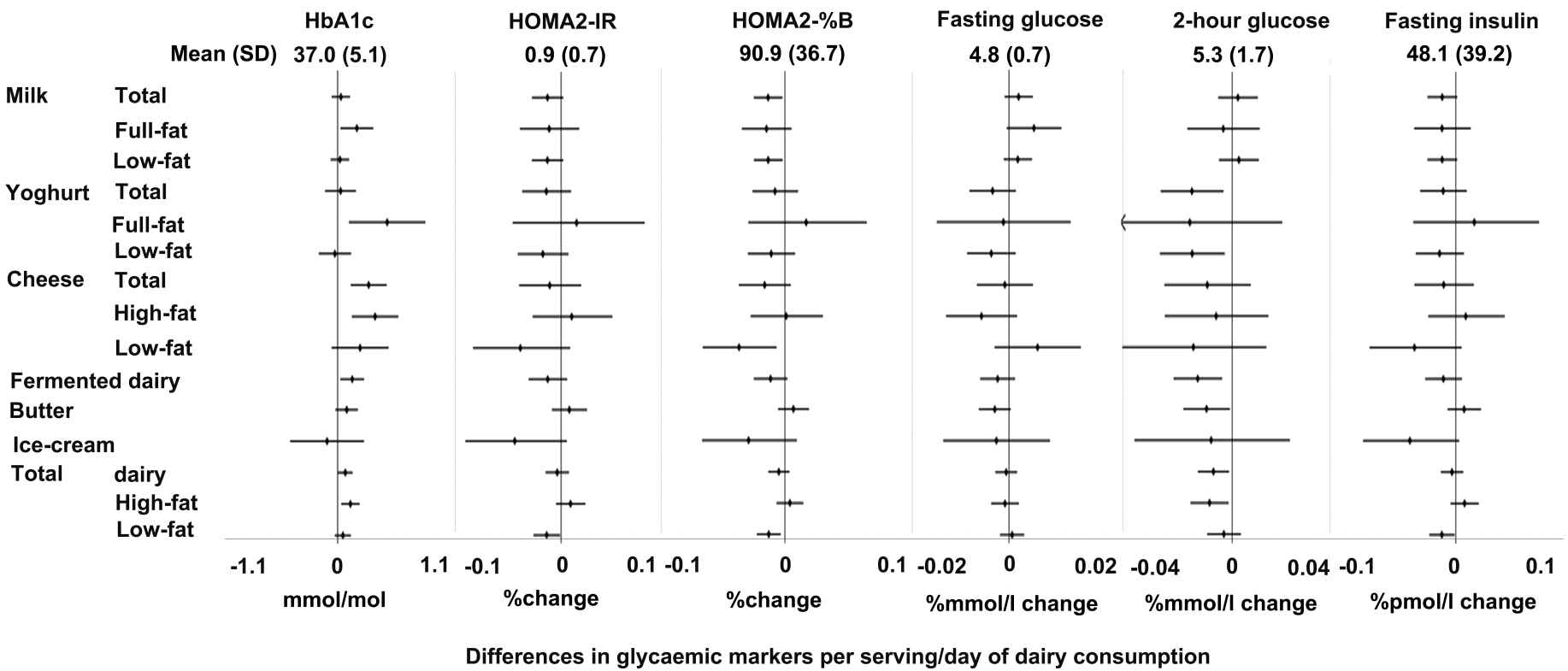


Fig. 4.3 Adjusted associations of types of dairy consumption (servings/day) with **glycaemic markers**. Statistically significant associations after correction for false discovery rate are marked with an asterisk. HbA1c: Haemoglobin A1c, HOMA-IR: Homeostasis Model Assessment for Insulin Resistance, HOMA-%B: Homeostasis Model Assessment for Beta cell function

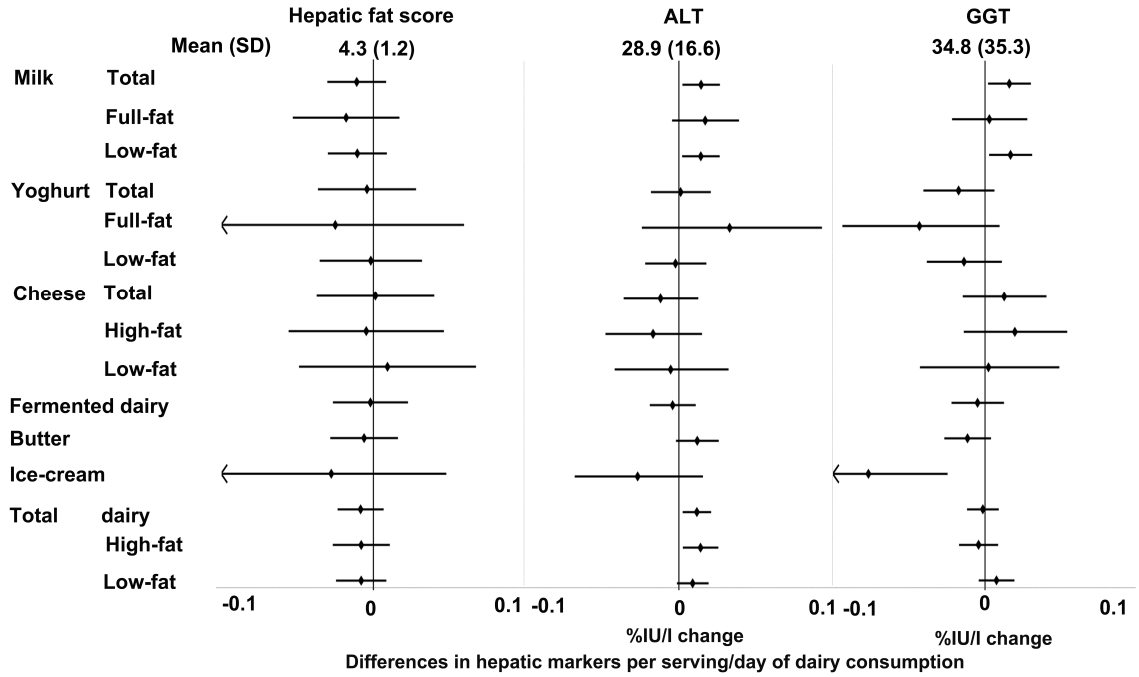


Fig. 4.4 Adjusted associations of types of dairy consumption (servings/day) with **markers of hepatic function**. Statistically significant associations after correction for false discovery rate are marked with an asterisk. ALT: Alanine Transaminase, GGT: Gamma-Glutamyl Transferase

4.4.4 Additional analyses

Results were not altered when analyses with 10 imputed datasets or complete-case analyses were done (data not shown). There was no indication of a non-linear association from analyses with restricted cubic splines or categorical exposures after correction for multiple testing. [**Figure 4.6** as example, shows the associations between 3 main exposures (milk, yoghurt and cheese) and 3 main outcomes (HbA1c, total-to-HDL-C ratio and CRP respectively). The *p*-value for non-linearity was <0.05, but this was not statistically significant after multiple test correction.

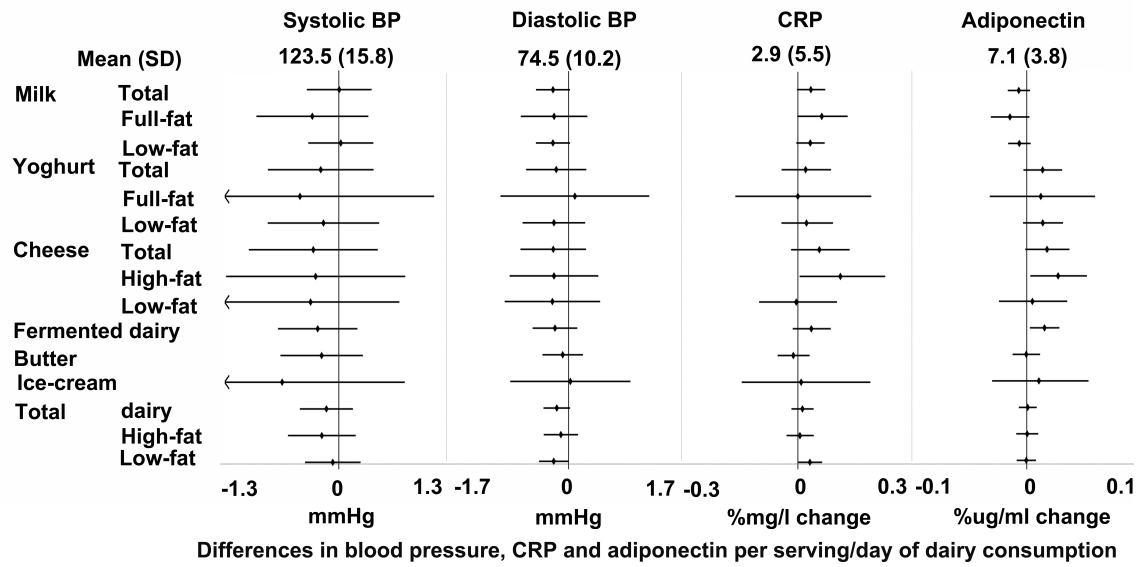


Fig. 4.5 Adjusted associations of types of dairy consumption (servings/day) with other physiologic markers including **systolic and diastolic blood pressure, CRP and adiponectin**. Statistically significant associations after correction for false discovery rate are marked with an asterisk. CRP:C-reactive protein

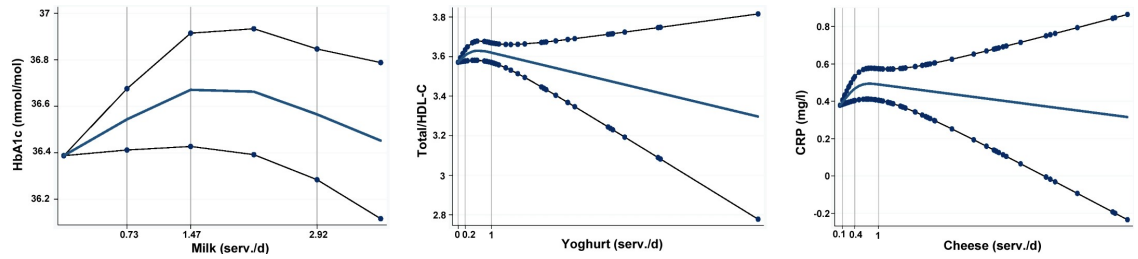


Fig. 4.6 **Non-linear associations** of dairy products with HbA1c, ratio of total to HDL-C and CRP. The plots were selected on the basis of nominally significant non-linearity. **A**. milk and HbA1c (p-nonlinear=0.015), **B**. yoghurt and ratio of total to HDL-C (p-nonlinear=0.015), and **C** cheese and CRP (p-nonlinear=0.015). HbA1c:Haemoglobin A1c, HDL-C:High-density lipoprotein cholesterol, CRP:C-reactive protein

4.5 Discussion

4.5.1 Summary of results

This study resulted in two novel findings. First, habitual daily consumption of one serving of low-fat dairy products was associated with a 3% lower ratio of VAT to SCAT as a marker of fat mass distribution, a measure which is associated with diabetes risk independently of total fat mass[174]. Second, habitual daily consumption of 1 glass of milk was associated with a 0.33 kg higher body lean mass.

4.5.2 Markers of adiposity

There are no previous studies on the association between dairy consumption and VAT/SCAT. An RCT showed a reduction in VAT among those consuming 6-7 servings of dairy products per day compared to those consuming less than 4 servings/day[78]. A cross-sectional study of twins reported an inverse association between low-fat fermented dairy products and VAT[175], but I have not identified any studies of the association between dairy consumption and SCAT.

Although total dairy consumption has been consistently associated with a lower body fat mass in RCTs[60–62], the number of studies for dairy subtypes is limited. I found no significant associations between any dairy type and total or peripheral fat mass or waist circumference and waist-to-hip ratio as proxies for fat mass distribution. The direction of the associations I observed between low-fat dairy consumption and total body fat mass, and waist circumference was consistent with that observed in previous studies and with that observed for VAT/SCAT in the current analysis. The higher dairy amounts used in trials compared to the consumption levels reported in observational studies could partly explain the lack of significance in certain associations.

Although no mechanism has been reported for an inverse association between low-fat dairy consumption and VAT/SCAT, a plausible explanation could be that the effects of dairy nutrients on fat mass are more pronounced in VAT than in SCAT. For example, VAT was reported to be more metabolically active with a more efficient glucose uptake than SCAT[176]. Relevant underlying mechanisms for these associations include higher lipolysis and lower lipogenesis due to lower parathyroid hormone and 1,25-$(OH)_2$ vitamin D levels related to calcium intake[177], and possibly a combination of higher satiety with lower appetite related to dairy protein intake[178–180]. Potential mechanisms for the associations of dairy products with body weight and composition are discussed in more detail in Chapter 9.

With these current analyses and specifically the observed positive association between milk consumption and body lean mass, we extend the previous understanding on the positive association between total dairy consumption and total lean mass[61] to include specific dairy subtypes. Dairy products and mainly milk have been consistently associated with bone health due to their nutrient content including calcium, phosphorous, vitamin D and protein, which might partly explain the positive association with lean mass[181]. Another potential mechanism is the increasing effect of milk on growth hormone[182, 183], which has been associated with a higher lean mass through a higher bone mineral density and muscle mass[184].

4.5.3 Markers of metabolic risk

Concerning lipid markers, I did not find an association between any dairy type and the ratio of total to HDL-C, though I observed a positive association between high-fat dairy and mainly butter consumption and total cholesterol and an inverse association of low-fat dairy and mainly milk consumption with HDL-C. The association of dairy consumption with the ratio of total to HDL-C has not been extensively studied previously, except for butter, which increased the ratio, and total and HDL-C separately when butter replaced hard margarine[87], but butter had no effect when it replaced olive oil or walnuts in RCTs[89, 95, 105]. Null associations have been reported between low-fat and high-fat dairy consumption and HDL-C[63].

My findings of a positive association of high-fat dairy and mainly butter consumption with LDL-C were consistent with findings of feeding trials, in which butter increased LDL-C when it substituted margarine[87], monounsaturated[89, 95, 105] or polyunsaturated[105] fat. However, in a meta-analysis of RCTs, low- and high-fat dairy consumption did not show an effect on LDL-C, but it is not clear whether butter was included in the high-fat dairy group in all the studies[63].

I found no association of any dairy type with triglycerides, ApoA1, ApoB or NEFA. With the exception of butter consumption which seems to have a neutral effect on triglyceride levels when it substitutes other fat sources[89, 105], the evidence is sparse for associations between other dairy types and the rest of the lipid markers.

As in previous research, associations between dairy consumption and glycaemic markers were null in the current study including that between low- or high-fat dairy consumption and fasting blood glucose or HOMA-IR in RCTs[63]. Evidence on associations between other dairy types and glycaemic markers is sparse. Dairy products have an insulinotropic effect in short-term trials due to their protein content, but this effect does not seem to be maintained in the long-term[185], which may explain the different findings between trials and prospective cohort studies.

I did not observe any association between dairy types and SBP or DBP, which is in line with prior findings from a meta-analysis of seven RCTs, which reported a neutral effect of low- and high-fat dairy consumption on SBP and DBP[63].

I also did not find any significant association of dairy consumption with hepatic fat, markers of hepatic function, CRP or adiponectin. While there is little research about dairy consumption and hepatic outcomes, our finding was consistent with a 6-month RCT showing no effect of milk consumption on hepatic fat in comparison to water[186]. The current findings for CRP and adiponectin were in agreement with a meta-analysis of six RCTs[63].

4.5.4 Strengths and limitations

This study has several strengths including its large sample size (n=12,065) and the inclusion of several dairy subtypes and diverse cardio-metabolic markers, which allowed the investigation of many potential pathways for cardio-metabolic disease. By employing DEXA and ultrasonography, we were able to use more accurate methods to assess VAT and SCAT than previous studies that used waist circumference and waist-to-hip ratio as proxies of central adiposity[74]. Our statistical approaches were thorough including the adjustment for important potential confounders including objectively measured physical activity, the derivation of estimates robust to outliers, and the handling of missing data with multiple imputation.

This study also has limitations. The cross-sectional design increases the risk of reverse causation and limits inference for causal association. Although the questionnaire used was assessed for validity in a similar population[158] and we adjusted for BMI as an established factor of dietary misreporting, we cannot exclude the possibility of error due to dietary misreporting caused by the use of self-reported methods of dietary assessment. In addition, the use of the food frequency questionnaire does not allow the discrimination of dairy types e.g.cheese or yoghurt to subtypes based on processing. This leads to loss of information, as there is a high heterogeneity among subtypes of yoghurt and cheese based on the way they are processed and produced with very different food matrices. Consumption of high-fat dairy products has a limited range and lower levels compared to low-fat dairy products, which might compromise the power to detect associations for high-fat dairy. Finally, although we adjusted our models for many potential confounders, we cannot exclude the possibility of residual confounding.

4.5.5 Conclusion

The novel finding of an inverse association between low-fat dairy products and VAT/SCAT suggests abdominal obesity as a potential pathway for the association of dairy consumption with cardio-metabolic disease. The association between milk consumption and a higher body lean mass could also be a potential explanation for the overall positive metabolic associations of dairy consumption. These findings are important for generating robust hypotheses and should be confirmed and further investigated in prospective studies, clinical trials and mechanistic studies.

Chapter 5

Associations of dairy consumption with cardio-metabolic markers: a prospective analysis in the EPIC Norfolk Study, UK

Summary

Background and aims: Accumulating evidence suggests that some types of dairy products are associated with lower cardio-metabolic risk. However, relevant pathways have not yet been investigated in long-term epidemiological studies on habitual dairy consumption and its relationship with metabolic risk factors. The aim of this study was to investigate prospective associations of habitual dairy consumption with markers of metabolic risk and adiposity.

Methods: We examined associations of changes in dairy consumption (servings/day) with parallel changes in metabolic markers using multiple linear regression in 15,612 adults followed up from 1993-1997 aged 40-78 years at baseline in the EPIC-Norfolk study, UK.

Results: An increase in low-fat fermented dairy products was associated with a lower increase in body weight and BMI over an average follow-up of 3.7 years. For example, a change in yoghurt consumption was associated with a lower increase in body weight by 0.23 kg (95% CI: -0.46, -0.01). An increase in high-fat dairy products (-0.14±1.18 servings/day) was associated with higher increases in body weight and BMI [b=0.13 (0.05, 0.21) and 0.04 (0.01, 0.07) respectively]. An increase in total dairy consumption (0.07±1.3) was positively associated with an increase in total cholesterol [0.02 mmol/l (0.003, 0.04)]. A similar association was observed between high-fat dairy (including butter) consumption and LDL-C, while increasing low-fat dairy consumption (0.06±1.02)

was inversely associated with increasing total and LDL-C. A change in full-fat milk (-0.11±0.62) was positively associated with HbA1c (p=0.027).

Conclusion: Our results support the differential associations of dairy types with metabolic pathways including associations of increasing low-fat fermented dairy products with lesser increase in adiposity and of increasing high-fat dairy products with greater increase in circulating lipids.

What is already known

- Dairy products have been associated with no or a lower risk of cardio-metabolic disease endpoints.
- Evidence from randomised controlled trials supports favourable effects of total dairy consumption on body weight and composition under conditions of energy restriction, increasing effects of butter, but more favourable effects of cheese on blood lipipds, no effects of total dairy consumption on glycaemia and no effects of total, high-fat or low-fat dairy consumption on blood pressure.
- Evidence on specific dairy types, effects on other cardio-metabolic markers such as inflammation and hepatic function and more long-term associations of habitual dairy consumption with cardio-metabolic markers is limited.

What this research adds

- Of all the cardio-metabolic markers examined, changes in dairy consumption were associated mainly with parallel changes in markers of adiposity and lipidaemia over 3.7 years of follow-up.
- An increase in low-fat fermented dairy products was associated with a lower increase in body weight and BMI. For example, a change in yoghurt consumption by 1 serving/day was associated with a lower increase in body weight by 0.23 kg.
- Increases of high-fat dairy consumption were associated with higher increase of total and low-density lipoprotein cholesterol (LDL-C), while increases of low-fat dairy consumption were associated with lower increases of total and LDL-C.

Publication

Trichia E, Imamura F, Khaw KT, Wareham NJ, Forouhi N G. Prospective associations of total and types of dairy products with markers of metabolic risk and adiposity: findings from the EPIC Norfolk study, UK (Manuscript under review by co-authors)

5.1 Previous evidence on prospective associations

Effects of total and types of dairy products on markers of metabolic risk and adiposity have been examined in several randomised controlled trials -RCTs- (Chapter 3). Even though RCTs are considered the gold standard for the investigation of causal associations, they also have some limitations (Chapter 3). For instance, most of these trials used mixed dairy products in their intervention, which makes it difficult to distinguish the net effects of specific dairy types. In addition, their average duration is limited to a few months. For example, the average duration of 37 RCTs included in a meta-analysis on the effects of dairy consumption on body weight and composition was 7.7 months[58]. It is of interest to examine long-term associations of habitual dairy consumption with cardio-metabolic markers, since cardio-metabolic disease is a chronic disease developing over several years. Furthermore, since cardio-metabolic disease is multifactorial, characterised by multiple biological pathways including body composition, lipidaemia, glycaemia, and blood pressure, it is of interest to investigate associations of total and types of dairy products with markers related to these pathways.

Some previous prospective cohort studies have investigated associations of habitual dairy consumption with cardio-metabolic markers, but results are sparse and inconsistent for specific dairy types (Chapter 3). Inconsistencies in findings might be related to differences in the definitions of dairy groups (total, low- and high-fat) and the use of the baseline consumption (n=9)[64–66, 68, 69, 71, 73, 80, 81], repeated measures (n=2)[67, 68] or change in dairy consumption (n=3)[66, 70, 72] across different studies. Associations of the parallel change (change across the same time points for both the exposure and the outcome) between dietary factors and anthropometric markers were shown to be closer to associations described in RCTs compared with associations between baseline diet and outcome change or associations between change in diet and prospective outcome change[187]. However, the number of studies assessing associations of parallel change is limited[66, 70, 72] especially for dairy types[70].

5.2 Study aims

The primary aim of this study was to investigate associations of changes in total and types of dairy products with parallel changes in markers of metabolic risk and adiposity, as potential pathways for the association of dairy products with cardio-metabolic disease. The secondary aim of this study was to investigate associations of the repeated measures of total and types of dairy consumption with prospective repeated measures of markers of metabolic risk and adiposity.

5.3 Methods

5.3.1 Study design and population

We evaluated data from the European Prospective Investigation into Cancer and Nutrition in Norfolk (EPIC-Norfolk) study in the UK. More information on the study design of the whole EPIC cohort[188] and Norfolk, UK arm[189] was previously published. Recruitment was done through general practices from 1993 to 1997 at baseline with exclusion criteria of having terminal or malignant disease, inability to attend a local clinic, alcoholism, psychiatric disorder, inadequate command of English or blindness. Participants were invited for a first follow-up between 1998 and 2000 and a second follow-up between 2004 and 2011. At each time point, participants provided general information, completed a dietary assessment and measurement of several physiological markers. From an initial sample of 25,639, we evaluated 15,612 adults after exclusion of participants with no dietary data at baseline or follow-up (n=673), participants censored before the first follow-up assessment (n=8,507) and participants with implausible values of total energy intake [<800 and >4000 kcal/day for men, and <500 and >3,500 kcal/day for women[190] (n=847)]. We also excluded participants with outliers of the changes in dairy consumption and cardio-metabolic marker for each association examined (outside the range of 3 SD from the mean). Participant informed consent was obtained and the study was approved by the Norwich District Ethics Committee.

5.3.2 Dietary assessment

Diet was assessed at baseline and first follow-up (1998-2000) with a 130-item semi-quantitative food frequency questionnaire (FFQ). Validity of the FFQ was internally assessed at baseline against 7-day diet diaries[158]. The questionnaire ascertained habitual consumption of foods over the past year with nine frequencies ranging from "never or less than once/month" to "6 times per day"; and additional questions on the type and amount of milk consumed, details on breakfast cereals, fat for cooking, visible fat on meat, and dietary supplements. The correlation coefficient between the dairy consumption from the FFQ and that from the 7-day diary at baseline was 0.56 for milk, 0.57 for yoghurt, 0.33 for cheese and 0.54 for butter. Dietary data were processed with the FETA software[159] , assessed in servings/day and grouped as shown in Table 4.1[160]. We considered total milk, yoghurt and cheese as primary exposures .

5.3.3 Assessment of markers of metabolic risk and adiposity

As primary outcomes we considered changes in body mass index (BMI), waist circumference, the ratio of total to high-density lipoprotein cholesterol (HDL-C), haemoglobin A1c (HbA1c), systolic and diastolic blood pressure. BMI was defined as weight (digital

scale, Salter, UK) divided by height squared (free-standing stadiometer). Waist and hip circumferences were measured to the nearest 0.1cm with a D loop non-stretch fibreglass tape to standardised protocol. Body fat was measured with bioelectrical impedance (Tanita, UK) at the first and second follow-up. The average of two blood pressure measurements was used (Accutorr sphygmomanometer, Datascope, UK). Non-fasting blood was collected and lipid markers i.e. total cholesterol, HDL-C and triglycerides (RA 1000, Bayer Diagnostics, Basingstoke), and glycaemic markers i.e. HbA1c (Diamat ion exchange HPLC, Bio-Rad Laboratories, Hemel Hempstead, UK) were measured. Low-density lipoprotein cholesterol (LDL-C) was calculated using the Friedewald formula[164]. Trained research nurses did all the assessments and the blood sampling. As part of the secondary outcomes, the metabolic risk z-score was calculated as previously described[167] by taking averages of z-scores of blood pressure [average of the z-scores of systolic (SBP) and diastolic blood pressure (DBP)], HbA1c, waist circumference, -1 x HDL-C and log-transformed triglycerides.

5.3.4 Assessment of socio-demographic and lifestyle factors

Socio-demographic factors including educational level, marital status and occupation, medical history (disease status and medication use) and smoking status across the different time points and baseline physical activity were assessed with the Health and Lifestyle questionnaire[188]. Physical activity at follow-up was assessed with the EPIC physical activity questionnaire (EPAQ2) designed based on validation studies using heart rate monitoring after calibration with estimates derived from doubly labelled water[191].

5.3.5 Statistical analyses

The ratio of total to HDL-C, and triglycerides were not normally distributed and were log-transformed. Missing values of the dairy exposures and covariates were imputed using multiple imputation by chained equations (MICE) specified for continuous and categorical variables individually[168], whereby we generated five imputed datasets.

In our primary analysis, we investigated the association of the change in dairy consumption between baseline and the first follow-up with the parallel change of the markers using multiple linear regression models. This approach was selected, because in observational research it gave the results most biologically plausible and the closest to those from RCTs for associations between diet and weight change[187].

As a positive control analysis, to confirm the plausibility of associations, we first evaluated the prospective association of changes in butter consumption with changes in LDL-C , which is a well-established association from RCTs[87].

Accounting for availability of covariates and their plausibility as confounders, we developed regression models with three different levels of adjustment for potential con-

founders. The first included socio-demographic factors i.e. age, sex, educational level, age at completion of full-time education, occupational and marital status, physical activity level, smoking status, total energy intake, medications (lipid-lowering, anti-hypertensive and hormone replacement therapy), and follow-up duration. Mutually-exclusive dairy products were adjusted for. The second model additionally included major food groups and dietary supplement use. The third additionally included BMI unless the outcome was BMI or the metabolic risk z-score. Models included baseline levels of covariates and their changes if applicable. Baseline outcome values were not adjusted for to avoid collider bias which could be larger than bias due to confounding of a baseline outcome[192–194].

Accounting for possibilities of both type I and type II errors, we interpreted results before and after false-discovery rate correction (two-sided alpha=0.05)[172, 195].

In secondary analyses, we conducted a longitudinal analysis for the associations of the repeated measures of dairy consumption from baseline and the first follow-up with the repeated measures of the markers from the first and the second follow-up using linear mixed models. Although this analysis is not theoretically as valid as the primary analysis, the decision to also proceed with it was based on the assumption that the parallel change analysis might be under-powered upon the observation that the changes in some dairy types were very small.

Pre-specified analyses were conducted additionally to test interactions with age, sex and BMI for all the associations. We further conducted analyses to assess robustness of our primary findings. We applied inverse probability weighting to assess potential bias due to healthy survivor effects or effects due to censoring over the follow-up after deriving the probability of censoring with logistic regression[196]. To assess stability of results, we applied multiple imputation with 10 datasets and complete-case analysis; analyses excluding participants with prevalent type 2 diabetes and additionally with hypertension, hyperlipidaemia or cardiovascular disease; and analyses adjusting for the baseline outcome. For all analyses we used Stata 14.2 (College Station, TX: StataCorp LP, 2015).

5.4 Results

5.4.1 Descriptive characteristics

Participants were followed for a mean±SD of 3.7±0.7 years. The mean±SD consumption of milk, yoghurt and cheese at baseline was 1.7±0.8, 0.3±0.4 and 0.5±0.4 servings/day respectively. The mean±SD of their change over the follow-up was -0.06±0.70, 0.02±0.40 and -0.04±0.40 servings/day respectively for milk, yoghurt and cheese (**Table 5.1**).

Changes in markers of metabolic risk and adiposity and lifestyle characteristics are presented in **Tables 5.2- 5.3**. For example average increases in weight and waist circum-

ferences were by 1.3±4.0 kg and 0.8±5.5 cm. In addition, total energy intake decreased by 88.6±478.5 kcal.

Socio-demographic, lifestyle, clinical and dietary characteristics varied by participants consuming different types of dairy products (**Table 5.4**). Non-consumers of milk reported a lower consumption of fruit, processed and whole-grain cereals, sugar-sweetened beverages and tea, and higher amounts of artificially sweetened beverages. High yoghurt consumers were more frequently women, of higher educational level, less frequently current smokers and consumed more fruit, vegetables, whole-grain cereals, artificially sweetened beverages, fruit juice and decaffeinated coffee and lower amounts of potatoes, sweet snacks and alcohol. High cheese consumers were of higher educational level and consumed more processed and whole-grain cereals, red meat, fruit juice and alcohol.

Table 5.1 Descriptive characteristics of total and types of dairy products at baseline, first follow-up and the change between baseline and first follow-up in the EPIC-Norfolk study (n=15,612)

		Baseline		**1st follow-up**		**Change**	
Dairy products (servings/d)		**Mean (SD)**					
Milk	Full-fat	0.33	(0.78)	0.23	(0.66)	-0.11	(0.62)
	Low-fat	1.39	(1.0)	1.43	(0.95)	0.04	(0.86)
	Total	1.73	(0.82)	1.67	(0.82)	-0.06	(0.71)
Yoghurt	Full-fat	0.04	(0.12)	0.04	(0.14)	0.00	(0.16)
	Low-fat	0.27	(0.39)	0.29	(0.41)	0.02	(0.40)
	Total	0.30	(0.41)	0.33	(0.42)	0.02	(0.41)
Cheese	High-fat	0.34	(0.29)	0.30	(0.27)	-0.03	(0.28)
	Low-fat	0.13	(0.27)	0.12	(0.26)	0.00	(0.28)
	Total	0.47	(0.40)	0.43	(0.37)	-0.04	(0.38)
Cream		0.07	(0.17)	0.07	(0.18)	0.00	(0.18)
Butter		0.43	(0.93)	0.44	(0.89)	0.00	(0.88)
Ice-cream		0.21	(0.28)	0.20	(0.30)	-0.01	(0.30)
Fermented dairy products		0.77	(0.61)	0.76	(0.61)	-0.01	(0.57)
Total dairy products	High-fat	1.18	(1.41)	1.05	(1.29)	-0.14	(1.18)
	Low-fat	1.82	(1.18)	1.89	(1.14)	0.06	(1.02)
	Total	3.22	(1.41)	3.15	(1.39)	-0.07	(1.32)

Table 5.2 Descriptive characteristics of markers of metabolic risk and adiposity at baseline, first follow-up, second follow-up and the change between baseline and first follow-up in the EPIC-Norfolk study (n=15,612)

	Baseline		**1st follow-up**		**2nd follow-up**		**Change**	
Anthropometric markers	**Mean (SD)**							
Weight (kg)	73.0	(12.8)	74.1	(13.2)	74.1	(14.0)	1.3	(4.0)
BMI (kg/m^2)	26.1	(3.7)	26.7	(4.0)	26.8	(4.3)	0.6	(1.4)
Waist circumference (cm)	87.3	(12.1)	88.0	(12.5)	94.2	(12.2)	0.8	(5.5)
Waist-to-hip ratio	0.8	(0.1)	0.8	(0.1)	0.9	(0.1)	0.0	(0.0)
Body fat (%) †	-	-	32.7	(11.3)	31.6	(8.1)	-	-
Lipid markers								
Total / HDL-cholesterol	4.7	(1.6)	4.4	(1.6)	3.7	(1.1)	-0.2	(1.2)
Total cholesterol (mmol/l)	6.2	(1.2)	6.1	(1.2)	5.4	(1.1)	-0.1	(1.0)
HDL-cholesterol (mmol/l)	1.4	(0.4)	1.5	(0.5)	1.5	(0.4)	0.1	(0.3)
LDL-cholesterol (mmol/l)	4.0	(1.0)	3.8	(1.0)	3.2	(1.0)	-0.2	(0.9)
Triglycerides (mmol/l)	1.8	(1.1)	1.9	(1.1)	1.7	(0.9)	0.1	(0.9)
Other markers								
HbA1c (mmol/mol)	34.9	(8.5)	36.6	(7.3)	40.1	(6.7)	1.6	(6.1)
Systolic blood pressure (mm Hg)	134.3	(17.9)	135.1	(18.2)	136.1	(16.7)	0.5	(14.9)
Diastolic blood pressure (mmHg)	82.0	(11.0)	81.9	(11.2)	78.2	(9.3)	-0.1	(10.5)
Metabolic risk z-score	-0.02	(0.59)	0.0	(0.61)	0.0	(0.58)	0.01	(0.34)

† Body fat measurements were not available at baseline
Abbreviations: BMI: Body mass index; DBP: Diastolic blood pressure; HbA1c: Haemoglobin A1c; HDL-C: High density lipoprotein cholesterol; LDL-C: Low-density lipoprotein cholesterol; SBP: Systolic blood pressure; SD: Standard deviation

Table 5.3 Descriptive characteristics of socio-demographic, behavioural, clinical and non-dairy dietary factors at baseline, first follow-up and the change between baseline and first follow-up in the EPIC-Norfolk study (n=15,612) †

		Baseline	1st follow-up	Change
Socio-demographic factors				
Age (years)		58.6 (8.9)	62.1 (9.0)	3.2 (0.8)
Sex *	Women	56.2	56.2	
Educational level *	Medium	41.9	41.9	
	High	14.7	14.6	
Age completing education (years) §		8.7 (12.2)	8.7 (12.2)	0.0 (0.0)
Socio-economic status *	Medium	16.7	16.7	
	High	46.7	46.7	
Marital status *§	Married	82.6	82.6	
	Widowed / Separated	13.4	13.5	
Lifestyle factors				
Smoking status *	Former smoker	41.4	42.7	
	Current smoker	9.4	8.1	
Physical activity *‖	Moderately inactive	29.6	15.9	
	Moderately active	24.3	20.3	
	Active	19.3	15.3	
Energy intake (kcal/d) ‖		2,018 (552)	1,929 (529)	-88.6 (478.5)
Medications / Supplements				
Lipid-lowering medication *	Yes	1.5	4.7	
Anti-hypertensive medication *	Yes	16.4	21.8	
Hormonal therapy *	Yes	12.3	12.5	
Dietary supplements *‖	Yes	49.6	55.5	
Non-dairy food dietary factors (g/d) ‖				
Fruits		250.2 (180.4)	263.6 (187.7)	12.2 (170.9)
Vegetables		241.8 (1240.)	240.8 (123.3)	-2.1 (110.8)
Potatoes		115.2 (60.5)	111.7 (59.5)	-3.7 (69.3)
Legumes		60.0 (37.6)	56.0 (35.8)	-3.9 (38.3)
Processed cereals		82.1 (54.1)	78.9 (49.6)	-2.9 (53.0)
Whole-grain cereals		78.9 (78.0)	69.1 (73.1)	-10.3 (75.1)
Poultry and eggs		37.9 (23.9)	37.5 (24.7)	-0.3 (25.9)
Red meat		62.0 (40.5)	57.7 (39.5)	-4.1 (42.4)
Processed meat		27.9 (22.9)	27.5 (22.3)	-0.4 (21.8)
Fish		37.7 (25.8)	37.4 (25.5)	-0.2 (25.4)
Sauces		19.5 (17.9)	19.3 (18.6)	-0.2 (20.4)
Margarine		16.6 (16.3)	14.4 (15.1)	-2.2 (16.6)
Nuts		2.5 (7.4)	2.5 (7.7)	(9.2)
Sweet snacks		116.6 (84.6)	108.3 (82.3)	-8.6 (72.2)
Sugar-sweetened beverages		33.1 (72.1)	33.7 (75.6)	0.9 (83.2)
Artificially sweetened beverages		36.9 (104.9)	36.0 (104.3)	-0.3 (101.1)
Fruit juice		50.9 (69.1)	55.1 (72.0)	4.0 (73.7)
Regular coffee		329.5 (320.6)	298.6 (301.2)	-27.9 (252.8)
Decaffeinated coffee		87.6 (207.1)	76.8 (191.7)	-10.2 (187.7)
Tea		632.0 (365.2)	617.1 (360.5)	-18.2 (250.6)
Alcoholic beverages		128.8 (232.2)	125.4 (221.9)	-1.5 (154.7)

†Continuous variables are presented as mean(SD) and categorical variables are presented as column percentages
‡ Total percentage of missing values: 13% at baseline, 59% (49% due to missing values of the physical activity variable) at follow-up and 60% at both when accounting for non-ovelapping missing values for all the variables
*Reference categories: sex:men; educational level: low; socio-economic status: low; physical activity: inactive; lipid-lowering medication: no; anti-hypertensive medication: no; hormonal therapy: no; dietary supplements: no
§ Missing values <5% at baseline and follow-up
‖ Missing values <5% at baseline, but 20-50% at follow-up

Table 5.4 Descriptive statistics of socio-demographic, behavioural, clinical and dietary factors by milk, yoghurt and cheese consumption at baseline: the EPIC-Norfolk study †

		Milk				**Yoghurt**				**Cheese**			
Mean (SD) (servings/d)		1.7 (0.8)				0.3 (0.4)				0.5 (0.4)			
Range (servings/d) ‡		0.0	2.9 - 3.7			0.0	0.8 - 6.1			0.0	0.7 - 5.7		
Participants (N)*		530		2,617		5,229		2,631		796		4,068	
Socio-demographic factors													
Age (years)		57.0	(8.8)	58.4	(8.7)	59.9	(9.1)	58.1	(8.3)	58.6	(8.8)	58.4	(9.1)
Sex §	Women	65.5		46.7		42.7		70.0		61.4		60.5	
Educational level §	Medium	40.0		45.3		41.5		41.7		40.2		43.3	
	High	16.6		13.3		9.9		16.2		10.2		19.0	
Socio-economic status §	Medium	14.9		16.2		15.7		17.0		17.1		16.7	
	High	49.1		45.8		40.7		50.1		42.1		51.2	
Marital status ‖	Married	83.0		81.1		83.1		80.7		78.3		80.4	
	Widowed / Separated	12.5		14.0		12.7		15.4		17.3		15.2	
Lifestyle factors													
Smoking status	Former smoker	41.9		40.1		45.0		38.4		37.4		40.7	
	Current smoker	10.6		12.1		13.4		5.4		12.3		8.1	
Physical activity § #	Moderately inactive	30.6		26.9		27.1		30.7		28.8		30.4	
	Moderately active	21.1		25.1		22.8		26.2		25.4		24.7	
	Active	22.6		21.7		18.8		20.7		16.0		20.9	
Energy intake (kcal/d) #		1,760	(543)	2,223	(554)	2,022	(573)	2,021	(514)	1,825	(566)	2,149	(558)
Medications / Supplements													
Lipid-lowering medication §	Yes	1.9		1.5		1.3		1.9		2.1		1.0	
Anti-hypertensive medication §	Yes	15.1		15.8		17.5		16.8		18.3		15.7	
Hormonal therapy §	Yes	16.2		10.4		8.5		15.8		13.9		13.2	
Dietary supplements § #	Yes	50.8		48.5		41.8		58.1		50.2		51.3	
Non-dairy dietary factors (g/d) #													
Fruits		277.5	(227.9)	246.2	(182.7)	208.7	(162.9)	324.4	(206.5)	242.7	(189.6)	280.7	(197.2)
Vegetables		255.9	(140.3)	236.4	(123.9)	217.0	(117.1)	275.3	(129.7)	224.7	(140.4)	269.2	(138.3)
Potatoes		107.0	(62.0)	119.9	(59.4)	124.7	(68.0)	107.5	(52.1)	112.9	(64.9)	113.8	(62.8)
Legumes		59.5	(53.1)	62.3	(39.4)	59.8	(38.9)	62.1	(38.3)	58.2	(37.3)	62.5	(39.9)
Processed cereals		71.3	(58.0)	86.9	(55.0)	82.7	(55.8)	79.2	(52.2)	73.2	(55.3)	84.9	(55.5)

Table 5.4 (continued)

	Milk				Yoghurt				Cheese			
Mean (SD) (servings/d)	1.7 (0.8)				0.3 (0.4)				0.5 (0.4)			
Range (servings/d) ‡	0.0		2.9 - 3.7		0.0		0.8 - 6.1		0.0		0.7 - 5.7	
Participants (N)*	530		2,617		5,229		2,631		796		4,068	
Whole-grain cereals	74.8	(75.8)	82.4	(81.0)	66.6	(78.0)	94.7	(79.4)	78.6	(83.0)	90.3	(78.8)
Poultry and eggs	37.2	(26.9)	38.6	(23.3)	36.7	(24.9)	39.1	(25.3)	38.9	(24.7)	37.7	(23.8)
Red meat	57.0	(42.2)	64.9	(42.4)	63.3	(41.9)	56.2	(39.9)	49.2	(38.7)	62.2	(41.8)
Processed meat	24.8	(21.4)	29.3	(24.2)	31.7	(25.6)	23.4	(20.8)	25.4	(25.6)	26.5	(24.2)
Fish	38.6	(31.4)	37.9	(27.0)	34.9	(24.8)	40.4	(27.9)	39.0	(29.4)	38.5	(27.9)
Sauces	17.9	(17.8)	20.2	(19.1)	18.6	(19.6)	20.0	(17.3)	15.3	(13.8)	21.0	(18.3)
Margarine	14.1	(15.2)	17.8	(16.6)	16.7	(16.8)	16.1	(15.6)	15.0	(15.8)	18.0	(16.9)
Nuts	3.2	(9.1)	2.5	(7.3)	2.4	(7.9)	2.6	(6.6)	2.6	(9.8)	3.1	(8.5)
Sweet snacks	102.4	(88.9)	128.5	(90.9)	128.7	(93.2)	98.8	(73.9)	113.2	(94.6)	116.7	(84.6)
Sugar-sweetened beverages	33.4	(70.3)	31.7	(62.4)	36.3	(78.4)	29.5	(68.0)	33.0	(73.7)	32.4	(68.7)
Artificially sweetened beverages	60.4	(156.0)	25.7	(78.2)	28.1	(92.5)	54.2	(126.2)	37.9	(119.2)	41.0	(110.6)
Fruit juice	50.1	(76.7)	47.5	(69.0)	40.4	(65.9)	62.6	(73.7)	45.9	(71.4)	57.4	(72.9)
Regular coffee	408.5	(383.2)	329.0	(333.7)	317.6	(329.9)	314.6	(316.9)	246.6	(306.1)	362.6	(328.6)
Decaffeinated coffee	118.5	(276.4)	87.7	(210.4)	70.3	(195.2)	114.5	(232.2)	76.7	(197.5)	98.4	(218.5)
Tea	388.1	(410.9)	720.4	(369.5)	661.5	(374.3)	601.3	(365.9)	643.0	(387.0)	611.6	(370.5)
Alcoholic beverages	143.2	(247.4)	111.0	(222.0)	150.7	(279.5)	96.3	(161.8)	88.2	(187.2)	134.3	(231.3)

†Continuous variables are presented as mean(SD) and categorical variables are presented as column percentages

‡ Five categories; Milk: Non-consumers, 0.7 servings/day, 1.5 servings/day, 2.2 servings/day, and 2.9-3.7 servings/day as derived from the food frequency questionnaire Yoghurt: Non-consumers and quartiles within consumers; Cheese: Non-consumers and quartiles within consumers

* Total percentage of missing values: 13% at baseline, 59% (49% due to missing values of the physical activity variable) at follow-up and 60% at both when accounting for non-overlapping missing values for all the variables

§ Reference categories: sex:men; educational level: low; socio-economic status: low; physical activity: inactive; lipid-lowering medication: no; anti-hypertensive medication: no; hormonal therapy: no; dietary supplements: no

|| Missing values <5% at baseline and follow-up

Missing values <5% at baseline, but 20-50% at follow-up

5.4.2 Dairy products and anthropometric markers

Low-fat fermented dairy products i.e. yoghurt (total and low-fat) and low-fat cheese were inversely associated with body weight and BMI (p<0.05). While body weight was increased by 1.3±4 kg, those who increased their habitual daily yoghurt consumption by 1 serving had a lower increase in body weight by 0.23 kg (95% CI: -0.46, -0.01) in the most adjusted model (**Figure 5.1**). Changes in low-fat yoghurt and low-fat cheese were inversely associated with changes in body weight [-0.31 kg (-0.56, -0.06) and -0.64 kg (-0.97, -0.31) respectively] and BMI [-0.23 kg/m^2 (-0.36, -0.11) and -0.13 kg/m^2 (-0.22, -0.04) respectively]. Changes in high-fat dairy products, full-fat milk and high-fat cheese were positively associated with changes in body weight, BMI, or both. For example, an increase in high-fat dairy consumption by 1 serving/day was associated with 0.13 kg (0.05, 0.21) higher body weight (**Figure 5.1**). No associations were observed between changes in any types of dairy products and changes in waist circumference or waist-to-hip ratio (Figure 5.1). Associations adjusted for a smaller number of covariates did not give materially different results (**Table A.12**).

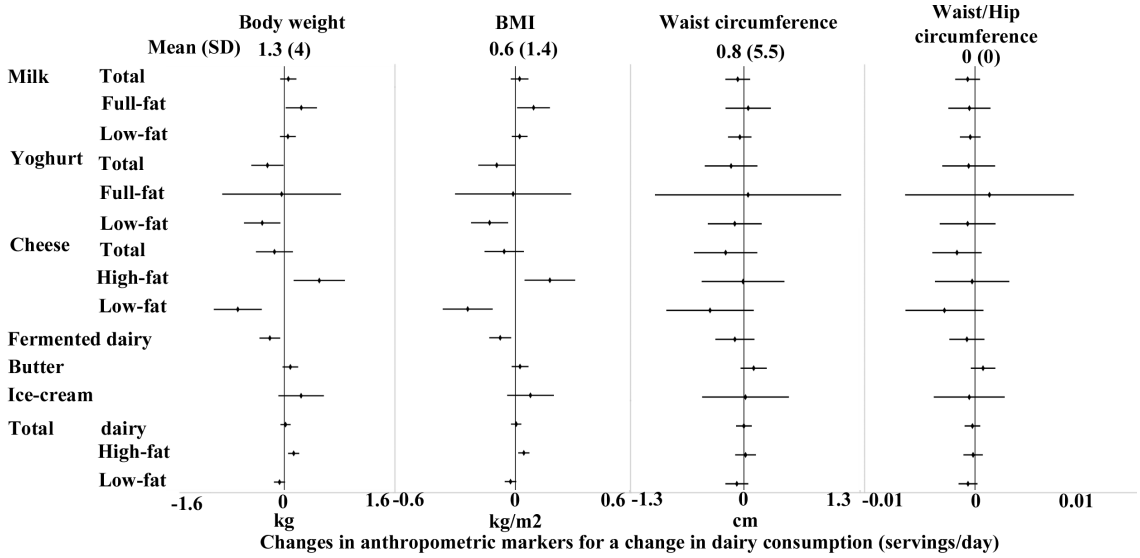


Fig. 5.1 Adjusted associations of changes in dairy consumption (servings/day) with changes in anthropometric markers. BMI: Body mass index

5.4.3 Dairy products and metabolic markers

Changes in butter were positively associated with changes in LDL-C [0.05 mmol/l (0.02, 0.07)] in the positive control analysis and also with changes in total cholesterol (**Figure 5.2**).

An increase in total dairy consumption was positively associated with an increase in total cholesterol [0.02 mmol/l (0.003, 0.04)], but not other lipids. Increasing levels of

high-fat dairy consumption were associated with increasing LDL-C levels [0.04 mmol/l (0.02, 0.06)]. Changes in total cheese consumption were not associated with changes in any lipid markers but increasing habitual high-fat cheese consumption by 1 serving/day was associated with 0.12 (0.04, 0.21) mmol/l higher total cholesterol, 0.04 mmol/l (0.01, 0.07) higher HDL-C and 0.09 mmol/l (0.02, 0.16) higher LDL-C. An increase of 1 serving/day of habitual low-fat dairy consumption was associated with a decrease of total and LDL-C by 0.03 mmol/l (-0.05, -0.01). Similar associations were observed for total and low-fat milk consumption and LDL-C. An increase of habitual total and low-fat yoghurt consumption was associated with 0.06 mmol/l (-0.12, -0.01) lower total cholesterol and 0.02 mmol/l (-0.04, -0.01) lower HDL-C respectively.

From associations between changes in dairy consumption and changes in non-lipid metabolic markers, only one significant association was observed between full-fat milk and HbA1c [0.52 mmol/mol (0.06, 0.97)] (**Figure 5.3**). Results from models of change for all the three different levels of adjustment used as described are presented in **Tables A.12-A.14**.

Overall, although the aforementioned associations were statistically significant at the nominal level, no association was significant when corrected for false-discovery rate (p>2x 10^{-5}).

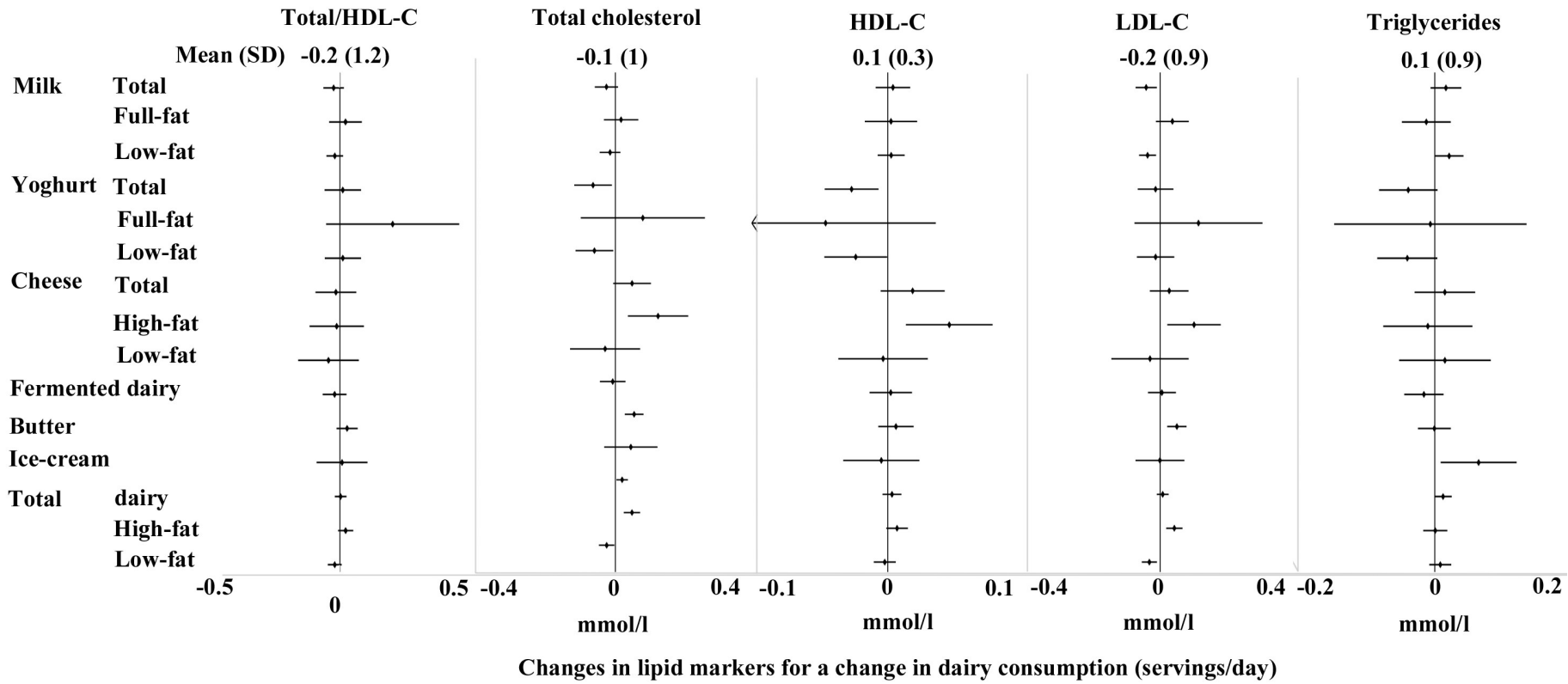


Fig. 5.2 Adjusted associations of changes in dairy consumption (servings/day) with changes in lipid markers. HDL-C: High density lipoprotein cholesterol, LDL-C: Low density lipoprotein cholesterol.

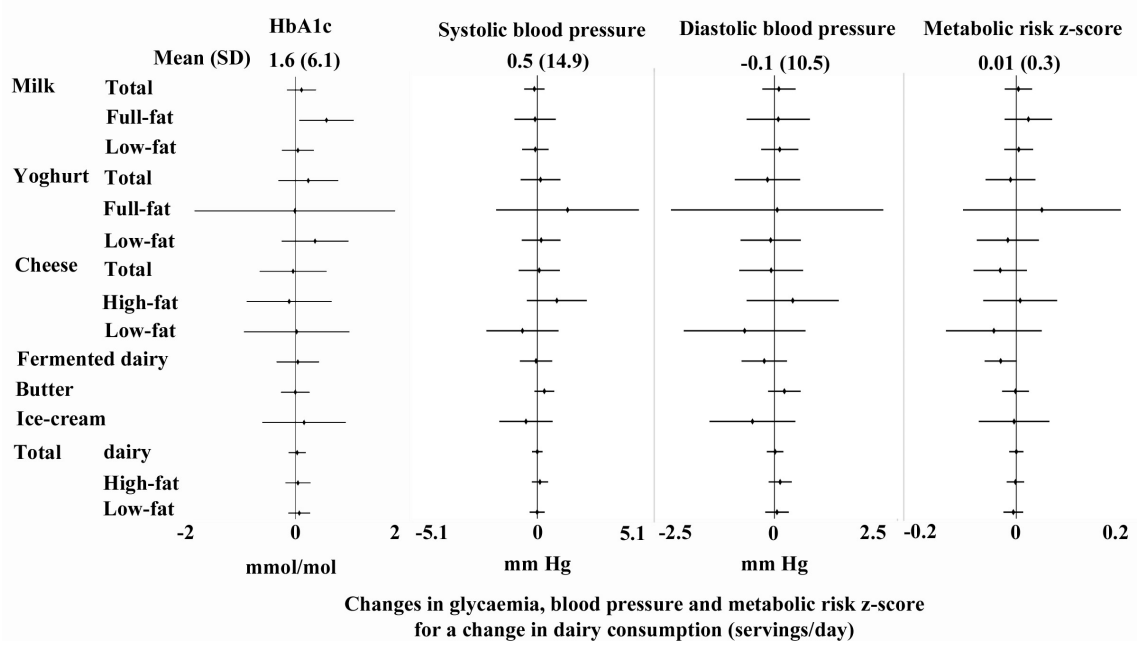


Fig. 5.3 Adjusted associations of changes in dairy consumption (servings/day) with changes in HbA1c, blood pressure and metabolic risk z-score. HbA1c: Haemoglobin A1c.

5.4.4 Longitudinal analyses

In secondary analyses, results from the most adjusted longitudinal models relating repeated measures of dairy products to repeated measures of cardio-metabolic markers at later time points are presented in **Tables 5.5- 5.7**.

Results from the positive control analysis showed a positive association between butter consumption and both LDL-C [0.32 mmol/l (0.15, 0.49)] as in the primary analysis and additionally triglycerides [1.34 mmol/l (0.71, 1.99)] (Table 5.6).

However, some of these associations were different in comparison to the primary results. Specifically, after correction for multiple testing, low-fat dairy consumption was on average positively associated with body weight and BMI [0.19 kg (0.11, 0.28)] and 0.16 kg/m^2 (0.1, 0.21) respectively] (Table 5.5). Similar associations were observed for milk (total and low-fat) and yoghurt (total and low-fat) (Table 5.5). Yoghurt consumption was also positively associated with waist circumference [0.54 cm (0.28, 0.79)] (Table 5.5). On the contrary, high-fat dairy consumption was inversely associated with the ratio of waist to hip circumference [-0.002 (-0.003, -0.001)] with similar associations for full-fat milk and high-fat cheese (Table 5.5).

Milk consumption was associated with higher ratio of total to HDL-C [1.18 (0.64, 1.72)] with similar associations observed for both full-fat and low-fat milk (Table 5.6). Low-fat dairy consumption was associated with lower HDL-C [-0.01 mmol/l (-0.02, -0.01)] (Table 5.6). Similar associations were found for total and low-fat milk.

Concerning the other markers of metabolic risk, no association was found between any dairy type and HbA1c after multiple test correction (Table 5.7). High-fat cheese consumption was inversely associated with SBP [-1.68 mmHg (-2.6, -0.77)], while full-fat milk was positively associated with DBP [0.5 mmHg (0.21, 0.79)] (Table 5.7). Finally, low-fat dairy consumption was positively associated with metabolic risk z-score [0.03 (0.02, 0.04)] with similar associations observed for total and low-fat milk (Table 5.7).

Table 5.5 Secondary analyses: Longitudinal associations of the repeated measures of total and types of dairy products at baseline (1993-1997) and first follow-up (1998-2000) with the repeated measures of the markers of body weight and composition at the first and the second (2004-2011) follow-up in the EPIC-Norfolk study †

	Weight (kg)			**BMI (kg/m^2)**			**Waist (cm)**			**Waist / Hip circumference**			**Body fat (%)**		
Mean (SD) at first follow-up	73.0 (12.8)			26.7 (4.0)			88.0 (12.5)			0.8 (0.1)			32.7 (11.3)		
Mean (SD) at second follow-up	74.0 (13.2)			26.8 (4.3)			94.2 (12.2)			0.9 (0.1)			31.6 (8.1)		
Participants (N)	14,145			14,107			14,199			14,218			14,019		
Dairy consumption (servings/d) ‡	b	95% CI		b	95% CI		b	95% CI		b	95% CI		b	95% CI	
Milk	0.16*	0.07	0.26	0.12*	0.06	0.19	0.01	-0.13	0.15	-0.001	-0.002	0.000	0.10	-0.02	0.21
Yoghurt	0.32*	0.15	0.49	0.27*	0.17	0.37	0.54*	0.28	0.79	0.002	0.000	0.004	0.22	0.03	0.41
Cheese	0.05	-0.12	0.23	0.00	-0.10	0.11	-0.16	-0.43	0.10	-0.002	-0.004	0.000	-0.22	-0.43	-0.02
Butter	-0.01	-0.10	0.08	0.00	-0.05	0.05	-0.12	-0.25	0.02	-0.001	-0.002	0.000	0.03	-0.07	0.13
Full-fat milk	0.09	-0.04	0.22	-0.01	-0.09	0.07	-0.28	-0.48	-0.08	-0.004*	-0.006	-0.002	-0.04	-0.18	0.11
Low-fat milk	0.16	0.07	0.25	0.13*	0.08	0.19	0.01	-0.13	0.15	-0.001	-0.002	0.000	0.12	0.02	0.23
Full-fat yoghurt	-0.01	-0.51	0.50	-0.17	-0.46	0.11	-0.06	-0.82	0.70	-0.003	-0.009	0.004	-0.19	-0.77	0.38
Low-fat yoghurt	0.35*	0.18	0.53	0.31*	0.21	0.41	0.59*	0.33	0.85	0.002	0.000	0.004	0.25	0.06	0.45
High-fat cheese	-0.22	-0.47	0.03	-0.18	-0.33	-0.03	-0.42	-0.79	-0.05	-0.005*	-0.008	-0.002	-0.77	-1.06	-0.48
Low-fat cheese	0.33	0.09	0.58	0.19	0.04	0.34	0.11	-0.26	0.48	0.001	-0.002	0.004	0.36	0.08	0.64
Ice-cream	0.11	-0.13	0.35	0.24	0.09	0.39	-0.22	-0.59	0.14	-0.002	-0.005	0.001	0.39	0.12	0.66
Fermented dairy products	0.20	0.07	0.32	0.15*	0.08	0.22	0.21	0.02	0.39	0.000	-0.002	0.002	0.01	-0.13	0.15
High-fat dairy products	0.01	-0.07	0.09	-0.03	-0.07	0.01	-0.17	-0.28	-0.05	-0.002*	-0.003	-0.001	-0.05	-0.13	0.04
Low-fat dairy products	0.19*	0.11	0.28	0.16*	0.10	0.21	0.09	-0.03	0.21	-0.001	-0.002	0.000	0.12	0.02	0.22
Total dairy products	0.10	0.03	0.17	0.08	0.03	0.12	-0.01	-0.11	0.09	-0.001	-0.002	0.000	0.07	-0.02	0.15

† Associations from the maximally adjusted linear mixed models are presented, which include: age (years), sex, educational level (low, medium, high), age at completion of full-time education (years), marital status (single, married, widowed or separated), socio-economic status based on occupation (low: technical/semi-routine and routine occupations; medium: lower managerial / intermediate occupations; high: professional / higher managerial occupations), individual follow-up time (years), physical activity level (inactive, moderately inactive, moderately active, active), smoking status (never, former and current smoker), lipid-lowering medication (Yes, No), anti-hypertensive medication (Yes, No), hormone-replacement therapy (Yes, No, Men), total energy intake (kcal/day), intakes (g/d) of fruit, vegetables, potatoes, legumes, nuts, processed cereals, whole-grain cereals, poultry and eggs, red meat, processed meat, fish, sauces, margarine, sweet snacks, sugar-sweetened beverages, artificially sweetened beverages, fruit juice, coffee, tea and alcoholic beverages, dietary supplement use (Yes, No) and BMI (kg/m^2; in associations other than those for weight and BMI). When repeated measures of the covariates were available at baseline and first follow-up, they were also used.
‡Servings as defined by Food Standards Agency 2002[197]: Milk- 1 average glass (200g); Yoghurt- 125g carton; Cheese- medium serving (40g); Single cream- 1 tablespoon (15g); Double cream- 1 tablespoon (30g); Butter- 1 teaspoon (10g); Ice-cream- 1 average scoop/tub (60g)
*$p < 9\text{x}10^{-4}$, which is the cut-off point as derived after False discovery rate correction
Abbreviations: BMI: Body mass index

Table 5.6 Secondary analyses: Longitudinal associations of the repeated measures of total and types of dairy products at baseline (1993-1997) and first follow-up (1998-2000) with the repeated measures of lipid markers at the first and the second (2004-2011) follow-up in the EPIC-Norfolk study †

	Total / HDL-C			**Total cholesterol (mmol/l)**			**HDL-C (mmol/l)**			**LDL-C (mmol/l)**			**Triglycerides (mmol/l)**		
Mean (SD) at first follow-up	4.4 (1.6)			6.1 (1.2)			1.5 (0.5)			3.8 (1)			1.9 (1.1)		
Mean (SD) at second follow-up	3.7 (1.1)			5.4 (1.1)			1.5 (0.4)			3.2 (1)			1.7 (0.9)		
Participants (N)	13,307			13,557			13,260			13,283			13,530		
Dairy consumption (servings/d) ‡	b	95% CI		b	95% CI		b	95% CI		b	95% CI		b	95% CI	
Milk	1.18*	0.64	1.72	-0.01	-0.03	0.02	-0.02*	-0.03	-0.01	0	-0.02	0.02	1.34*	0.71	1.99
Yoghurt	-0.39	-1.35	0.58	0	-0.04	0.04	0.01	-0.01	0.02	0	-0.04	0.04	-0.26	-1.38	0.86
Cheese	-0.3	-1.33	0.74	0	-0.05	0.05	0.01	0	0.03	-0.01	-0.05	0.04	-1.35	-2.55	-0.13
Butter	0.22	-0.28	0.72	0.04*	0.02	0.07	0.01	0	0.01	0.04*	0.01	0.06	-0.06	-0.65	0.54
Full-fat milk	1.59*	0.84	2.34	0.04	0.01	0.08	-0.01	-0.03	0	0.03	0	0.06	1.92*	1.02	2.83
Low-fat milk	1.14*	0.61	1.67	-0.01	-0.03	0.02	-0.02*	-0.03	-0.01	0	-0.02	0.02	1.28*	0.63	1.93
Full-fat yoghurt	-1.63	-4.43	1.24	0.12	-0.01	0.26	0.06	0.02	0.1	0.07	-0.06	0.2	-0.53	-3.77	2.82
Low-fat yoghurt	-0.23	-1.22	0.77	-0.01	-0.06	0.04	0	-0.01	0.02	0	-0.04	0.04	-0.24	-1.39	0.92
High-fat cheese	-1.86	-3.23	-0.47	-0.03	-0.09	0.04	0.03	0.01	0.05	-0.05	-0.11	0.01	-2.15	-3.73	-0.53
Low-fat cheese	1.33	-0.11	2.78	0.03	-0.04	0.09	0	-0.03	0.02	0.03	-0.02	0.09	-0.48	-2.15	1.23
Ice-cream	2.92*	1.52	4.34	0.06	0	0.13	-0.02	-0.04	0	0.08	0.02	0.14	1.21	-0.38	2.83
Fermented dairy products	-0.34	-1.03	0.34	0	-0.03	0.03	0.01	0	0.02	0	-0.03	0.03	-0.78	-1.59	0.03
High-fat dairy products	0.31	-0.11	0.73	0.04	0.01	0.06	0	0	0.01	0.03	0.01	0.04	0.19	-0.3	0.69
Low-fat dairy products	0.79	0.32	1.27	0	-0.02	0.02	-0.01*	-0.02	-0.01	0	-0.02	0.02	0.76	0.2	1.32
Total dairy products	0.56	0.16	0.96	0.02	0	0.04	0	-0.01	0	0.02	0	0.03	0.34	-0.12	0.81

† Associations from the maximally adjusted linear mixed models are presented, which include: age (years), sex, educational level (low, medium, high), age at completion of full-time education (years), marital status (single, married, widowed or separated), socio-economic status based on occupation (low: technical/semi-routine and routine occupations; medium: lower managerial / intermediate occupations; high: professional / higher managerial occupations), individual follow-up time (years), physical activity level (inactive, moderately inactive, moderately active, active), smoking status (never, former and current smoker), lipid-lowering medication (Yes, No), anti-hypertensive medication (Yes, No), hormone-replacement therapy (Yes, No, Men), total energy intake (kcal/day), intakes (g/d) of fruit, vegetables, potatoes, legumes, nuts, processed cereals, whole-grain cereals, poultry and eggs, red meat, processed meat, fish, sauces, margarine, sweet snacks, sugar-sweetened beverages, artificially sweetened beverages, fruit juice, coffee, tea and alcoholic beverages, dietary supplement use (Yes, No) and BMI (kg/m^2). When repeated measures of the covariates were available at baseline and first follow-up, they were also used.

‡ Servings as defined by Food Standards Agency 2002[197]: Milk- 1 average glass (200g); Yoghurt- 125g carton; Cheese- medium serving (40g); Single cream- 1 tablespoon (15g); Double cream- 1 tablespoon (30g); Butter- 1 teaspoon (10g); Ice-cream- 1 average scoop/tub (60g)

*$p < 9\text{x}10^{-4}$, which is the cut-off point as derived after False discovery rate correction

Abbreviations: HDL-C: High density lipoprotein cholesterol; LDL-C: Low density lipoprotein cholesterol

Table 5.7 Secondary analyses: Longitudinal associations of the repeated measures of total and types of dairy products at baseline (1993-1997) and first follow-up (1998-2000) with the repeated measures of HbA1c, blood pressure and a z-score for metabolic risk at the first and the second (2004-2011) follow-up in the EPIC-Norfolk study †

	HbA1c (mmol/mol)			**Systolic blood pressure (mmHg)**			**Diastolic blood pressure (mmHg)**			**Metabolic risk z-score**		
Mean (SD) at first follow-up	36.6 (7.3)			135 (18)			82 (11)			0.0 (0.6)		
Mean (SD) at second follow-up	40.1 (6.7)			136 (17)			78 (9)			0.0 (0.58)		
Participants (N)	8,886			14,172			14,175			8,816		
Dairy consumption (servings/d) ‡	b	95% CI		b	95% CI		b	95% CI		b	95% CI	
Milk	-0.05	-0.19	0.09	0.22	-0.11	0.55	0.16	-0.04	0.37	0.03*	0.02	0.05
Yoghurt	0.00	-0.24	0.24	0.33	-0.31	0.96	0.10	-0.29	0.49	0.02	0.00	0.05
Cheese	-0.31	-0.59	-0.03	-0.82	-1.46	-0.19	-0.19	-0.60	0.21	-0.03	-0.06	-0.01
Butter	0.00	-0.13	0.13	-0.10	-0.41	0.21	0.05	-0.15	0.24	0.00	-0.01	0.02
Full-fat milk	-0.21	-0.41	-0.01	0.31	-0.16	0.78	0.50*	0.21	0.79	0.01	-0.01	0.03
Low-fat milk	-0.04	-0.18	0.10	0.25	-0.08	0.59	0.16	-0.05	0.37	0.04*	0.02	0.05
Full-fat yoghurt	-0.03	-0.77	0.71	-0.19	-1.99	1.60	0.15	-1.01	1.32	-0.06	-0.13	0.01
Low-fat yoghurt	0.00	-0.25	0.25	0.39	-0.22	1.00	0.09	-0.30	0.48	0.03	0.00	0.05
High-fat cheese	-0.63	-1.01	-0.25	-1.68*	-2.60	-0.77	-0.18	-0.73	0.38	-0.08*	-0.11	-0.04
Low-fat cheese	-0.02	-0.40	0.36	0.11	-0.76	0.99	-0.15	-0.71	0.41	0.01	-0.03	0.05
Ice-cream	-0.14	-0.51	0.23	-0.22	-1.09	0.66	-0.06	-0.60	0.49	0.05	0.01	0.08
Fermented dairy products	-0.12	-0.31	0.06	-0.20	-0.64	0.24	-0.03	-0.31	0.24	0.00	-0.02	0.02
High-fat dairy products	-0.09	-0.2	0.03	-0.09	-0.35	0.18	0.16	-0.01	0.32	0.00	-0.02	0.01
Low-fat dairy products	-0.04	-0.16	0.07	0.19	-0.10	0.49	0.13	-0.05	0.30	0.03*	0.02	0.04
Total dairy products	-0.05	-0.15	0.05	-0.04	-0.28	0.20	0.05	-0.10	0.20	0.01	0.00	0.02

† Associations from the maximally adjusted linear mixed models are presented, which include: age (years), sex, educational level (low, medium, high), age at completion of full-time education (years), marital status (single, married, widowed or separated), socio-economic status based on occupation (low: technical/semi-routine and routine occupations; medium: lower managerial / intermediate occupations; high: professional / higher managerial occupations), individual follow-up time (years), physical activity level (inactive, moderately inactive, moderately active, active), smoking status (never, former and current smoker), lipid-lowering medication (Yes, No), anti-hypertensive medication (Yes, No), hormone-replacement therapy (Yes, No, Men), total energy intake (kcal/day), intakes (g/d) of fruit, vegetables, potatoes, legumes, nuts, processed cereals, whole-grain cereals, poultry and eggs, red meat, processed meat, fish, sauces, margarine, sweet snacks, sugar-sweetened beverages, artificially sweetened beverages, fruit juice, coffee, tea and alcoholic beverages, dietary supplement use (Yes, No) and BMI (kg/m^2; in associations other than those for metabolic risk z-score). When repeated measures of the covariates were available at baseline and first follow-up, they were also used.

Table 5.7 (continued)

‡ Servings as defined by Food Standards Agency 2002citeFSA20021: Milk- 1 average glass (200g); Yoghurt- 125g carton; Cheese- medium serving (40g); Single cream- 1 tablespoon (15g); Double cream- 1 tablespoon (30g); Butter- 1 teaspoon (10g); Ice-cream- 1 average scoop/tub (60g)
*$p < 9 \times 10^{-4}$, which is the cut-off point as derived after False discovery rate correction
Abbreviations: BMI: Body mass index; HbA1c: Haemoglobin A1c

5.4.5 Additional analyses

We conducted stratified analyses by age, sex and BMI, when the interactions were significant (p<0.05) (**Tables 5.8-5.10**). An increase in high-fat cheese consumption was associated with a higher increase in body weight and BMI only among participants within the 50-60 years age group [0.87 kg (0.33, 1.41), *p-int*=0.037 and 0.28 kg/m^2 (0.08, 0.48), *p-int*=0.027 respectively] (Table 5.8). Stratified results by sex showed an inverse association between an increase in low-fat milk and an increase in the ratio of total to HDL-C only among men [-0.06 (-0.1, -0.02)] (*p-int*=0.02), whereas an increase of cheese consumption was inversely associated with an increase in waist circumference only among women [-0.48 (-0.88, -0.08), *p-int*=0.009] (Table 5.9). Despite interactions with BMI being significant (*p-int*<0.05), associations within strata of BMI were null (Table 5.10).

Although the magnitude or significance of the associations from the complete-case analysis are different in certain cases, the directions of associations did not change (Tables A.15-A.17). From the main dairy types, the coefficient which substantially changed is the one for the association between the change in yoghurt consumption and the change in body weight, which became stronger [-0.47 kg (-0.77, -0.17); Table A.15].

Results from the specified sensitivity analyses in regards to quality of imputation, prevalent diseases, and possible attrition bias were not substantially different compared to results from the primary analysis (results not shown).

Table 5.8 Stratified results by age for significant interactions in associations of the change in total and types of dairy products with the change in cardio-metabolic markers from baseline to the first follow-up after a mean of 3.7 years in the EPIC-Norfolk study †

	Strata	**< 50 years**			**50-60 years**			**≥60 years**			
	N range ‡	2,795 - 2,830			4,620 - 4,665			6,493 - 6,506			
Dairy consumption (servings/day) §	**Cardio-metabolic marker**	b	95% CI		b	95% CI		b	95% CI		*Pint*
High-fat cheese	Weight (kg)	0.47	-0.38	1.31	0.87	0.33	1.41	0.09	-0.39	0.56	0.037
High-fat cheese	BMI (kg/m^2)	0.18	-0.11	0.47	0.28	0.08	0.48	0.04	-0.13	0.21	0.027
Low-fat dairy products	Waist (cm)	-0.16	-0.39	0.07	-0.12	-0.29	0.05	0.05	-0.10	0.21	0.019

† Associations from the maximally adjusted linear regression models are presented, which include: age (years), sex, educational level (low, medium, high), age at completion of full-time education (years), marital status (single, married, widowed or separated), socio-economic status based on occupation (low: technical/semi-routine and routine occupations; medium: lower managerial / intermediate occupations; high: professional / higher managerial occupations), individual follow-up time (years), physical activity level (inactive, moderately inactive, moderately active, active), smoking status (never, former and current smoker), lipid-lowering medication (Yes, No), anti-hypertensive medication (Yes, No), hormone-replacement therapy (Yes, No, Men), total energy intake (kcal/day), intakes (g/d) of fruit, vegetables, potatoes, legumes, nuts, processed cereals, whole-grain cereals, poultry and eggs, red meat, processed meat, fish, sauces, margarine, sweet snacks, sugar-sweetened beverages, artificially sweetened beverages, fruit juice, coffee, tea and alcoholic beverages, dietary supplement use (Yes, No) and BMI (kg/m^2; in associations other than those for metabolic risk z-score). When repeated measures of the covariates were available at baseline and first follow-up, their change was also included.

‡ A sample size range is given instead of a single sample size due to the exclusion of participants outside the range of mean±3SD of the change in dairy consumption

§ Servings as defined by Food Standards Agency 2002[197]: Milk- 1 average glass (200g); Yoghurt- 125g carton; Cheese- medium serving (40g); Single cream- 1 tablespoon (15g); Double cream- 1 tablespoon (30g); Butter- 1 teaspoon (10g); Ice-cream- 1 average scoop/tub (60g)

Abbreviations: BMI: Body mass index

Table 5.9 Stratified results by sex for significant interactions in associations of the change in total and types of dairy products with the change in cardio-metabolic markers from baseline to the first follow-up after a mean of 3.7 years in the EPIC-Norfolk study †

	Strata	**Men**			**Women**			
	N range ‡	5404 - 6087			7263 - 7953			
Dairy consumption (servings/day) §	**Cardio-metabolic marker**	b	95% CI		b	95% CI		Pint
Low-fat milk	Total / HDL-C	-0.06	-0.1	-0.02	0.02	-0.02	0.05	0.02
Low-fat yoghurt	Waist / Hip circumference	0	-0.01	0	0	0	0	0.027
High-fat cheese	Waist (cm)	0.43	-0.23	1.08	-0.31	-0.85	0.22	0.021
Total cheese	Waist (cm)	0.21	-0.32	0.75	-0.48	-0.88	-0.08	0.009

† Associations from the maximally adjusted linear regression models are presented, which include: age (years), sex, educational level (low, medium, high), age at completion of full-time education (years), marital status (single, married, widowed or separated), socio-economic status based on occupation (low: technical/semi-routine and routine occupations medium: lower managerial / intermediate occupations; high: professional / higher managerial occupations), individual follow-up time (years), physical activity level (inactive, moderately inactive, moderately active, active), smoking status (never, former and current smoker), lipid-lowering medication (Yes, No), anti-hypertensive medication (Yes, No), hormone-replacement therapy (Yes, No, Men), total energy intake (kcal/day), intakes (g/d) of fruit, vegetables, potatoes, legumes, nuts, processed cereals, whole-grain cereals, poultry and eggs, red meat, processed meat, fish, sauces, margarine, sweet snacks, sugar-sweetened beverages, artificially sweetened beverages, fruit juice, coffee, tea and alcoholic beverages, dietary supplement use (Yes, No) and BMI (kg/m^2; in associations other than those for metabolic risk z-score). When repeated measures of the covariates were available at baseline and first follow-up, their change was also included.
‡ A sample size range is given instead of a single sample size due to the exclusion of participants outside the range of mean$\pm$3SD of the change in dairy consumption
§ Servings as defined by Food Standards Agency 2002[197]: Milk- 1 average glass (200g); Yoghurt- 125g carton; Cheese- medium serving (40g); Single cream- 1 tablespoon (15g); Double cream- 1 tablespoon (30g); Butter- 1 teaspoon (10g); Ice-cream- 1 average scoop/tub (60g)

Table 5.10 Stratified results by BMI for significant interactions in associations of the change in total and types of dairy products with the change in cardio-metabolic markers from baseline to the first follow-up after a mean of 3.7 years in the EPIC-Norfolk study †

	Strata	**< 25 kg/m^2**			**≥25 kg/m^2**			
	N range ‡	5518 - 5899			7179 - 8064			
Dairy consumption (servings/day) §	**Cardio-metabolic marker**	b	95% CI		b	95% CI		Pint
Yoghurt	Total cholesterol (mmol/l)	0.06	-0.16	0.28	0.13	-0.06	0.31	0.014
Yoghurt	LDL-C (mmol/l)	0.00	-0.002	0.002	0.001	-0.001	0.003	0.011
Butter	Waist (cm)	-0.03	-0.09	0.02	0.003	-0.05	0.05	0.003
Butter	Waist / Hip circumference	0.00	-0.11	0.11	-0.01	-0.12	0.1	0.032
Fermented dairy products	Total cholesterol (mmol/l)	-0.07	-0.21	0.07	0.08	-0.05	0.22	0.028
High-fat dairy products	Waist (cm)	-0.01	-0.09	0.06	-0.01	-0.07	0.05	0.04
Total dairy products	Waist (cm)	-0.07	-0.15	0.01	-0.07	-0.13	0.003	0.005

† Associations from the maximally adjusted linear regression models are presented, which include: age (years), sex, educational level (low, medium, high), age at completion of full-time education (years), marital status (single, married, widowed or separated), socio-economic status based on occupation (low: technical/semi-routine and routine occupations medium: lower managerial / intermediate occupations; high: professional / higher managerial occupations), individual follow-up time (years), physical activity level (inactive, moderately inactive, moderately active, active), smoking status (never, former and current smoker), lipid-lowering medication (Yes, No), anti-hypertensive medication (Yes, No), hormone-replacement therapy (Yes, No, Men), total energy intake (kcal/day), intakes (g/d) of fruit, vegetables, potatoes, legumes, nuts, processed cereals, whole-grain cereals, poultry and eggs, red meat, processed meat, fish, sauces, margarine, sweet snacks, sugar-sweetened beverages, artificially sweetened beverages, fruit juice, coffee, tea and alcoholic beverages, dietary supplement use (Yes, No) and BMI (kg/m^2; in associations other than those for metabolic risk z-score). When repeated measures of the covariates were available at baseline and first follow-up, their change was also included.
‡ A sample size range is given instead of a single sample size due to the exclusion of participants outside the range of mean±3SD of the change in dairy consumption
§ Servings as defined by Food Standards Agency 2002[197]: Milk- 1 average glass (200g); Yoghurt- 125g carton; Cheese- medium serving (40g); Single cream- 1 tablespoon (15g); Double cream- 1 tablespoon (30g); Butter- 1 teaspoon (10g); Ice-cream- 1 average scoop/tub (60g)
Abbreviations: BMI: Body mass index; LDL-C: Low-density lipoprotein cholesterol

5.5 Discussion

5.5.1 Summary of results

In the pre-specified primary analysis on associations of changes in dairy consumption with parallel changes in cardio-metabolic markers, we observed an inverse association of changes in low-fat fermented dairy consumption, with changes in body weight and BMI over an average of 3.7 years. In contrast, there was a positive association of changes in high-fat dairy consumption with changes in body weight and BMI. Increasing low-fat dairy consumption was inversely associated with an increase in both total and LDL-C, while increasing total and high-fat dairy consumption (especially butter or high-fat cheese), was associated with a greater increase in total cholesterol, LDL-C and HDL-C.

5.5.2 Findings in context of previous evidence

Anthropometric markers

Our results on associations between changes in dairy consumption and changes in anthropometric markers only agree with some other prospective cohort studies. Results from these studies are mixed, most of them indicating either inverse or null associations and sparse for the different dairy types or groups they examine.

We did not find any association between the change in total or low-fat dairy consumption and changes in body weight or BMI, whereas we observed a positive association between changes in high-fat dairy consumption including full-fat milk and changes in weight and BMI. Of the previous studies on prospective associations between dairy consumption and body weight, six investigated associations of total dairy consumption[64–69, 81], five investigated associations of low-fat dairy consumption[66, 67, 69–71, 80] and four examined associations of high-fat dairy consumption[66, 67, 69, 80]. Results were conflicting for total and high-fat dairy products with four of them indicating an inverse[66–69] and five of them a null association[64–67, 70]. Null associations were reported in all the five studies examining associations of low-fat dairy products[66, 67, 69–71, 80]. We also observed inverse associations between changes in the consumption of low-fat fermented dairy products i.e. total and low-fat yoghurt and low-fat cheese and changes in weight and BMI. Our results agree with the results from three prospective studies on the association between yoghurt and weight change[65, 67, 70]. One study did not find any association between total or low-fat yoghurt and weight change[73].

We observed no associations between changes in any dairy type and changes in waist circumference and the ratio of waist to hip circumference. Our results are in agreement with the majority of the prospective studies, as null associations were reported in four out of five studies on total dairy consumption[64, 65, 67, 68], all three studies on low-fat dairy consumption[67, 71, 80] and all four studies on cheese[65, 67, 68, 71] in relation to the

change in waist circumference. In contrast, inverse associations with the change in waist circumference were reported from two studies on yoghurt[65, 67], one study on milk[65] and one study on butter[80].

Results from RCTs are mainly available for total dairy products as most of the interventions included a mixture of the three main dairy types i.e. milk, yoghurt and cheese[58]. According to a recent meta-analysis of RCTs, an average total dairy consumption of 2.6±1 serving/day decreased body weight by 0.64 kg during an average period of 7.7±7.9 months in 16 trials which applied energy restriction[58]. A decreasing effect was also reported for waist circumference in trials with energy restriction[58]. The same meta-analysis, reported an increasing effect of total dairy products on body weight in trials without any energy restriction[58], potentially because in the case of energy restriction, the same restriction is applied to both the intervention and the control group, whereas in the case of no energy restriction, if the randomisation is not successful or the compliance differs between the intervention and the control group, any effect might be due to differences in energy intake.

Potential mechanisms for an inverse association of low-fat fermented dairy products, such as yoghurt and low-fat cheese with body weight, have been reported in relevant studies. There is evidence that yoghurt has beneficial effects on gut function, which has been linked to obesity and type 2 diabetes[198], owing to the fermentation process it has gone through[199]. In addition, the content of most nutrients in yoghurt is higher than that of milk, as yoghurt is more condensed and their bioavailability has been proposed to be enhanced due to a lower gastric pH caused by the higher acidity of yoghurt[32]. That means that potential effects from nutrients e.g. an enhanced regulation of satiety and appetite by dairy protein[178–180] might be more pronounced in yoghurt than in milk.

The underlying biological mechanism for the positive associations of the change in high-fat dairy and milk consumption with the change in body weight and BMI after adjustment for total energy intake is not clear. There is evidence of an increasing effect of total dairy consumption on body lean mass with a simultaneous decreasing effect on body fat mass, resulting in a decreasing total effect on body weight[58]. It is harder to disentangle this balance of effects when examining dairy types, as the evidence is more limited, so a potential explanation could be that for high-fat dairy products the increasing effect on lean mass prevails over the decreasing effect on body fat mass, leading to a slight body weight increase.

Metabolic markers and blood pressure

We did not observe any association between the change in dairy consumption and the change in the ratio of total to HDL-C. Past evidence is limited except for butter, which has been shown to increase the ratio when it replaces soft margarine[87] or olive oil[89]. Our results of a positive association between the change in high-fat dairy consumption including butter and high-fat cheese and the change in total and LDL-C and HDL-C

(for high-fat cheese only) overall agree with evidence from RCTs. In a meta-analysis of 20 clinical trials, LDL-C was reduced after replacement of butter with soft or hard margarine[87]. Butter also increases LDL-C when it substitutes food sources of mono-[89] or poly-unsaturated fatty acids[105].

Concerning associations of cheese consumption with lipids, a meta-analysis of five RCTs concluded that cheese consumption over a period of 2-8 weeks results in a lower increase in cholesterol (total, LDL, HDL) than butter of the same polyunsaturated fat to saturated fat content[83]. In our study we observed similar associations between butter and LDL-C (0.05 mmol/l decrease per 10 g of butter or 5.2 g of saturated fat) and high-fat cheese and LDL-C (0.09 mmol/l decrease per 40 g of cheese or 8.67 g of saturated fat). However, considering the nature of the FFQ, it is possible that there is measurement error in the estimates, which could result in slightly different estimates than the true.

We also observed inverse associations of the change in low-fat dairy consumption with the change in total (mainly due to yoghurt) and LDL-C (mainly due to milk). Evidence on associations of other dairy products with lipids is sparse. A meta-analysis of nine RCTs reported null effects of low- or high-fat dairy products on LDL-C or HDL-C[63]. It is not clear though whether butter was included in the high-fat dairy interventions. A positive association between high-fat dairy products and lipids is also supported from evidence on the effect of saturated fat on lipids. Saturated fat has been consistently reported to increase total, LDL-C and HDL-C, when it substitutes carbohydrates[200], mono- or poly-unsaturated fatty acids[201]. Inverse associations of low-fat dairy products with blood lipids could be attributed to calcium, which decreases lipogenesis and increases lipolysis, an effect which might be more pronounced in low-fat than high-fat dairy products[177].

We found null associations between the change in dairy consumption and the change in HbA1c, except for milk which was positively associated. Overall, our results agree with evidence from a meta-analysis of RCTs, where neither low- nor high-fat dairy products were associated with markers of glycaemia such as fasting blood glucose or homeostasis model assessment for insulin resistance (HOMA-IR)[63]. Evidence on other dairy products and glycaemic markers is insufficient to draw any conclusions. Dairy products have been characterised as insulinotropic in the short-term, an effect that they do not seem to hold in the long-term, which potentially explains the differential results according to the study design[185]. We did not identify any significant associations between the change in dairy consumption and the change in blood pressure in accordance with results from a meta-analysis of seven RCTs in normotensive people[63].

5.5.3 Longitudinal analyses

Although some of the longitudinal associations in our secondary analysis were significant after FDR correction, some of the results from associations with anthropometric markers

were less concordant with results from other prospective cohort studies and RCTs, while associations with lipids were more consistent.

For example, our positive control analysis of the association between butter and LDL-C led to a positive association, which is consistent with evidence from RCTs as described before (section 5.5.2). An example of disagreement with past evidence is the association between yoghurt consumption and body weight, which was positive in the longitudinal analysis even after adjustment for energy intake and other dietary variables, but inverse in the analysis of parallel change. As described above and according to results from other prospective cohort studies on the association between yoghurt consumption and body weight, either inverse[65, 67, 70] or null[73] associations were reported.

The explanation of the results from our longitudinal analyses is not clear, but I can speculate on possible reasons. Longitudinal associations might be biased if not adjusted for the baseline outcome e.g. baseline body weight to account for baseline differences in body weight. However, it is not advised to adjust for baseline outcome in observational studies, where it is often not possible to know whether the baseline exposure e.g. dairy consumption or other factors related to the exposure has influenced the baseline outcome[192], so that adjustment for that could lead to collider bias and spurious associations[194]. Consequently, we trust that our main analysis of parallel change is a valid approach to assess prospective associations, but we cannot quantify any potential bias introduced to our longitudinal analysis. In this analysis, we showed that such biases might be stronger for anthropometric outcomes, maybe due to a greater inherent behavioural component, which is an important point of consideration for prospective analyses in nutritional epidemiology.

5.5.4 Strengths and limitations

This study has several strengths. The prospective design of our study reduces the probability of reverse causation bias. The repeated measures of dairy consumption and cardio-metabolic markers at the same time points allowed us to perform an analysis of parallel change, which has been shown to give the most biologically plausible results compared with other analytical approaches in observational prospective cohort studies[187]. We used a comprehensive set of types and groups of dairy products in the associations examined and we investigated multiple potential pathways for the associations between dairy consumption and cardio-metabolic disease. We also adopted a rigorous statistical approach including the use of multiple imputation to handle missing data and the use of inverse probability weighting in our secondary analyses to examine the presence of healthy survivor's effect[202].

This study also has limitations. The observational nature of the study does not allow us to make any causal inferences, but we can generate hypotheses to be further tested in RCTs. Diet was assessed with an FFQ, which is a subjective, self-reported method, usually accompanied by a degree of measurement error, which we could not quantify in the present

study. Although the analysis of parallel change has been identified as a suitable analytical approach to investigate associations between diet and body weight[187], the change in dairy consumption observed in our population over the 3.7 years of follow-up might not have been as large and variable as needed to detect some associations. Finally, although we adjusted for several potential confounders, we cannot eliminate the possibility of residual confounding.

5.5.5 Conclusion

The current analysis of parallel changes in dairy consumption and markers of metabolic risk, showed differential associations with adiposity and lipidaemia for different dairy types, which extends previous understanding. The main result of an inverse association between an increase in low-fat fermented dairy products and an increase in body weight is a potential pathway for the previously described association with type 2 diabetes[160]. These findings contribute to greater understanding of the differential associations of dairy products with cardio-metabolic health and should be further confirmed in clinical settings and other populations.

We also showed that longitudinal analyses i.e. associations between exposure at certain time points and outcome at later time points might include bias, as shown in previous studies, but especially for anthropometric outcomes, which are affected more by behavioural components and are thus more prone to bias. Knowing that this analysis might entail bias, we pre-specified the analysis of parallel change as the primary analysis and we based our conclusions on this analysis.

Chapter 6

Biomarkers of dairy consumption: Part I. Development and validation of metabolite scores

Summary

Background and aims: Measurement error in self-reported dietary assessment is a well-recognised limitation of nutritional research, and there is interest in the use of objectively measured nutritional biomarkers independent of such errors. To improve understanding of the links between dairy products and cardio-metabolic disease, we aimed to develop and validate metabolite scores predicting consumption of different dairy types and to investigate their predictive value over blood fatty acids, which have been identified as dairy fat biomarkers.

Methods: We evaluated metabolomic profiles using the targeted Biocrates platform among 10,281 participants of the Fenland study for the discovery (n=6,035) and validation (n=4,246) of metabolite scores predicting each of total and types of dairy products separately, accounting for socio-demographic, lifestyle and clinical factors. In the internal validation set, we examined how well the metabolite scores with or without phospholipid fatty acids (C15:0, C17:0 and trans-16:1n-7, as candidate biomarkers of dairy consumption) could predict milk, yoghurt, cheese, butter, and total dairy consumption assessed from a food frequency questionnaire. Next, we evaluated data from an untargeted metabolomics platform (Metabolon) among 1,440 participants of the diabetes case-cohort set nested within the EPIC Norfolk study to externally validate the metabolite scores for prediction of dairy consumption.

Results: The area under the curve (AUC) statistic for consumption of the selected dairy products ranged from 0.68 to 0.81 in the internal validation (p<0.05 for prediction by metabolite scores). Addition of the odd-chain saturated fatty acids (OCSFAs) in the discovery analysis resulted in an AUC range of 0.72-0.84. In the external validation set,

AUCs of 0.60, 0.62 and 0.64 respectively were observed for milk, butter and total dairy (p<0.05 for milk only), increasing to 0.65, 0.66 and 0.77 respectively (p<0.05 for all) when OCSFAs were included in the models. For yoghurt and cheese, AUCs were 0.69 and 0.66 respectively in multivariable models, but each of the metabolite scores did not significantly predict each dairy type in the external validation before and after adding phospholipid fatty acids.

Conclusions: A set of metabolites could predict milk, butter and total dairy consumption with internal and external validity. External validity of yoghurt and cheese was not confirmed. These findings indicate that the use of metabolomics is a promising approach for the identification of novel biomarkers of dairy consumption and we recommend the replication of our approach in other populations and the use of more sets of metabolites and biological samples e.g. metabolites related to the gut microbiome, which might better reflect fermented products like yoghurt and cheese.

What is already known

- Measurement error of self-reported methods of dietary assessment is a common limitation in nutritional epidemiology.
- The use of nutritional biomarkers is independent of this measurement error.
- The odd-chain saturated fatty acids (OCSFAs; C15:0 and C17:0) and trans-16:1n-7 have been suggested as candidate biomarkers of dairy fat, but with limitations.
- To overcome limitations of candidate biomarkers of dairy fat, it is of interest to identify novel biomarkers, which can be potentially achieved with the use of metabolomics.

What this research adds

- In a discovery analysis, SM-OH C14:1 was one of the top metabolite signals predicting all dairy types, SM C16:1 predicted milk and total dairy products, and LPC a C17:0 predicted cheese, butter and total dairy products.
- We developed metabolite scores predicting total and each dairy type in the discovery and internal validation set.
- In the external validation set, metabolite scores significantly predicted milk, butter and total dairy products, but not yoghurt or cheese.
- The metabolite scores had predictive value in addition to the phospholipid fatty acids for all the dairy types, while fatty acids had predictive value in addition to the metabolite scores only for butter in the internal validation set.

Publication

Trichia E, Imamura F, Koulman A, Brage S, Griffin SJ, Langenberg C, Khaw KT, Wareham NJ, Forouhi N G. Development and validation of dairy prediction models using metabolomics in two UK cohorts and the associations of derived metabolite scores with type 2 diabetes risk (Manuscript under preparation)

6.1 Measurement error in dietary assessment

The measurement error, which accompanies the self-reported methods of dietary assessment is a common limitation in nutritional epidemiology[203]. Several approaches have been proposed to reduce and partly account for measurement error. First, adjustment of associations between diet and any outcome for energy intake estimated from the same dietary assessment method as the exposure -thus having correlated measurement errors- may reduce measurement error by partly cancelling out the correlated errors [152, 204]. When the method of dietary assessment is the food frequency questionnaire (FFQ), this only gives a crude estimate of total energy intake. However, it is still important to do the adjustment, primarily to partly account for the confounding related to it, but also to partly account for the correlated measurement errors[204].

In addition, dietary misreporting has been shown to be differential across methods of dietary assessment and groups of people. Although there is no gold standard of self-reported dietary assessment methods, the error derived from multiple-day food diaries and 24-hour recalls is on average smaller than that from FFQ[152, 205, 206]. If only one method of dietary assessment is available in a study, it is not possible to address the error which is specific to this method. However, taking into consideration several predictors of dietary misreporting might partly account for measurement error. The most well-characterised predictor of energy misreporting is body mass index (BMI) with under-reporters being overall overweight or obese and over-reporters usually being underweight[145–152]. The ability of high BMI to predict dietary under-reporting is not limited to the current status, but extends to a high-BMI history or generally BMI fluctuations across time[149, 151]. Other proposed predictors of energy under-reporting are older age[145, 147, 148, 151], women[147, 151, 206], high percentage of body fat[151, 206], lower education[147, 151] and some psychosocial factors including social desirability, dissatisfaction with body image, dieting and restrained eating[150, 151]. Energy over-reporting has been less studied, but some suggested younger age[151], men[147], current smoking[147] and low socio-economic status[148] as predictors additionally to underweight.

In some studies, energy under-reporters also reported lower intake of total fat[148, 149, 151], saturated fat[149], carbohydrates[151], sugar[149–151], added salt[148], fibre[149, 151], frequency of snacks[148] and consumption of fried foods[148], whereas they reported higher intakes of total protein[148–150], fibre[148], calcium[148], iron[148], vitamin C[148], and folate[148]. Considering that dairy products are important sources of saturated fat, protein and calcium, as we described in section 2.5.4, self-reported dairy consumption is also prone to measurement error. In addition, some dairy types, specifically milk and butter are included in many composite foods, which might be one additional level of complexity and source of error. As we reported in section 2.5.4, we identified large differences in the amounts of milk and butter consumption when we accounted for their content in composite foods compared to previously reported estimates.

Even if consistent efforts have been made to account for measurement error, no universal solution has been identified so far, as it is often impossible to know all the potential sources and correlates of error which the self-reported nature of dietary assessment methods entails. Consequently, there have been attempts to replace this subjective element with more objective methods of dietary assessment including the doubly labelled water technique for the estimation of total energy intake and the identification of biomarkers of dietary intake, which reflect nutrient intakes or food consumption.

6.2 Nutritional biomarkers

The assessment of nutritional biomarkers also entails challenges and measurement errors. However, these are of a different nature and thus uncorrelated with the measurement errors from the self-reported dietary assessment methods[204]. Although the use of nutritional biomarkers overcomes the errors related to the subjective nature of the self-reported dietary assessment, there are still some points to consider when designing or conducting a nutritional biomarker study[204]:

- Between-person variation due to genetic, hormonal, homeostatic, metabolic and gut microbiome-related differences, as well as the interplay between them.
- The range of intake of a nutrient or food in a population relative to the plateau that the corresponding biomarker might reach in the biological sample.
- Lifestyle factors e.g. smoking, physical activity and alcohol consumption.
- Pathological conditions and use of medications.
- Differential bioavailability of biomarker-related food components.
- The specificity of the biomarker to a nutrient or food. Specificity is more difficult to be achieved for foods with overlapping nutrient profiles.
- Whether a biomarker can reflect diet in an absolute way (recovery biomarkers) or a relative way (concentration biomarkers). Ideally, a dose-response quantification is desirable, which can be achieved with recovery biomarkers such as biomarkers detected in 24-hour urine or doubly labelled water. However, the majority of biomarkers are concentration biomarkers, which are still informative for ranking individuals with low or high habitual consumption.
- The reference period of a biomarker, which depends on the kinetics of the food component and the type of biological specimen the biomarker is stored in. When it is of interest to assess the habitual intake of people in an epidemiological study, biomarkers with a short half-life are useful only for foods which are consumed on a

regular basis, so that we do not have high inter-individual variability depending on the day of tissue collection[207].

- Differences in food production, storage, processing or preparation.
- Technical errors related to laboratory methods.

The uncorrelated errors between nutritional biomarkers and self-reported dietary assessment and the useful information we can obtain from both methods, imply that biomarkers cannot replace self-reported dietary assessment, but both are useful tools to complement each other and used in combination.

6.3 Odd-chain and trans-16:1n-7 fatty acids as biomarkers of dairy consumption

Results from associations of dairy consumption with cardio-metabolic disorders are mainly derived from studies using self-reported methods of dietary assessment, so it is of interest to identify potential biomarkers of dairy consumption. The first studies, which explored such biomarkers, used a hypothesis driven approach to investigate fatty acids as potential biomarkers of the ranking of dairy consumption (concentration biomarkers). This hypothesis, which included the odd-chain saturated fatty acids (OCSFAs) C15:0 (pentadecanoic acid) and C17:0 (heptadecanoic acid), was generated from the observation that these fatty acids are produced by microbial fermentation or microbial de novo lipogenesis in the ruminant duodenum and are transferred to the ruminant milk, which is then consumed by humans, but cannot be produced endogenously in the human body[208]. Before that observation, the only use of OCSFAs was as internal standards in the chemical analyses of fatty acids with gas or liquid chromatography[209, 210]. Instead the focus was on even-chain fatty acids, which constitute more than 99% of fatty acids in human blood plasma in contrast to the very small amounts of OCSFAs[208]. Likewise, OCSFAs constitute on average 1.6% of the total saturated fat in milk, whereas C14:0, C16:0 and C18:0 constitute 11.2%, 28.9% and 11.1% respectively[35]. However, C16:0 is more abundant in meat (e.g. 36% of the saturated fat in beef) and C18:0 is more abundant in cocoa products (approximately 30% of total saturated fat) and meat (approximately 20% of total saturated fat)[35], so their specificity as dairy biomarkers was questionable. In addition, coconut oil contains more C14:0 than milk (18% of total fatty acids in coconut oil)[35], but it has been studied as a potential biomarker of dairy fat, since it can be assumed that consumption levels of coconut oil are on average lower than that of dairy products in certain populations.

From a literature review and citations of relevant papers, I identified 27 studies[154, 201, 211–235] examining crude correlations or adjusted associations of potential biomarkers of dairy fat intake including **C14:0** (n=12)[213, 216, 217, 220, 223–225, 228–230, 232,

235], **C15:0** (n=24)[211–220, 222–235], **C17:0** (n=20)[211–220, 222, 225, 228–235], **trans-16:1n-7** (n=12)[154, 201, 217, 219, 221, 223–226, 229, 230, 232] and **trans-18:1n-7** (vaccenic acid; n=2)[217, 229]. These associations were with total dairy products (n=16)[211–213, 215, 216, 219, 220, 222, 224–226, 228, 230, 231, 233, 235], high-fat dairy products (n=11)[154, 201, 214, 221, 223–225, 229, 231, 232, 234], low-fat dairy products (n=6)[201, 223–225, 231, 232], full-fat milk (n=5)[201, 218, 221, 223, 224], low-fat milk (n=5)[201, 221, 223, 224, 231], cheese (n=5)[212, 215, 221, 226, 231], high-fat cheese (n=3)[201, 223, 224], low-fat cheese (n=3)[201, 223, 224] and butter (n=8)[201, 212, 218, 221, 223, 224, 227, 231] consumption, whereas for yoghurt the number of studies identified was small[223, 231]. The first study on this hypothesis examined the correlations of estimated dairy consumption with C15:0 and C17:0 in adipose tissue[211]. The estimated correlations in this study were high enough (r=0.34-0.61 and r=0.2-0.35 between total dairy consumption and C15:0 or C17:0 respectively) to generate further interest in OCSFAs as potential biomarkers of dairy consumption[211] (**Table 6.1**).

When comparing results from such studies, it is important to account for the biological sample used as different samples have different reference time periods. There is a constant exchange of fatty acids between the different tissues from the stage right after their intake from the diet to their introduction to the blood circulation in the form of chylomicrons, their uptake from different tissues to their storage in the adipose tissue[236]. Due to this constant movement, the fatty acid content of the different tissues is correlated[236]. However, according to the reference period of intake we are interested in, we might achieve a higher correlation with dietary intake if we choose the most relevant sample, but also the most precise method of dietary assessment. For example, it seems that dairy consumption from multiple food diaries had overall higher correlations with blood fatty acids than consumption from multiple 24-hour recalls[213] or FFQ[215], and intakes from 24-hour recalls in turn had higher correlations than intakes from FFQ[220] (Table 6.1). The biological samples used in the majority of the studies to measure fatty acids are adipose tissue, blood (total fatty acids or non-esterified fatty acids or fatty acids from phospholipids, cholesteryl esters or triglycerides) and erythrocytes[236]. Biomarkers in adipose tissue reflect average intake over 1 – 1.5 year, biomarkers in erythrocytes reflect average intake over a few months -as the life cycle of erythrocytes is approximately 120 days- and biomarkers in serum and plasma samples reflect more short-term intake over a few hours to weeks[237].

Of the studies identified (Table 6.1) , seven analysed adipose tissue samples[211, 213–216, 222, 235], four erythrocytes[217, 218, 225, 232], 23 serum or plasma samples (phospholipids, n=13[154, 201, 212, 213, 216, 219–221, 223, 224, 228, 229, 233]; cholesteryl esters, n=2[212, 213]; total plasma, n=5[217, 225, 230, 232, 234]; total serum, n=3[215, 226, 227]; whole blood, n=1[231]). The majority of the studies assessed dairy consumption with an FFQ, (n=20)[154, 201, 211, 214, 215, 217–228, 231, 232, 235],

while seven studies used a food diary[211–213, 215, 216, 229, 233] and two studies used a 24-hour recall[213, 220]. C15:0 was moderately correlated with total and high-fat dairy consumption and this correlation was stronger than that of C17:0, while the sum of the two biomarkers showed a slightly stronger correlation[211, 213] (Table 6.1). Correlations of C14:0 lie on the same levels as those of C17:0. Trans-16:1n-7 was also correlated with dairy consumption and the correlation was weaker than that of C15:0, but stronger overall than that of C17:0. Correlations with low-fat dairy products and low-fat dairy subtypes were overall low, ranging from -0.01 to 0.17.

6.4 Limitations of odd-chain saturated fatty acids and trans-16:1n-7 as candidate biomarkers of dairy consumption

Although results from existing studies show OCSFAs and trans-16:1n-7 as promising biomarkers of dairy fat, there are several potential limitations of their use as dairy biomarkers. First, the specificity of OCSFAs as dairy biomarkers has been questioned both because they have been associated with other foods and because they do not seem to be specific to individual dairy types. After OCSFAs are produced in the ruminant duodenum by microbial fermentation not only are they transferred to milk, but they are also contained in ruminant meat[208]. Thus, the correlation with dairy consumption might be population-specific depending also on the consumption levels of other food items, such as ruminant meat (e.g. beef), but also fish which can be important food sources of OCSFAs when consumed in relatively high amounts[238, 239]. OCSFAs have also been inversely correlated with alcohol consumption[216, 218, 219, 228, 232] with correlations ranging from -0.1 for erythrocyte C15:0[232] to -0.44 for phospholipid C15:0[216] and from -0.14 for erythrocyte C17:0[218] to -0.52 for phospholipid C17:0[216]. A potential explanation for these correlations is that alcohol consumption increases the substrate for the production of the even-chain fatty acids C16:0 and C18:0, which is acetyl-CoA through the conversion of ethanol to acetate in liver and the subsequent conversion of acetate to acetyl-CoA[240]. Because fatty acids are measured in relative concentrations, it can be inferred that an increase of even-chain saturated fatty acids will lead to a decrease of the relative concentrations of other fatty acids i.e. OCSFAs.

Second, while OCSFAs and trans-16:1n-7 in contrast to even-chain saturated fatty acids were thought not to be produced endogenously, but to be synthesised only by the ruminant bacteria[241], there has been evidence that they may be derived from alternative biological pathways as well. This became evident for C17:0 when it was observed that while the ratio of C15:0 to C17:0 in milk is approximately 2:1, the same ratio in human sample tissues overall is reversed and approximately 1:2 and thus C17:0 must be derived

Table 6.1 Previously published crude correlations or adjusted associations of fatty acids C14:0, C15:0, C17:0 and trans-16:1n-7 with total and high-fat dairy products

	C14:0		C15:0		C17:0		Trans-16:1n-7	
Dairy products	**Total**	**High-fat**	**Total**	**High-fat**	**Total**	**High-fat**	**Total**	**High-fat**
Blood								
Assessment of changes †								
Albani, 2015[231]	-	-	↑	↑	↑	↑	-	-
Plasma								
Correlations								
Yakoob,2016[232]	-	-	-	0.26	-	0.19	-	0.24
Yakoob,2014[225]	0.08	0.05	-	0.19	-	0.12	-	0.10
Sun, 2007[217] ‡	0.18, 0.17	-	0.28, 0.29	-	0.13, 0.13	-	0.19, 0.20	-
Assessment of changes †								
Jenkins, 2017[234]	-	-	-	↑	-	↔	-	-
Abdullah, 2015[230]	↔	-	↑	-	↑	-	↔	-
Serum								
Correlations								
Santaren, 2014[226]	-	-	0.20	-	-	-	0.00	-
Brevik, 2005[215] §	-	-	0.37, 0.27	-	0.22, 0.09	-	-	-
Phospholipids								
Correlations								
Da Silva, 2014[154]	-	-	-	-	-	-	-	0.15
de Oliveira Otto,2013[224]	0.14	0.09	0.22	0.16	-	-	0.07	0.13
Mozaffarian, 2013[223]	-	0.10	-	0.15	-	-	-	0.15
Micha, 2010[201]	-	-	-	-	-	-	-	0.39
Saadatian-Elahi, 2009[220] ‖	0.44, 0.25	-	0.33, 0.15	-	-0.08, 0.13	-	-	-
Thiebaut,2009[219]	-	-	0.13	-	0.10	-	0.00	-
Rosell, 2007[216]	0.27	-	0.43	-	0.23	-	-	-
Wolk, 2001[213] #	0.41, 0.36	-	0.40, 0.46	-	0.17, 0.26	-	-	-
Smedman, 1999[212]	-	-	-	0.34	-	0.00	-	-
Assessment of changes †								
Weitkunat, 2017[233]	-	-	↔	-	↔	-	-	-
Nestel, 2014[229]		↔	-	↑	-	↔	-	↑
Mozzafarian, 2010[221]	-	-	-	-	-	-	-	↑
Cholesteryl esters								
Correlations								
Wolk, 2001[213] #	0.35, 0.39	-	0.32, 0.40	-	0.22, 0.36	-	-	-
Smedman, 1999[212]	-	-	-	0.46	-	0.00	-	-
Erythrocytes								
Correlations								
Yakoob,2016[232]	-	-	-	0.16	-	0.13	-	0.21
Yakoob,2014[225]	-0.01	0.01	-	0.07	-	-0.01	-	0.09
Sun, 2007[217] ‡	0.12, 0.11	-	0.22, 0.23	-	0.11, 0.14	-	0.18, 0.18	-
Adipose tissue								
Correlations								
Laursen, 2018[235]	0.34	-	0.39	-	0.25	-	-	-
Aslibekyan,2012[222]	-	-	0.34	-	0.16	-	-	-
Rosell, 2007[216]	0.47	-	0.52	-	0.3	-	-	-
Brevik, 2005[215] §	-	-	0.39, 0.25	-	0.06, 0.16	-	-	-
Baylin, 2002[214]	-	-	-	0.31	-	0.31	-	-
Wolk, 2001[213] #	0.48, 0.58	-	0.58, 0.69	-	0.20, 0.22	-	-	-
Wolk, 1998[211]**	-	-	0.61, 0.34	-	0.35, 0.20	-	-	-

*Five randomised controlled trials[229–231, 233, 234], eight prospective cohort studies[201, 211–213, 221, 223, 224, 232], one case-cohort study[235], three nested case-control studies[217, 219, 225], one case-control study[222], and six cross-sectional studies[154, 214–216, 220, 226] were included

†Changes were assessed with linear regression[221, 229–231] or t-tests [233, 234]

‡The first correlation is derived from cross-sectional data and the second from 4-year prospective data

§The first correlation is derived from 14-day food diary data and the second from food frequency questionnaire data

‖The first correlation is derived from 24-hour recall data and the second from food frequency questionnaire data

#The first correlation is derived from 21-week food diary data and the second from 14 24-hour recall data

**The first correlation is derived from four one-week food diary data and the second from food frequency questionnaire data

from an additional source to a higher extent than C15:0[242]. Even-chain fatty acids are produced from repeated condensation of malonyl-CoA with acetyl-CoA[243]. Under presence of propionate, a short-chain fatty acid produced by the gut microbiome, propionyl-CoA is produced, which competes with acetyl-CoA for the condensation of malonyl-CoA and which then produces OCSFAs as shown in **Figure 6.1**[233, 244, 245].Through this mechanism, it has also been proposed that higher fibre intake might lead to higher production of OCSFAs[233]. A second alternative pathway is α-oxidation of even-chain saturated fatty acids C16:0 and C18:0 to produce C15:0 and C17:0 respectively[208]. This pathway is better characterised for C17:0 from two animal studies[234]. In the first study, infusion with C18:0 led to higher serum levels of C17:0 in rats[234]. In the second study, supplementation with phytol, which produces phytanic acid, a substrate competing with α-oxidation of C18:0 to produce C17:0, resulted in lower serum levels of C17:0[234]. Concerning trans-16:1n-7, it has been suggested that it is partly derived from the partial β-oxidation of trans-18:1n-7[246].

Schematic representation of the odd-chain saturated fatty acid production removed for copyright reasons. Copyright holder is Karolin Weitkunat.

Fig. 6.1 Mechanism of endogenous odd-chain saturated fatty acid production from propionate. Figure adapted from Weitkunat et al[233]

Third, technical error from the laboratory methods employed can lead to higher variation. The low levels of OCSFAs in sample tissues might lead to unstable measurements[247]. In addition, although trans-18:1n-7 is the predominant trans fatty acid in milk, human sample contents of trans-16:1n-7 and trans-18:1n-7 are similar[217, 248]. It has been suggested that depending on the choice of gas chromatography capillary column and the temperature, there might be co-elution of trans-16:1n-7 and iso C17:0 -a branched-chain fatty acid- leading to erroneously higher concentrations of trans-16:1n-7[248].

Finally, a natural source of variation by country can be different ruminant feeding practices related to energy, fat and fibre intake of the animals[33].

In order to account for the limitations of the known potential biomarkers of dairy fat intake, it has been attempted to identify novel biomarkers through metabolomics.

6.5 Use of metabolomics for the identification of novel nutritional biomarkers

Different study designs have been employed to identify or validate a biochemical compound as a nutritional biomarker including animal studies, observational studies and dietary interventions[204]. The traditional method followed in these studies has been to examine associations between food consumption or nutrient intake and one or more candidate biomarkers under a hypothesis-driven deductive approach. Although the inductive, data-driven approach, has been more conservatively used and has even received criticism in the context of research quality due to the lack of an initial hypothesis, it has also been identified as very important for the advancement of science and the generation of new hypotheses[249]. Remarkable examples of how inductivism has contributed to scientific advancements include Darwin's theory of evolution and Watson and Crick's discovery of the DNA structure[249].

With the use of the "-omics" technologies i.e. genomics, transcriptomics, proteomics and metabolomics, spreading in several fields of systems biology research, hypothesis-free, exploratory analyses can be used to identify potential biomarkers of dietary intake. Since metabolomics, which is the study of low-molecular weight metabolites (usually <1,500 Da), is the final downstream product of the genome and thus more directly linked to the phenotype compared to the other "-omics" approaches, it is particularly useful for the identification of potential biomarkers[250]. The use of metabolomics has been exponentially increasing for the last few decades due to the development of high-throughput methods including nuclear magnetic resonance (NMR) and mass spectrometry (MS)[250].

6.6 Metabolomics in epidemiology and nutrition research

Following this trend, the usefulness of metabolomics has been evident also in food and nutrition research with several applications from the identification of food compounds also including contaminants, allergens, toxins, quality and geographical origin[251] to contributions towards personalised nutrition (identification of person's nutritype) and the identification of potential biomarkers of dietary intake for the advancement of dietary assessment[252]. To identify novel nutritional biomarkers, one needs to explore that part of the human metabolome which includes all those metabolites directly derived from foods or produced from the ingestion and metabolism of the food consumed, also known as the food metabolome[207]. The large number of compounds constituting the food metabolome (> 25,000) and the multiple stages of food metabolism that these compounds might be related to, makes the characterisation of the food metabolome highly complex[207]. The fate of a food compound can be one of the following[207]:

- Digestion in the mouth, stomach or small intestine into simple nutrients readily available for absorption. In this case, the food compound is not a very useful biomarker as usually the simple nutrients can be found in multiple foods or even produced endogenously making it hard to trace their dietary source.
- Processing by the gut microbiota. In this case, the end products of the processing might be included in the human metabolome and although some of them might be used as dietary biomarkers, the pathways involved are more complex.
- Absorption and metabolism usually in the liver or kidneys. In this case, the compound and/or its metabolites are released in the circulation or stored tissues and it is more likely that the food compound is a more specific biomarker of a food.

According to a recent literature review of metabolomics studies for the identification of dietary biomarkers, metabolomics has been more extensively used for the identification of biomarkers of fruits, vegetables, meat, fish, bread, whole grain cereals, nuts, wine, coffee, tea, cocoa and chocolate and the classification of individuals according to the dietary pattern they mostly adhere to e.g. vegetarian, Mediterranean diet, a prudent or a Western dietary pattern[207, 252].

6.7 Previous studies on the exploration of dairy biomarkers using metabolomics

So far, 12 studies on the investigation of potential novel dairy biomarkers from metabolomics analyses have been identified[227, 229, 253–262] (**Table 6.2**). Of those, six have assessed blood metabolomic profiles with MS[227, 229, 255, 259, 261] or H-NMR[254], seven

have assessed urine metabolomic profiles with MS[256, 262] or H-NMR[253, 257, 258, 260, 262] and one study has used H-NMR in faecal samples[257] (Table 6.2). Most of the studies were randomised controlled trials (RCTs) and two of the studies were observational (cross-sectional[261] or nested case-control[227]). Of these studies, two included only patients with irritable bowel syndrome[254, 255], one included children[253], one included twins[261] and three included overweight or obese people[229, 258, 260].

Despite the metabolite signals identified from these studies as summarised in Table 6.2, which can be used for the exploration of candidate biomarkers of dairy consumption, several limitations can be identified. The two studies by Pedersen et al have included patients with irritable bowel syndrome[254, 255] with limited generalisability, as the bioavailability of nutrients in these patients might be limited due to symptoms like diarrhoea, which may result in limited absorption. The study by Hjerpsted et al, which assessed potential biomarkers of cheese consumption, used butter with the same amount of fat as a control, which automatically excludes the identification of potential biomarkers of dairy fat[256]. The metabolite signals derived from this study were mainly products of tryptophan or tyrosine metabolism, supporting a higher content of these amino acids in cheese than in butter[256]. Although these amino acids might be specific to cheese when compared with butter, they are not specific to cheese when compared with other foods such as meat[256]. The same holds for other amino acids or products of amino acid metabolism that were reported from other studies such as tyrosine[257, 261], phenylalanine, valine and trimethyl-N-aminovalerate (product of lysine or proline)[261]. In addition, some metabolites reported are not specific to dairy products, although sensitive to dairy consumption. Such metabolites are urinary citrate, which might be produced from the citric acid contained in dairy products, but it may also originate from cranberry juice, grape juice or tea and even produced endogenously[257, 260]; urinary creatinine, creatine and urea, which are markers of protein catabolism[257, 258, 260]; urinary trimethylamine-N-oxide, which is a gut metabolite of choline, a nutrient contained in dairy products, eggs, fish and meat[260]; urinary 3-phenyllactic acid, which is a product of the metabolism of lactic acid bacteria and also contained in honey as a preservative[262]; and urinary hippurate[253, 257, 260], which is a product of the gut microbiota[263].

Some studies identified certain lipid classes[227, 229, 259, 261], such as several shpingomyelins [259, 261] and phospholipids[229, 259, 261], as potential dairy biomarkers. These metabolites could be identified as potential biomarkers of dairy fat, but since they constitute broad lipid classes, lipids specific to dairy consumption need to be identified. In addition, such metabolites may not be relevant for very low-fat dairy products such as skimmed milk, for which non-lipid molecules are more likely to be potential biomarkers. Finally, urinary lactose, galactose and galactonate as potential biomarkers of milk consumption and urinary 3-phenyllactic acid as potential biomarker of cheese consumption reported in the study by Munger et al. could be very specific for this study[262]. As

the authors mentioned, excess amounts of lactose and galactose coming from the 600 ml of milk used in the intervention may result in increased urinary excretion of these and also galactonate, which might not be relevant for lower habitual consumption levels of milk[262]. Thus, while these could be good biomarkers, they might not be useful for dietary assessment or clinical applications in real-life settings.

Overall, single molecules in metabolomics are less likely to serve as specific biomarkers to individual dairy types. Thus, instead of investigating single molecules as potential biomarkers as done in previous metabolomics or nutritional biomarker studies, we aimed to develop metabolite scores. The combination of selected molecules to create scores might be more specific to the consumption of an individual food compared to single biomarkers[207]. Especially since the candidate biomarkers of dairy consumption do not discriminate between different dairy types, the use of a compound score of multiple biomarkers may effectively reflect individual dairy types.

Table 6.2 Discovery studies of potential novel biomarkers of dairy consumption using metabolomics

Author, year	Dairy type	Study design	Sample population	Biomarkers identified	Direction of association
MS*					
Blood					
Pallister, 2017[261]	total dairy products	cross-sectional	twins	trimethyl-N-aminovalerate, SM*(OH) C14:1	↑
Nestel, 2014[229]	high-fat dairy products	RCT*	overweight / obese	LPC* 15:0, LPC 17:0, LPC(O) 20:0, LPC(O) 22:1	↑
Meikle, 2015[259]	high-fat dairy products (control: soy products)	RCT, cross-over	-	SM 32:0, SM 34:0, PC* 29:0, PC 28:0, PC 30:0, PC 32:1, PC 36:3, PC 38:3, PC 38:5, PI 34:1, PI 36:4, PI 38:4	↑
Guertin, 2014[227]	butter	nested case-control study	-	methyl palmitate (15 or 2), 10-undecenoate	↑
Pallister, 2017[261]	butter	cross-sectional	twins	SM(OH) C14:1, PC aa C28:1	↑
Pallister, 2017[261]	milk	cross-sectional	twins	trimethyl-N-aminovalerate, uridine, SM(OH) C14:1, PC aa C28:1, phenylalanine, tyrosine, valine	↑
Pallister, 2017[261]	milk	cross-sectional	twins	1,5-Anhydrogucitol, erythronate	↓
Pedersen, 2011[255]	probiotic milk	RCT, parallel	IBS* patients	lactate, glutamine, proline, creatinine, creatine, aspartic acid	↑
Urine					
Munger, 2017[262]	milk	RCT, crossover	-	lactose, galactose	↑
Hjerpsted, 2014[256]	cheese (control: butter)	RCT, crossover	-	indoxyl sulfate, xanthurenic acid, tyramine sulfate, 4-hidroxyphenylacetic acid, isovalerylglutamic acid, isovaleryl-glycine, tiglylglycine, isobutyrylglycine	↑
Munger, 2017[262]	cheese (control: soy-based drink)	RCT, crossover	-	3-phenyllactic acid	↑
H-NMR*					
Blood					
Pedersen, 2010[254]	probiotic milk	RCT, parallel	IBS patients	L-lactate, 3-hydroxybutyrate	↑
Urine					
Zheng, 2016[260]	high-fat dairy products (control: diet low in high-fat dairy products)	RCT, parallel	overweight / obese	citrate, creatinine, urea	↑
Zheng, 2016[260]	high-fat dairy products (control: diet low in high-fat dairy products)	RCT, parallel	overweight / obese	trimethylamine-N-oxide, hippurate	↓
Bertram, 2007	milk	RCT	children	hippurate	↓
Munger, 2017[262]	milk (control: soy-based drink)	RCT, crossover	-	galactonate	↑
Zheng, 2015 (1)[257]	semi-skimmed milk (control: water)	RCT, crossover	-	citrate, creatinine, urea	↑
Zheng, 2015 (2)[258]	skim milk (control: water)	RCT, parallel	overweight	urea	↑
Zheng, 2015 (1)[257]	cheese (control: milk)	RCT, crossover	-	prolinebetaine, tyrosine, hippurate	↑
Feaces					
Zheng, 2015 (1)[257]	semi-skimmed milk (control: water)	RCT, crossover	-	acetate, butyrate, propionate, glycerol	↑
Zheng, 2015 (1)[257]	cheese (control: milk)	RCT, crossover	-	acetate, butyrate, propionate, malonate	↑

*Abbreviations: H-NMR: Proton Nuclear Magnetic Reasonance; IBS: Irritable Bowel Syndrome; LPC: Lyso-phosphatidylcholine; MS:Mass Spectrometry; RCT: Randomised Cotrolled Trial; PC: Phosphatidylcholine; SM: Sphingomyelin

6.8 Study aims

The aims of this study were to

1. Identify metabolites predicting dairy consumption using metabolomics
2. Develop and validate metabolite scores that reflect consumption of total and types of dairy products
3. Examine whether the metabolite scores could provide predictive ability in addition to C15:0, C17:0 and trans-16:1n-7, which have been identified as candidate dairy biomarkers

6.9 Methods

The reporting of the process for the development and validation of the dairy prediction models was based on the TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis) guidelines[264].

6.9.1 Study design and population

Targeted metabolomics data from the Fenland study comprised the discovery set for the exploration of potential biomarkers of dairy consumption and development of the metabolite scores. Untargeted metabolomics data from the EPIC Norfolk incident diabetes case-cohort study comprised the external validation set. The discovery and validation sets were selected to investigate associations between metabolite scores and type 2 diabetes risk in EPIC Norfolk study (Chapter 7), for which the metabolite scores must be derived from an independent set.

Discovery and internal validation sets

A summary of the study design is presented in **Figure 6.2**. In the present analysis, we evaluated 10,281 participants from the baseline assessment of the Fenland study (section 4.3) after exclusion of those with no metabolomics data (n=1,751), and no dietary data (n=14), pregnant women (n=3), men with energy intake less than 800 kcal, or more than 4,000 kcal or women with energy intake less than 500 kcal or more than 3,500 kcal (n=200) and participants with more than the 50% of the metabolites missing (n=186). The sample was split into two subsets: a discovery set and a validation set (internal validation). The validation set was selected based on the availability of measurements for the fatty acids C15:0, C17:0 and trans-16:1n-7 (n=4,246), because the third aim of this study was to assess the additive value of the metabolite scores to models with fatty acids as candidate dairy biomarkers and the scores would be derived ideally from an independent set. Fatty

acids were measured in samples randomly chosen by picking every other box up to 5,000 samples covering the whole cohort (approximately 12,000 samples). The rest of the samples constituted the discovery set (n=6,035).

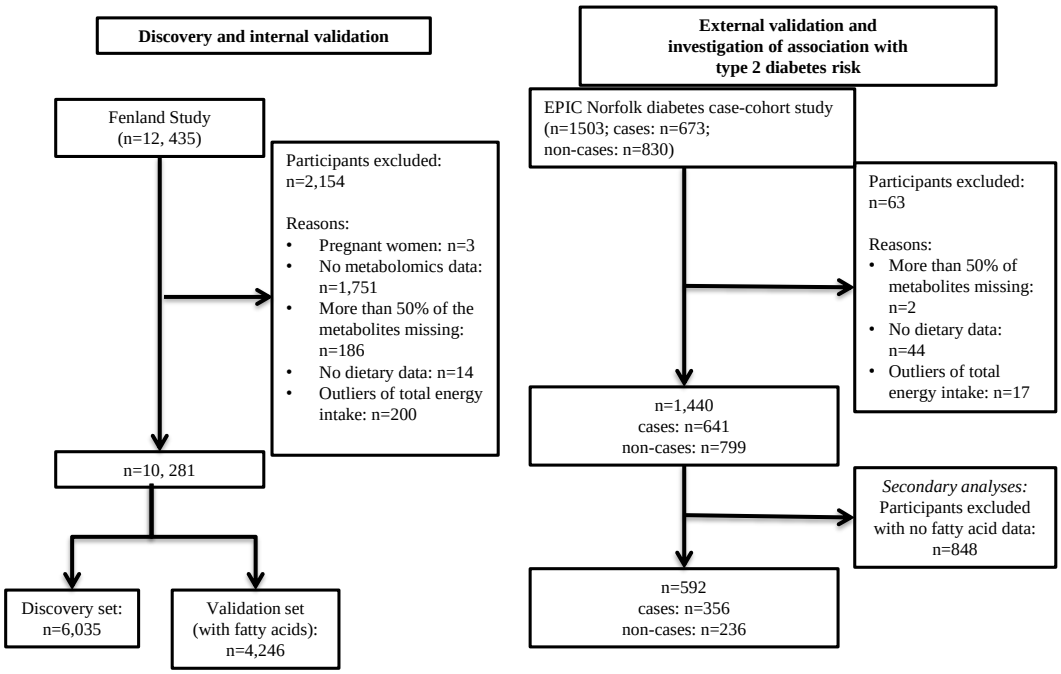


Fig. 6.2 Flow diagram of the inclusion process of participants in the discovery, internal and external validation sets

External validation set

The external validation was performed in a diabetes case-cohort study of 1,440 participants (initial sample: 1,503) nested within the EPIC Norfolk study (section 5.3) after exclusion of participants who were missing dietary data at baseline or follow-up (n=44), men with energy intake less than 800 kcal or more than 4,000 kcal or women with energy intake less than 500 kcal or more than 3,500 kcal (n=17) and participants with more than 50% of the metabolites missing (n=2). In the secondary analysis, where OCSFAs were also included in the discovery, we further excluded participants with no fatty acid measurements (n=848) leaving 592 participants for this analysis (356 incident diabetes cases and 236 non-cases).

6.9.2 Metabolite and fatty acid measurement

As mentioned (section 6.5) the two laboratory techniques for metabolomic analyses are MS and NMR. MS is usually coupled with liquid (LC) or gas (GC) chromatography. In large epidemiological studies, LC might be preferable, with much smaller preparation times than that of GC, which requires a derivatisation step in the beginning, in order to make the compounds volatile[265]. However, LC has some disadvantages compared to

GC. For example, ion suppression, which is the unsuccessful ionisation of some molecules due to the competition with other molecules is more common in LC than GC resulting in GC being more sensitive especially for smaller compounds[265]. Both MS and NMR have advantages and disadvantages and no method can assess the global metabolome. NMR is a fast method (2-3 minutes/sample), as it does not require any prior derivatisation or separation and it is quantitative, but it has low sensitivity, is expensive, requires relatively large samples and cannot detect inorganic compounds[266]. Although NMR was more popular initially, MS has been gaining more ground in epidemiological studies, due to the lower and decreasing cost, the requirement of small samples, and its high sensitivity, which has made it very useful also in the identification of nutritional biomarkers, which are expected to be detected in low concentrations[265]. However, MS might have low reproducibility, because of the variation in the pH and the column temperature and column ageing or contamination[267]. These problems can be handled with the use of a buffer such as formic acid or trifluoroacetic acid to maintain the pH, the use of a thermostat and the use of a pump system to re-equilibrate the column[267]. In addition, MS is less quantitative and slower (20-30 minutes/sample) than NMR[266]. In both of the studies, metabolomics profiling was conducted with MS. More details on the specific methods are presented in the following sections.

The reporting of the methods for the metabolite measurement is based on the proposed reporting standards by the Chemical Analysis Working Group as part of the Metabolomics Standards Initiative[268].

Metabolomics assay in the discovery and internal validation sets

Fasting blood samples were collected after an overnight fast. For the preparation of blood plasma, blood was mixed with heparin, an anticoagulant, which does not interfere with the analytical samples[269]. Samples were stored in -80°C in the MRC Epidemiology Unit Biorepository Freezer until analysed. Before the analysis, the samples were removed from the freezer, placed on roller mixers at 30 rpm to thaw for 15-20 minutes and centrifuged for 1 minute at 2,000 rpm. Calibration standards and internal standards were also briefly centrifuged and quality control (QC) samples were centrifuged for 5 minutes at 2,750 rpm. All vials were then shaken for 15 minutes at 1,200 rpm and vortexed and then transferred to 96-well plates with rows A-H and columns 1-12. Derivatisation agent (50 ul) was added to all the wells. The internal standards (10 ul) were added to all wells for quantification apart from A1, which included the blank (deionised water with phosphate buffer saline). Positions B2-D2 contained the zero samples (blank samples with internal standards). Positions E1-C2 contained the calibration samples, positions D2-F2 contained the commercial QC samples, positions G2 and H12 contained the pooled QC samples and positions H3-G12 contained the study samples. The sample processing was done on a Hamilton STAR liquid handling station (Hamilton Robotics Ltd, Birmingham, UK).

Targeted metabolomic profiling of 10,684 blood plasma samples was performed with a commercial kit (Absolute IDQ p180 kit, Biocrates Life Sciences AG, Innsbruck, Austria). Flow injection analysis (FIA) MS (AB SCIEX 5500 Qtrap mass spectrometer, Sciex Ltd, Warrington, UK), isocratic with methanol, in positive ionisation mode was performed for lipids and acylcarnitines , FIA-MS in negative ionisation mode was performed for hexose, while for amino acids and biogenic amines, ultra-performance liquid chromatography MS (UPLC-MS; Waters ltd, Machester, UK couples to ABSciex 5500 Qtrap mass spectrometer, Sciex Ltd, Warrington, UK) was performed with a 5-minute gradient elution starting with 100% water and changing to 95% acetonitrile and 0.2% formic acid over a Waters Acquity UPLC BEH C18 column (2.1 x 50 mm, 1.7mm).

The raw metabolomics data were processed with the MetIDQ software (provided with the kit). The data went through checks for outliers, non-detects, peak picking, normalisation using the QC samples (metabolite values were divided by the ratio of the mean value of QC samples within a plate to the mean value of QC samples across all the plates) and finally batch correction. Normalisation is very important to distinguish between biological variation and noise due to column degradation, changes in room temperature or pH of the mobile phase, matrix effects or the repeated use of the ion source and results can be substantially different when it is not applied[265].

Values under detection limits were replaced with numbers randomly selected from a uniform distribution between 0.1 and a minimum observed within batch. After exclusion of 13 metabolites based on the results from the QC and the batch correction, 174 metabolites were left including amino acids (n=22), biogenic amines (n=12), acylcarnitines (n=40), phosphatidylcholines (PCs) (n=74), lysophosphatidylcholines (LPCs) (n=14), sphingomyelins (SMs) (n=11) and hexose. Amino acids and biogenic amines were quantified and expressed in absolute scale, whereas the lipids, acylcarnitines and hexose were semi-quantified due to lack of standards and thus expressed in a relative scale.

Metabolomics assay in the external validation set

Non-fasting blood samples were collected at baseline from the 1,503 participants of the diabetes case-cohort study nested within the EPIC Norfolk study. Samples were stored in -80°C until one day before the analysis and overnight in liquid nitrogen the night before the analysis. Deproteinisation was done with methanol, which has been shown to promote an optimised extraction and deproteinisation[270], under vigorous shaking for two minutes (Glen Mills GenoGrinder 2,000). The samples were then centrifuged and placed on a TurboVap (Zymark) to remove the organic solvent. The order of the study samples on the plates was randomised and QC samples (two commercial QC standards and four pooled QC samples) were evenly distributed across the study samples. A blank sample and internal standards, which compensate for the technical noise that might be introduced to the analytical process[269] were also used.

Untargeted metabolomic profiling was performed in 1,503 blood citrated plasma samples (DiscoveryHD4® platform, Metabolon, Inc.; Waters ACQUITY UPLC, Thermo Scientific Q-Exactive spectrometer with a heated electrospray ionization (HESI-II) source and Orbitrap mass analyzer with 35,000 mass resolution). Samples were divided into five fractions of which two were used for reverse phase UPLC tandem MS (MS/MS) with positive ion mode electrospray ionization (ESI; mobile phases with 0.1% formic acid in water and 0.1% formic acid in methanol), one was used for reverse phase UPLC MS/MS with negative ion mode ESI (mobile phases with 6.5 mM ammonium bicarbonate in water, pH=8 and 6.5 mM ammonium bicarbonate in 95% methanol and 5% water) , one was used for hydrophilic interaction liquid chromatography (HILIC) UPLC MS/ MS with negative ion mode ESI (mobile phases with 10 mM ammonium formate in 15% water, 5% methanol, 80% acetonitrile and 10 mM ammonium formate in 50% water and 50% acetonitrile) and one was kept as a backup. The UPLC column used was a 2.1 mm x 100 mm Water BEH C18 1.7 mm at 40°C and the HILIC column was a 2.1 mm x 150 mm Waters BEH Amide 1.7mm at 40°C. Gradient elution with water, methanol, 0.05% pentafluoropropionic anhydrate and 0.1% formic acid was used in all the columns.

Metabolite identification was based on the retention index, accurate mass matched to the Metabolon internal library of metabolites within a range of ± 10 ppm and the standards used. The raw metabolomics data were processed and went through normalisation by "block correction" i.e. setting median to 1 and adjusting each data point proportionately and QC checks with in-house software from Metabolon, Inc. After QC procedures, 940 metabolites where quantified, of which 308 were unknown.

Approximate metabolite matching between the targeted and untargeted platforms

Due to the different assays used in the Fenland study (targeted platform) and the EPIC Norfolk study (untargeted platform), some of the molecules included in one dataset could not be exactly matched with similar molecules of the other dataset. Thus, in order to perform the external validation, matching of the metabolites between the two metabolomic platforms was performed.

Fatty acids assay

Plasma phospholipid fatty acids (38 individual fatty acids) were measured in 4,791 participants in the Fenland study and 592 participants in the EPIC Norfolk diabetes case-cohort study. For the measurement of fatty acids, the same method was used in both studies. Phospholipids were isolated from total plasma lipids with solid-phase extraction and then free fatty acids were extracted with hydrolysis and derivatisation. The free fatty acids were methylated to form fatty acid methyl-esters, which are volatile and were then used in GC (7890N; Agilent Technologies) with a 30 m capillary column with a diameter of 0.25 mm. Sample was processed with sequential multipurpose sampler systems (Gerstel GmbH &

Co. KG, Mulheim an der Ruhr, Germany). One blank sample and two QC samples were used in all the batches: one from human plasma pooled from the study samples and one from horse plasma. Standards of all fatty acids were used for the GC calibration. Study and QC samples were stored in -80°C before analysis. Fatty acids were identified by comparing their retention times with those of the standards. Fatty acids were expressed in relative amounts. More details of the methods applied were published previously[271].

6.9.3 Metabolite annotation

Nomenclature of amino acids followed the recommendations by the IUPAC-IUB Commission on Biochemical Nomenclature[272]. For all the lipids, "Cx:y" indicates x carbon atoms and y double bonds. For the acylcarnitines, x denotes the number of carbon atoms of the carboxylic acid, which is esterified with carnitine. Thus, carnitine is denoted as C0. For PCs there is the additional indication of "aa" when it contains two fatty acids or "ae", when it contains one fatty acid and one fatty alcohol. So a PC might be denoted as PC aa Cx:y or as PC ae Cx:y. If it is an LPC, then this becomes LPC aa Cx:y or LPC ae Cx:y. Finally, SMs are denoted as SM Cx:y or SM-OH Cx:y if it is a hydroxy-SM.

6.9.4 Dietary assessment

In both studies, an 130-item semi-quantitative FFQ was used for the dietary assessment reflecting habitual dietary intakes over the past year. The FFQ was validated against 7-day food diaries in the EPIC Norfolk study[158]. Participants could choose one from the nine frequencies of dairy consumption ranging from "never or less than once/month" to "6 times per day" and provide more details on the type and amount of milk consumed. Correlations between the dairy consumption as reported in the questionnaires and as reported in the 7-day diaries were 0.56 for milk, 0.57 for yoghurt, 0.33 for cheese and 0.54 for butter. Questionnaire data were processed with the FETA software[159]. Dairy outcomes included milk, yoghurt, cheese, butter and total dairy products (sum of milk, yoghurt, cheese, butter and cream).

6.9.5 Statistical analysis

Data processing

Further metabolomic and fatty acid data pre-treatment was performed, which is an important step to emphasise biological information and interpretability[273]. This involved log-transformation to reduce heteroscedasticity, centering by subtracting the mean to leave only the relevant variation and auto-scaling by dividing by the standard deviation to make the metabolites comparable[273]. Missing values were imputed using multiple imputation with chained equations[168] and due to the low number of missing values (**Tables 6.5**

and A.18), a single imputed dataset was used in subsequent analyses. Data were processed in Stata 14.2 (College Station, TX: StataCorp LP, 2015).

Due to the skewed distributions of most of the dairy products, which did not approximate the normal distribution even after transformation, binary variables of low and high consumers were created for milk, yoghurt, cheese, butter, and total dairy products (sum of milk, yoghurt, cheese, cream and butter). Milk consumption, which was more frequently consumed than the other dairy types, was divided into <1 serving/day and ≥ 2 servings/day. Yoghurt, cheese and butter consumption were divided into <1 serving/week and ≥ 1 serving/day. Total dairy consumption was divided into <1 serving/day and ≥ 3 servings/day. The log-transformed energy densities of milk and total dairy products (in 2,000 kcal of total energy intake), which approximated the normal distribution were also used as outcomes.

Methods for metabolomics data analysis

Metabolomics data can be analysed with the use of multiple univariate models for each metabolite or the use of a multivariable model including all the metabolites. The use of a multivariable method has the advantage of accounting for the correlations between the explanatory variables i.e. the metabolites and thus their relations, which might be more informative than the associations with each metabolite separately[274]. In addition, we can avoid the use of multiple test correction, which in some cases may increase the risk of false negatives[274].

Multivariable methods can be categorised into unsupervised and supervised. Unsupervised methods do not get any input about which variable we would like to base our predictions on, in contrast to supervised methods. Unsupervised methods reported in metabolomics studies include principal component analysis (PCA), which can be used in a hypothesis-free exploration or for dimensionality reduction, but also other machine learning methods such as k means clustering and parallel factor analysis[275]. Although unsupervised methods such as PCA are used widely and routinely in nutritional and other metabolomics studies[252], they are not very useful for nutritional biomarkers identification, because we are interested in the discrimination of the sample based on a specific dietary consumption profile, which is not very likely the same as the discrimination that the unsupervised methods will show[275]. Supervised methods used in metabolomics studies include regression and its variations e.g. penalised regression, discriminant analysis methods e.g. partial least squares-discriminant analysis (PLS-DA), which is the most extensively used method in nutritional metabolomics studies[252] or other machine learning methods such as random forests, support vector machines and artificial neural networks[275].

All the methods have their advantages and disadvantages and the choice of a method is often a trade-off between interpretability, easy and fast application, and flexibility. For example, while ordinary regression is easily interpretable, in our case it might not be very

useful. We expect high multicollinearity between metabolites, which might be handled less effectively with the ordinary regression. PLS-DA is a method, which easily over-fits the data, so a well-designed validation and caution with the interpretation of the results are needed, when this method is applied[274]. Machine learning techniques can effectively deal with high-dimensional data, they do not require assumptions that other methods do, and they empirically have a high prediction accuracy but at the same time some of them introduce the risk of over-fitting, are time consuming and difficult to apply and interpret[276–281].

Regularisation or penalisation or shrinkage methods apply a penalty on the regression coefficients leading to biased effect estimates, but simultaneously decreasing the variance introduced due to mutlicollinearity between the different explanatory variables (e.g. metabolites)[280]. The three main types of penalised regression are the Ridge, LASSO (Least Absolute Shrinkage and Selection Operator) and elastic net regression and a comparison of their characteristics is presented in **Table 6.3**. The first penalised method was Ridge regression, which applies a penalty on the sum of squares of the regression coefficients (L2 norm), keeps all the coefficients in the model (no shrinkage to zero) and applies similar weights to highly correlated variables[282]. After Ridge regression, the LASSO regression was introduced, which applies a penalty to the absolute values of the regression coefficients (L1 norm), shrinks some coefficients to zero, thus dropping variables out of the model and keeps only one out of a group of perfectly correlated variables[283]. LASSO regression gives a better prediction accuracy than Ridge regression, because it decreases the variation after variable selection[283]. On the other hand, this variable selection attribute makes LASSO less stable than Ridge[283]. Another limitation of the LASSO is that the sample size poses a restriction to the number of variables that can be selected to be kept in the model[283]. The third main type of penalised methods among several variations of the LASSO[284], is elastic net, which is a combination of the former two methods, as it applies both an L1 and an L2 penalty at a ratio, which can be defined[285]. Elastic net combines the advantages of both methods and overcomes their disadvantages, and it has been characterised as the stabilised version of LASSO, as it applies equal weights on absolutely correlated variables, without dropping them, thus giving the ability to investigate group effects and the number of variables to keep in the model is independent of the sample size[284]. These penalties can be used both in linear and logistic regression models.

Discovery and validation of dairy prediction models

Elastic net regression had the lowest prediction error when compared with Ridge, LASSO, principal components, partial least squares, support vector and random forest regression in a metabolomics study[280]. Due to its suitability for high-dimensional data with multicollinearity, its advantages over the other penalised methods and its relative simplicity

Table 6.3 Comparison of characteristics of the three main penalised regression methods

	Ridge	**LASSO**	**Elastic Net**
Penalty	$\sum \beta^2$ (L2)	$\|\beta\|$ (L1)	$\sum \beta^2$ and $\|\beta\|$ (L1+L2)
Shrinkage to zero	No	Yes	Yes
Perfectly correlated variables	Equal weights	Keeps only one in the model	Equal weights
Relative prediction accuracy	Lower	Higher	Higher
Relative stability	Higher	Lower	Higher
Restriction on number of variables by sample size	No	Yes	No

Abbreviations: LASSO: Least Absolute Shrinkage and Selection Operator

compared to other machine learning methods, elastic net regression was applied for the development of dairy prediction models.

These models were used for each dairy type specified in three sets of analyses. The first set included all the 174 metabolites from the targeted metabolomics platform of the Fenland study. The second set included only the 82 metabolites, which were matched with the metabolites from the untargeted metabolomics platform of the EPIC Norfolk diabetes case-cohort study. The third set included the 82 matched metabolites and the two OCSFAs C15:0 and C17:0.

For binary dairy outcomes, elastic net logistic regression with a stochastic gradient descent (SGD) classifier was used in Python Scikit Learn module v0.19.1[286]. Prior to application, fine-tuning of the SGD algorithm was performed for all the different models by level of adjustment (**Table 6.4**) using the Scikit Learn GridSearchCV function, which used exhaustive combinations of input values for several parameters of the algorithm and 3-fold cross-validation to give the set of parameters which resulted in the best predictions. In our case, we tuned the parameters of L1 ratio (inputs: 0.05, 0.15, 0.25, 0.35, 0.45, 0.55, 0.65, 0.75, 0.85), which defines whether the elastic net penalisation will be closer to a Ridge or a LASSO penalty, alpha (inputs:10^{-15}, 10^{-10}, 10^{-8}, 10^{-4}, 0.001, 0.01, 1, 5, 10, 20), which is a constant that multiplies the penalty, and the maximum number of iterations for the model to reach convergence (inputs: 5, 50, 100, 500, 1,000). For linear dairy outcomes, elastic net linear regression with 3-fold cross validation was used (Scikit Learn function: ElasticNetCV).

Five different levels of adjustment were used in the models (**Table 6.4**). Model 1 (base model) included age (continuous in years), sex, test site (Cambridge, Ely, Wisbech), smoking status (never smoker, ex-smoker, current smoker), physical activity energy expenditure (continuous in kj/kg/day), lipid-lowering medication (Yes, No), hormone replacement therapy (Yes-women, No-women, Men), BMI (continuous in kg/m^2). Model 2 included the metabolites (continuous). Model 3 combined model 1 and model 2. Model 4 additionally included educational level (low, medium, high), socioeconomic status based on occupation (low- technical/semi-routine and routine occupations; medium- lower managerial / intermediate occupations; high- professional/higher managerial occupations) and anti-

Table 6.4 Predictors included in the models for the prediction of dairy consumption

Model	Predictors
Discovery and validation of prediction models	
Model 1	age, sex, test site, physical activity, smoking status, lipid-lowering medication, HRT, BMI
Model 2	Metabolites
Model 3	Model 1 + model 2
Model 4	Model 3 + educational level, occupational status, anti-hypertensive medications
Model 5	Model 3 + total energy intake, fruit, vegetables, cereals, red meat, processed meat, fish, margarine, sweet snacks, SSBs, caffeinated coffee, tea, alcoholic beverages, dietary supplements
Predictive value of metabolite scores for dairy consumption in addition to C15:0, C17:0 and trans-16:1n-7	
Model 1	age, sex, test site, physical activity, smoking status, lipid-lowering, HRT, BMI
Model 2	C15:0, C17:0, trans-16:1n-7 fatty acids
Model 3	Metabolite score
Model 4	Model 1 + model 2
Model 5	Model 1 + model 3
Model 6	Model 4 + total energy intake, fruit, vegetables, cereals, red meat, processed meat, fish, margarine, sweet snacks, SSBs, caffeinated coffee, tea, alcoholic beverages, dietary supplements
Model 7	Model 1 + model 2 + model 3

Abbreviations: BMI: Body mass index; HRT: Hormone-replacement therapy; SSBs: Sugar-sweetened beverages

hypertensive medication (Yes, No). Model 5 was additionally adjusted for dietary variables including total energy intake (kcal/day), dietary supplements (Yes, No) and intakes (g/day) of fruit, vegetables, cereals, red meat, processed meat, fish, margarines, sweet snacks, sugar-sweetened beverages, coffee, tea, alcoholic beverages and dairy products other than the dairy outcome.

Metabolite scores were created from the sum of the metabolites weighted by the elastic net coefficients. Apart from the scores with all the metabolites, we also created scores with the top metabolites only, as the main contributors to dairy prediction and which are more likely to be validated in an independent set without noise that the rest of the metabolites might add. As top metabolites, we defined those that were above the mean+2*SD of the regression coefficients. As per the TRIPOD guidelines, the same models were developed also without any penalisation (Appendix 1 Tables A.21, A.24, A.27)[264].

The internal and external validation were performed by applying the discovery elastic net coefficients as weights to create the metabolite scores in the internal and external validation sets respectively and their inclusion in logistic prediction models (without penalisation).

Assessment of the predictive value of the metabolite scores in addition to C15:0, C17:0 and trans-16:1n-7

The subset used as the internal validation set in the previous analysis, was also used to evaluate whether the metabolite scores have predictive value over the use of C15:0, C17:0 and trans-16:1n-7, which have been described as candidate biomarkers of dairy consumption. Logistic and linear regression models were developed without penalisation with the Statsmodels v0.9.0 Python module. Seven different sets of predictors were

used in the models (Table 6.4). We selected covariates which could potentially predict dairy consumption or potentially confound the association between metabolites and dairy consumption and also be relatively easily available in a clinical setting. Dietary factors were included in secondary models, because the overall scope of this project was to develop metabolite scores, which could be used independent of self-reported dietary data. Model 1 included age, sex, test site, smoking status, physical activity energy expenditure, lipid-lowering medication, hormone replacement therapy and BMI as described in the previous analysis. Model 2 included C15:0, C17:0 and trans-16:1n-7. Model 3 included the metabolite scores standardised. Model 4 was a combination of model 1 and model 2. Model 5 was a combination of model 1 and model 3. Model 6 was model 4 additionally adjusted for the same dietary variables as reported in the previous analysis. Model 7 was the combination of model 1, model 2 and model 3. Metabolite scores including the top metabolites only were used in secondary analyses. Further secondary analyses included additional adjustment for C14:0 and additional adjustment of model 4 for educational level, socioeconomic status and anti-hypertensive medication. Likelihood ratio tests were used to statistically compare the different nested models (with or without the fatty acids and with or without the metabolite scores).

Metrics for the evaluation of the prediction models

According to the TRIPOD guidelines, it is suggested to estimate and report primarily metrics for calibration or discrimination. Areas under the curve (AUC) were estimated (Scikit Learn function: roc_auc_score) and the receiver operating characteristic (ROC) curves (1-specificity against sensitivity) were produced for logistic prediction models (Scikit Learn function: roc_curve). As metrics of the overall performance, R^2 was calculated for the linear models and the accuracy (Scikit Learn function:accuracy_score).

Prediction models were developed and validated in Python 3.6.3 using Jupyter Notebooks [287].

6.10 Results

6.10.1 Descriptive characteristics

Descriptive characteristics of participants in the discovery, internal and external validation sets are presented in (**Table 6.5**).

Participants in the discovery set were 46.8% women with mean (SD) age of 48.9 (7.4) years.Almost half of the participants were of medium educational level and high socioeconomic status. Almost one third of them were current smokers with a mean (SD) BMI of 26.9 (4.8) kg/m^2. Distributions of socio-demographic, lifestyle and dietary variables were similar between the discovery and internal validation sets.

Participants in the diabetes case-cohort study nested within the EPIC Norfolk study were older, of lower educational level and socio-economic status, less frequently smokers, more frequently on anti-hypertensive medication and HRT and consumed more milk, margarines, sweet snacks, tea and less yoghurt and red meat than participants in the discovery and internal validation sets within the Fenland study.

6.10.2 Metabolite matching between the two cohorts

In total, we matched 82 metabolites. Exact matching was possible for 21 out of 22 amino acids, 10 out of 12 biogenic amines and 15 out of 40 acylcarnitines, as shown in **Table 6.6**. For the lipid compounds i.e. PCs, LPCs and SMs, the application of FIA-MS/MS did not allow for differentiation of the fatty acids contained in these lipids based on number of carbon atoms and their type of bond. For this reason, the lipids reported from the targeted metabolomics platform are the sum of all the isobaric (same weight within a range of $\pm$0.5 Da) and isomeric (same number of atoms, but different structure) compounds of these lipids. Likewise, specific hexose sugars could not be identified, so their sum was reported as hexose. On the other hand, the untargeted metabolomics platform identified specific isobaric/isomeric compounds for some lipids, while for others, it identified sum of compounds, but with a higher specificity than the targeted platform. Approximate metabolite matching was performed based on the number of carbon atoms, the number of double bonds and the molecular mass (also accounting for positive or negative ion mode). After this process, approximate matching was possible for nine out of 14 LPCs, 17 out of 74 PCs, eight out of 11 SMs and hexose, that was matched with three sugars (Table 6.6).

Table 6.5 Descriptive characteristics of socio-demographic, lifestyle and clinical factors in the discovery and validation set of the Fenland study and the external validation set of the EPIC Norfolk diabetes case cohort study*

		Discovery set		Internal validation set		External validation set	
Participants (N)		6,035		4,246		1,440	
Socio-demographic factors							
Age (years)		48.9	7.4	47.8	7.3	60.1	9.0
Sex (ref. Men)	Women	46.8		45.5		48.9	
Testsite (ref. Cambridge) †	Ely	35.3		34.9			
	Wisbech	27.2		29.5			
Educational level (ref. Low) ‡ §	Medium	45.9		45.9		37.6	
	High	34.4		33.7		12.3	
Socio-economic status (ref. Low) ‡ §	Medium	19.7		19.5		16.0	
	High	53.8		53.2		42.1	
Lifestyle factors							
Smoking status (ref. Never) ‡ §	Former smoker	53.9		54.2		43.0	
	Current smoker	33.7		32.3		12.7	
Physical activity (ref. Inactive) † ‡	Energy expenditure (kj/kg/d)	53.7	22.1	54.4	22.4		
	Moderately inactive					26.9	
	Moderately active					21.9	
	Active					16.7	
Energy intake (kcal/d) §		1,924	571	1,939	579	2023	578
BMI (kg/m^2) ‡ §		26.9	4.8	26.8	4.7	27.6	4.5
Medications / Supplements							
Lipid-lowering medication (ref. No) ‡	Yes	4.3		3.5		1.9	
Anti-hypertensive medication (ref. No) ‡	Yes	7.2		7.2		23.6	
Hormone replacement therapy (ref. No for women / Men) ‡	Yes	2.9		2.8		16.0	
Dietary supplements ‡ §	Yes	41.0		42.4		54.2	
Types of dairy products							
Milk §	Energy density	1.5	1	1.6	1.0	1.8	0.9
	< 1 serving/d	38.8		37.8		21.5	
	≥2 serving/d	27.0		27.1		41.0	
Yoghurt §	< 1 serving/wk	35.6		36.6		53.5	
	≥1 serving/d	13.1		12.6		7.3	
Cheese §	< 1 serving/wk	27.9		28.0		24.8	
	≥1 serving/d	8.1		9.0		8.2	
Butter §	< 1 serving/wk	51.9		53.9		63.4	
	≥1 serving/d	17.8		18.3		19.4	
Total dairy products §	Energy density	3.0	1.4	3.0	1.4	3.1	1.4
	< 1 servings/d	6.7		7.0		3.9	
	≥3 servings/d	39.1		39.4		43.8	
Non-dairy dietary factors (g/d)							
Fruits §		240.6	203.3	244.8	198.7	241.1	198.9
Vegetables §		258.1	143.2	253.4	135.5	233.1	120.3
Cereals §		169.3	100.6	169.2	96.4	154.4	85.5
Red meat §		74.2	46.7	72.9	48.4	64.3	46.0
Processed meat §		31.7	26.9	32.3	29.3	31.4	25.5
Fish §		42.9	33.4	43.2	35.4	37.2	25.5
Margarines §		7.5	10.0	7.8	10.5	17.2	16.7
Sweet snacks §		90.6	69.1	91.9	70.7	115.9	85.9
Sugar-sweetened beverages §		43.2	95.1	39.9	84.7	41.0	84.8
Coffee §		360.6	348.3	356.6	347.2	385.1	337.4
Tea §		485.7	371.2	492.3	373.2	636.2	370.0
Alcoholic beverages §		151.2	249.5	149.7	238.1	120.7	229.9

*The mean and SD is presented for continuous variables and column percentages are presented for categorical variables

†Testsite is not applicable in the EPIC Norfolk study. Physical activity was objectively assessed in the Fenland study and expressed as a continuous variable for physical activity energy expenditure. Physical activity level was assessed with a questionnaire in the EPIC Norfolk study and categorised into four categories.

‡Missing values for each variable were < 3% with total non-overlapping missing values of 4% across all variables in the discovery set and 6.4% in the internal validation set

§Missing values for each variable were < 2% in the external validation set with total non-overlapping missing values of 3.8% across all variables in the external validation set

Table 6.6 Metabolite matching between the targeted metabolomics platform (Biocrates) in the Fenland study and the untargeted platform (Metabolon) in the EPIC Norfolk diabetes case-cohort study

Metabolite class	Metabolite name (targeted)	Metabolite abbreviation (targeted)	Matched metabolite name (untargeted)	Matched metabolite abbreviation (untargeted)	Precision of matching
Amino acid	Alanine	Ala	alanine	Ala	Exact*
Amino acid	Arginine	Arg	arginine	Arg	Exact*
Amino acid	Asparagine	Asn	asparagine	Asn	Exact*
Amino acid	Aspartate	Asp	aspartate	Asp	Exact*
Amino acid	Citrulline	Cit	citrulline	Cit	Exact*
Amino acid	Glutamine	Gln	glutamine	Gln	Exact*
Amino acid	Glutamate	Glu	glutamate	Glu	Exact*
Amino acid	Glycine	Gly	glycine	Gly	Exact*
Amino acid	Histidine	His	histidine	His	Exact*
Amino acid	Isoleucine	Ile	isoleucine	Ile	Exact*
Amino acid	Leucine	Leu	leucine	Leu	Exact*
Amino acid	Lysine	Lys	lysine	Lys	Exact*
Amino acid	Methionine	Met	methionine	Met	Exact*
Amino acid	Ornithine	Orn	ornithine	Orn	Exact*
Amino acid	Phenylalanine	Phe	phenylalanine	Phe	Exact*
Amino acid	Proline	Pro	proline	Pro	Exact*
Amino acid	Serine	Ser	serine	Ser	Exact*
Amino acid	Threonine	Thr	threonine	Thr	Exact*
Amino acid	Tryptophan	Trp	tryptophan	Trp	Exact*
Amino acid	Tyrosine	Tyr	tyrosine	Tyr	Exact*
Amino acid	Valine	Val	valine	Val	Exact*
Biogenic amines	Phenylethylamine	PEA	-	-	No matching
Biogenic amines	Acetylornithine	AcOrn	N-delta-acetylornithine	AcOrn	Exact*
Biogenic amines	Symmetric dimethylarginine	SDMA	dimethylarginine (SDMA + ADMA)	DMA	Exact*
Biogenic amines	alpha-Aminoapidic acid	alpha-AAA	2-aminoadipate	alpha-AAA	Exact*
Biogenic amines	Creatinine	Creatinine	creatinine	creatinine	Exact*
Biogenic amines	Kynurenine	Kynurenine	kynurenine	kynurenine	Exact*
Biogenic amines	Methioninesulfoxide	Met-SO	methionine sulfoxide	Met-SO	Exact*
Biogenic amines	cis-hydroxyproline	c4OHPro	-	-	No matching
Biogenic amines	trans-hydroxyproline	t4OHPro	trans-4-hydroxyproline	t4OHPro	Exact*
Biogenic amines	Sarcosine	Sarcosine	sarcosine (N-Methylglycine)	sarcosine	Exact*
Biogenic amines	Serotonin	Serotonin	serotonin	serotonin	Exact*
Biogenic amines	Spermidine	Spermidine	-	-	No matching
Biogenic amines	Taurine	Taurine	taurine	taurine	Exact*

Table 6.6 (continued)

Metabolite class	Metabolite name (targeted)	Metabolite abbreviation (targeted)	Matched metabolite name (untargeted)	Matched metabolite abbreviation (untargeted)	Precision of matching
Acylcarnitines	Carnitine	C0	carnitine	C0	Exact*
Acylcarnitines	Acetylcarnitine	C2	acetylcarnitine	C2	Exact*
Acylcarnitines	Propionylcarnitine	C3	propionylcarnitine	C3	Exact*
Acylcarnitines	Propenoylcarnitine	C31	-	-	No matching
Acylcarnitines	Hydroxypropionylcarnitine	C3OH	-	-	No matching
Acylcarnitines	Butyrylcarnitine	C4	butyrylcarnitine	C4	Exact*
Acylcarnitines	Butenylcarnitine	C41	-	-	No matching
Acylcarnitines	Hydroxybutyrylcarnitine	C3DCC4OH	hydroxybutyrylcarnitine	C3DCMC5OH	Exact*
Acylcarnitines	Valerylcarnitine	C5	-	-	No matching
Acylcarnitines	Tiglylcarnitine	C51	tiglyl carnitine	C51	Exact*
Acylcarnitines	Glutaconylcarnitine	C51DC	-	-	No matching
Acylcarnitines	Glutarylcarnitine	C5DCC6OH	-	-	No matching
Acylcarnitines	Methylglutarylcarnitine	C5MDC	3-methylglutarylcarnitine (2)	C5MDC	Exact*
Acylcarnitines	Hydroxyvalerylcarnitine	C3DCMC5OH	-	-	No matching
Acylcarnitines	Hexanoylcarnitine	C6C41DC	hexanoylcarnitine	C6C41DC	Exact*
Acylcarnitines	Hexenoylcarnitine	C61	-	-	No matching
Acylcarnitines	Pimelylcarnitine	C7DC	-	-	No matching
Acylcarnitines	Octanoylcarnitine	C8	octanoylcarnitine	C8	Exact*
Acylcarnitines	Nonaylcarnitine	C9	-	-	No matching
Acylcarnitines	Decanoylcarnitine	C10	decanoylcarnitine	C10	Exact*
Acylcarnitines	Decenoylcarnitine	C101	-	-	No matching
Acylcarnitines	Decadienylcarnitine	C102	-	-	No matching
Acylcarnitines	Dodecanoylcarnitine	C12	laurylcarnitine	C12	Exact*
Acylcarnitines	Dodecanedioylcarnitine	C12DC	-	-	No matching
Acylcarnitines	Dodecenoylcarnitine	C121	-	-	No matching
Acylcarnitines	Tetradecanoylcarnitine	C14	myristoylcarnitine	C14	Exact*
Acylcarnitines	Tetradecenoylcarnitine	C141	myristoleoylcarnitine	C141	Exact*
Acylcarnitines	Hydroxytetradecenoylcarnitine	C141OH	-	-	No matching
Acylcarnitines	Tetradecadienlycarnitine	C142	-	-	No matching
Acylcarnitines	Hydroxytetradecadienylcarnitine	C142OH	-	-	No matching
Acylcarnitines	Hexadecanoylcarnitine	C16	palmitoylcarnitine	C16	Exact*
Acylcarnitines	Hexadecenoylcarnitine	C161	-	-	No matching
Acylcarnitines	Hydroxyhexadecenoylcarnitine	C161OH	-	-	No matching
Acylcarnitines	hexadecadienylcarnitine	C162	-	-	No matching

Table 6.6 (continued)

Metabolite class	Metabolite name (targeted)	Metabolite abbreviation (targeted)	Matched metabolite name (untargeted)	Matched metabolite abbreviation (untargeted)	Precision of matching
Acylcarnitines	Hydroxyhexadecadienylcarnitine	C162OH	-	-	No matching
Acylcarnitines	Hydroxyhexadecanoylcarnitine	C16OH	-	-	No matching
Acylcarnitines	Octadecanoylcarnitine	C18	stearoylcarnitine	C18	Exact*
Acylcarnitines	Octadecenoylcarnitine	C181	-	-	No matching
Acylcarnitines	Hydroxyoctadecenoylcarnitine	C181OH	-	-	No matching
Acylcarnitines	Octadecadienylcarnitine	C182	linoleoylcarnitine	C182	Exact*
Lysophosphatidylcholines	lysoPhosphatidylcholine acyl C14:0	LPC a C14:0	1-myristoyl-GPC (14:0)	lysoPCaC140	Approximate
Lysophosphatidylcholines	lysoPhosphatidylcholine acyl C16:1	LPC a C16:1	1-palmitoleoyl-GPC (16:1)	lysoPCaC16:1	Approximate
Lysophosphatidylcholines	lysoPhosphatidylcholine acyl C17:0	LPC a C17:0	1-margaroyl-GPC (17:0)	lysoPCaC17:0	Approximate
Lysophosphatidylcholines	lysoPhosphatidylcholine acyl C18:0	LPC a C18:0	1-stearoyl-GPC (18:0)	lysoPCaC18:0	Approximate
Lysophosphatidylcholines	lysoPhosphatidylcholine acyl C18:2	LPC a C18:2	1-linoleoyl-GPC (18:2)	lysoPCaC18:2	Approximate
Lysophosphatidylcholines	lysoPhosphatidylcholine acyl C20:4	LPC a C20:4	1-arachidonoyl-GPC (20:4)	lysoPCaC20:4	Approximate
Lysophosphatidylcholines	lysoPhosphatidylcholine acyl C26:1	LPC a C26:1	-	-	No matching
Lysophosphatidylcholines	lysoPhosphatidylcholine acyl C16:0	LPC a C16:0	1-palmitoyl-GPC (16:0)	lysoPCaC16:0	Approximate
Lysophosphatidylcholines	lysoPhosphatidylcholine acyl C18:1	LPC a C18:1	1-oleoyl-GPC (18:1)	lysoPCaC18:1	Approximate
Lysophosphatidylcholines	lysoPhosphatidylcholine acyl C20:3	LPC a C20:3	1-eicosatrienoyl-GPC (20:3)	lysoPCaC20:3	Approximate
Lysophosphatidylcholines	lysoPhosphatidylcholine acyl C24:0	LPC a C24:0	-	-	No matching
Lysophosphatidylcholines	lysoPhosphatidylcholine acyl C26:0	LPC a C26:0	-	-	No matching
Lysophosphatidylcholines	lysoPhosphatidylcholine acyl C28:0	LPC a C28:0	-	-	No matching
Lysophosphatidylcholines	lysoPhosphatidylcholine acyl C28:1	LPC a C28:1	-	-	No matching
Phosphatidylcholines	phosphatidylcholine diacyl C24:0	PC aa C24:0	-	-	No matching
Phosphatidylcholines	phosphatidylcholine diacyl C26:0	PC aa C26:0	-	-	No matching
Phosphatidylcholines	phosphatidylcholine diacyl C28:1	PC aa C28:1	-	-	No matching
Phosphatidylcholines	phosphatidylcholine diacyl C30:0	PC aa C30:0	1-myristoyl-2-palmitoyl-GPC (14:0/16:0)	PC(14:0/16:0)	Approximate
Phosphatidylcholines	phosphatidylcholine diacyl C32:0	PC aa C32:0	1,2-dipalmitoyl-GPC (16:0/16:0)	PC(16:0/16:0)	Approximate
Phosphatidylcholines	phosphatidylcholine diacyl C32:1	PC aa C32:1	1-palmitoyl-2-palmitoleoyl-GPC (16:0/16:1)	PC(16:0/16:1)	Approximate
Phosphatidylcholines	phosphatidylcholine diacyl C32:2	PC aa C32:2	1,2-dipalmitoleoyl-GPC (16:1/16:1)	PC(16:1/16:1)	Approximate
Phosphatidylcholines	phosphatidylcholine diacyl C32:3	PC aa C32:3	-	-	No matching
Phosphatidylcholines	phosphatidylcholine diacyl C34:1	PC aa C34:1	1-palmitoyl-2-oleoyl-GPC (16:0/18:1)	PC (16:0/18:1)	Approximate
Phosphatidylcholines	phosphatidylcholine diacyl C34:2	PC aa C34:2	1-palmitoyl-2-linoleoyl-GPC (16:0/18:2)	PC(16:0/18:2)	Approximate
Phosphatidylcholines	phosphatidylcholine diacyl C34:3	PC aa C34:3	-	-	No matching
Phosphatidylcholines	phosphatidylcholine diacyl C34:4	PC aa C34:4	-	-	No matching
Phosphatidylcholines	phosphatidylcholine diacyl C36:0	PC aa C36:0	-	-	No matching
Phosphatidylcholines	phosphatidylcholine diacyl C36:1	PC aa C36:1	1-stearoyl-2-oleoyl-GPC (18:0/18:1)	PC(18:0/18:1)	Approximate

Table 6.6 (continued)

Metabolite class	Metabolite name (targeted)	Metabolite abbreviation (targeted)	Matched metabolite name (untargeted)	Matched metabolite abbreviation (untargeted)	Precision of matching
Phosphatidylcholines	phosphatidylcholine diacyl C36:2	PC aa C36:2	1-stearoyl-2-linoleoyl-GPC (18:0/18:2)	PC(18:0/18:2)	Approximate
Phosphatidylcholines	phosphatidylcholine diacyl C36:3	PC aa C36:3	1-palmitoyl-2-dihomo-linolenoyl-GPC (16:0/20:3)	PC (16:0/20:3)	Approximate
			1-oleoyl-2-linoleoyl-GPC (18:1/18:2)	PC(18:1/18:2)	Approximate
Phosphatidylcholines	phosphatidylcholine diacyl C36:4	PC aa C36:4	1-palmitoyl-2-arachidonoyl-GPC (16:0/20:4)	PC(16:0/20:4)	Approximate
Phosphatidylcholines	phosphatidylcholine diacyl C36:5	PC aa C36:5	1-palmitoyl-2-eicosapentaenoyl-GPC (16:0/20:5)	PC(16:0/20:5)	Approximate
Phosphatidylcholines	phosphatidylcholine diacyl C36:6	PC aa C36:6	-	-	No matching
Phosphatidylcholines	phosphatidylcholine diacyl C38:0	PC aa C38:0	-	-	No matching
Phosphatidylcholines	phosphatidylcholine diacyl C38:1	PC aa C38:1	-	-	No matching
Phosphatidylcholines	phosphatidylcholine diacyl C38:3	PC aa C38:3	-	-	No matching
Phosphatidylcholines	phosphatidylcholine diacyl C38:4	PC aa C38:4	1-stearoyl-2-arachidonoyl-GPC (18:0/20:4)	PC(18:0/20:4)	Approximate
Phosphatidylcholines	phosphatidylcholine diacyl C38:5	PC aa C38:5	1-palmitoyl-2-docosapentaenoyl-GPC (16:0/22:5n6)	PC(16:0/22:5n6)	Approximate
			1-palmitoyl-2-docosapentaenoyl-GPC (16:0/22:5n3)	PC(16:0/22:5n3)	Approximate
Phosphatidylcholines	phosphatidylcholine diacyl C38:6	PC aa C38:6	-	-	No matching
Phosphatidylcholines	phosphatidylcholine diacyl C40:1	PC aa C40:1	-	-	No matching
Phosphatidylcholines	phosphatidylcholine diacyl C40:2	PC aa C40:2	-	-	No matching
Phosphatidylcholines	phosphatidylcholine diacyl C40:3	PC aa C40:3	-	-	No matching
Phosphatidylcholines	phosphatidylcholine diacyl C40:4	PC aa C40:4	-	-	No matching
Phosphatidylcholines	phosphatidylcholine diacyl C40:5	PC aa C40:5	-	-	No matching
Phosphatidylcholines	phosphatidylcholine diacyl C40:6	PC aa C40:6	-	-	No matching
Phosphatidylcholines	phosphatidylcholine diacyl C42:0	PC aa C42:0	-	-	No matching
Phosphatidylcholines	phosphatidylcholine diacyl C42:1	PC aa C42:1	-	-	No matching
Phosphatidylcholines	phosphatidylcholine diacyl C42:2	PC aa C42:2	-	-	No matching
Phosphatidylcholines	phosphatidylcholine diacyl C42:4	PC aa C42:4	-	-	No matching
Phosphatidylcholines	phosphatidylcholine diacyl C42:5	PC aa C42:5	-	-	No matching
Phosphatidylcholines	phosphatidylcholine diacyl C42:6	PC aa C42:6	-	-	No matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C30:0	PC ae C30:0	-	-	No matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C30:1	PC ae C30:1	-	-	No matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C30:2	PC ae C30:2	-	-	No matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C32:1	PC ae C32:1	-	-	No matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C32:2	PC ae C32:2	-	-	No matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C34:0	PC ae C34:0	-	-	No matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C34:1	PC ae C34:1	1-pentadecanoyl-2-oleoyl-GPC (15:0/18:1)	PC(15:0/18:1)	Approximate
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C34:2	PC ae C34:2	1-pentadecanoyl-2-linoleoyl-GPC (15:0/18:2)	PC(15:0/18:2)	Approximate
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C34:3	PC ae C34:3	-	-	No matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C36:0	PC ae C36:0	-	-	No matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C36:1	PC ae C36:1	1-margaroyl-2-oleoyl-GPC (17:0/18:1)	PC(17:0/18:1)	Approximate

Table 6.6 (continued)

Metabolite class	Metabolite name (targeted)	Metabolite abbreviation (targeted)	Matched metabolite name (untargeted)	Matched metabolite abbreviation (untargeted)	Precision of matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C36:2	PC ae C36:2	1-margaroyl-2-linoleoyl-GPC (17:0/18:2)	PC(17:0/18:2)	Approximate
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C36:3	PC ae C36:3	-	-	No matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C36:4	PC ae C36:4	-	-	No matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C36:5	PC ae C36:5	-	-	No matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C38:0	PC ae C38:0	-	-	No matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C38:1	PC ae C38:1	-	-	No matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C38:2	PC ae C38:2	-	-	No matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C38:3	PC ae C38:3	-	-	No matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C38:4	PC ae C38:4	-	-	No matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C38:5	PC ae C38:5	-	-	No matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C38:6	PC ae C38:6	-	-	No matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C40:1	PC ae C40:1	-	-	No matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C40:2	PC ae C40:2	-	-	No matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C40:3	PC ae C40:3	-	-	No matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C40:4	PC ae C40:4	-	-	No matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C40:5	PC ae C40:5	-	-	No matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C40:6	PC ae C40:6	-	-	No matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C42:0	PC ae C42:0	-	-	No matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C42:1	PC ae C42:1	-	-	No matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C42:2	PC ae C42:2	-	-	No matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C42:3	PC ae C42:3	-	-	No matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C42:4	PC ae C42:4	-	-	No matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C42:5	PC ae C42:5	-	-	No matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C44:3	PC ae C44:3	-	-	No matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C44:5	PC ae C44:5	-	-	No matching
Phosphatidylcholines	phosphatidylcholine acyl-alkyl C44:6	PC ae C44:6	-	-	No matching
Sphingomyelins	shpingomyeline C16:0	SM C16:0	palmitoyl sphingomyelin (d18:1/16:0)	SM(d18:1/16:0)	Approximate
Sphingomyelins	shpingomyeline C16:1	SM C16:1	sphingomyelin (d18:2/16:0, d18:1/16:1)	SM(d18:2/16:0, d18:1/16:1)	Approximate
Sphingomyelins	shpingomyeline C18:0	SM C18:0	stearoyl sphingomyelin (d18:1/18:0)	SM(d18:1/18:0)	Approximate
Sphingomyelins	shpingomyeline C18:1	SM C18:1	sphingomyelin (d18:1/18:1, d18:2/18:0)	SM(d18:1/18:1, d18:2/18:0)	Approximate
Sphingomyelins	shpingomyeline C20:2	SM C20:2	-	-	No matching
Sphingomyelins	shpingomyeline C24:0	SM C24:0	-	-	No matching
Sphingomyelins	shpingomyeline C24:1	SM C24:1	sphingomyelin (d18:1/24:1, d18:2/24:0)	SM(d18:1/24:1, d18:2/24:0)	Approximate
Sphingomyelins	hydroxysphingomyeline C14:1	SM-OH C14:1	sphingomyelin (d18:1/15:0, d16:1/17:0)	SM(d18:1/15:0, d16:1/17:0)	Approximate

Table 6.6 (continued)

Metabolite class	Metabolite name (targeted)	Metabolite abbreviation (targeted)	Matched metabolite name (untargeted)	Matched metabolite abbreviation (untargeted)	Precision of matching
Sphingomyelins	hydroxysphingomyeline C16:1	SM-OH C16:1	-	-	No matching
Sphingomyelins	hydroxysphingomyeline C22:1	SM-OH C22:1	tricosanoyl sphingomyelin (d18:1/23:0)	SM(d18:1/23:0)	Approximate
Sphingomyelins	hydroxysphingomyeline C22:2	SM -OH C22:2	sphingomyelin (d18:2/23:0, d18:1/23:1, d17:1/24:1)	SM(d18:2/23:0, d18:1/23:1, d17:1/24:1)	Approximate
Hexose	hexose	Hexose	mannose	mannose	Approximate
			fructose	fructose	Approximate
			glucose	glucose	Approximate

*The term exact matching refers to matching of the same molecule, although all the molecules in the untargeted platform were expressed in a relative scale, while amino acids and biogenic amines were expressed in an absolute scale in the targeted platform

6.10.3 Discovery of metabolites predicting dairy consumption

Parameters producing the best prediction in 3-fold cross-validation and used to fine-tune the elastic net discovery models are presented in (**Table 6.7**). As described (section 6.9.5), three metabolite sets were used for the development of dairy prediction models: the total set of metabolites from the targeted metabolomics platform of the Fenland study (n=174), the subset of the metabolites, which could be approximately matched with the metabolites from the untargeted platform in the EPIC Norfolk study (main approach; n=82; Table 6.6) and the subset of the overlapping metabolites with the addition of the OCSFAs (n=84). Top metabolite signals from the main (metabolites, socio-demographic and lifestyle factors excluding diet) logistic elastic net regression models for each approach and for each dairy type are presented in **Figure 6.3**. Coefficients estimated in the logistic and linear elastic net regression models and models without penalisation are presented in Appendix 1 (**Tables A.19- A.27**).

Table 6.7 Parameters used in the elastic net models in the discovery set for the different set of metabolites

Dairy products	Model	All metabolites from the targeted platform		Metabolites ovelapping between the two platforms		Metabolites overlapping between the two platforms and odd-chain saturated fatty acids	
		Alpha*	L1 ratio †	Alpha*	L1 ratio †	Alpha*	L1 ratio †
Binary dairy variables							
Total milk	Metabolites	0.01	0.75	0.01	0.45	0.01	0.15
	+ socio-demographic, lifestyle ‡	0.01	0.25	0.01	0.05	0.01	0.05
	+ diet §	0.01	0.25	5	0.05	0.1	0.85
	Socio-demographic, lifestyle only ‡	0.01	0.05	0.01	0.05	10^{-8}	0.25
Total yoghurt	Metabolites	0.01	0.55	0.001	0.85	0.001	0.45
	+ socio-demographic, lifestyle ‡	0.01	0.15	0.01	0.25	0.0001	0.65
	+ diet §	0.0001	0.85	0.0001	0.25	0.001	0.85
	Socio-demographic, lifestyle only ‡	0.0001	0.45	0.0001	0.45	0.001	0.85
Total cheese	Metabolites	0.01	0.85	0.01	0.55	0.01	0.45
	+ socio-demographic, lifestyle ‡	0.01	0.85	0.01	0.75	10^{-8}	0.85
	+ diet §	0.01	0.55	0.001	0.45	0.01	0.85
	Socio-demographic, lifestyle only ‡	10^{-10}	0.05	10^{-10}	0.05	10^{-15}	0.35
Butter	Metabolites	0.01	0.25	0.01	0.45	0.01	0.35
	+ socio-demographic, lifestyle ‡	0.01	0.45	0.01	0.45	0.01	0.25
	+ diet §	0.01	0.75	0.01	0.75	0.01	0.75
	Socio-demographic, lifestyle only ‡	10^{-10}	0.15	10^{-10}	0.15	0.01	0.85
Total dairy products	Metabolites	0.01	0.05	0.01	0.15	0.005	0.05
	+ socio-demographic, lifestyle ‡	0.01	0.85	0.01	0.75	0.01	0.35
	+ diet §	0.01	0.85	0.0001	0.85	0.001	0.25
	Socio-demographic, lifestyle only ‡	0.0001	0.05	0.0001	0.05	10^{-15}	0.05
Continuous dairy variables ‖							
Total milk	Metabolites	0.006	0.5	0.003	0.5	0.003	0.5
	+ socio-demographic, lifestyle ‡	0.006	0.5	0.003	0.5	0.005	0.5
	+ diet §	0.06	0.5	0.06	0.5	0.08	0.5
	Socio-demographic, lifestyle only ‡	0.005	0.5	0.005	0.5	0.09	0.5
Total dairy products	Metabolites	0.005	0.5	0.004	0.5	0.004	0.5
	+ socio-demographic, lifestyle ‡	0.007	0.5	0.004	0.5	0.005	0.5
	+ diet §	0.05	0.5	0.05	0.5	0.05	0.5
	Socio-demographic, lifestyle only ‡	0.005	0.5	0.005	0.5	0.0002	0.5

*Alpha is the constant that multiplies the penalty in penalised regression. For models with the binary dairy variables, values were selected after 3-fold cross-validation using as input in the GridSearch function of Python the values: 10^{-15}, 10^{-10}, 10^{-8}, 10^{-4}, 0.001, 0.01, 1, 5, 10, 20.

†L1 ratio describes whether the elastic net penalisation will be closer to a Ridge (L1 ratio closer to 0) or a LASSO (L1 ratio closer to 1). For models with the binary dairy variables, values were selected after 3-fold cross-validation using as input in the GridSearch function of Python the values: 0.05, 0.15, 0.25, 0.35, 0.45, 0.55, 0.65, 0.75, 0.85

‡Socio-demographic and lifestyle factors include: age, sex, test site, smoking status, physical activity, lipid lowering medication, hormone-replacement therapy, body mass index

§Dietary factors include: total energy intake, fruit, vegetables, cereals, red meat, processed meat, fish, margarine, sweet snacks, sugar-sweetened beverages, caffeinated coffee, tea, alcoholic beverages and dietary supplements

‖For linear models with the continuous dairy variables, parameters were chosen with cross-validation embedded within the elastic net linear regression function.

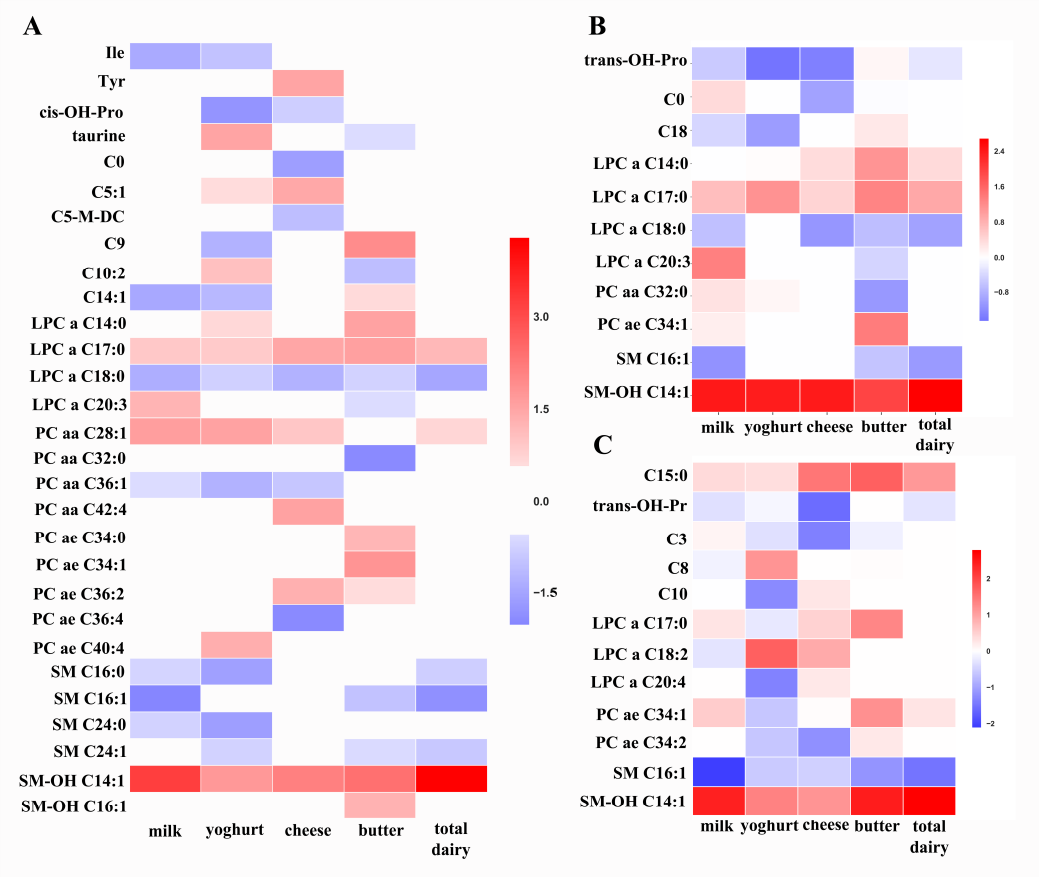


Fig. 6.3 Heatmaps of the top signals (>mean+2*SD of the absolute values of the elastic net coefficients) from the discovery analysis with elastic net logistic regression models with age, sex, test site, smoking status, physical activity, lipid lowering medication, hormone-replacement therapy, body mass index and **A)** 174 metabolites or **B)** 82 metabolites overlapping with the metabolites from the external validation set or **C)** 82 metabolites overlapping with the metabolites from the external validation set and the odd-chain saturated fatty acids.

The metabolite, which contributed the most to prediction of most dairy types was SM-OH C14:1 (Figure 6.3). In all the three analyses, SM C16:1 was the second top signal in total and the top negative signal for total dairy products and for milk apart from the analysis with the 82 overlapping metabolites, in which it was the third for milk. LPC a C17:0 was one of the top metabolite signals positively predicting cheese, butter and total dairy products in the analysis with the 174 metabolites (Figure 6.3A), and butter and total dairy products in the analysis with the 82 metabolites (Figure 6.3B). Cis-OH-Pro was the top (negative and in total) signal for yoghurt in the analysis with the 174 metabolites (Figure 6.3A), but could not be matched with any metabolite from the untargeted platform (Table 6.6). Instead, in the analysis with the 82 metabolites, trans-OH-Pro was the top signal negatively predicting both yoghurt and cheese (Figure 6.3B).

When OCSFAs were added to the discovery analysis (Figure 6.3C), LPC a C17:0 was one of the top metabolite signals for butter only. C15:0 was the second top signal for

cheese and butter and the third top signal for total dairy products (positive associations), whereas C17:0 was not identified as a top signal for any dairy type (Tables A.25, A.26).

6.10.4 Effect of adjusting for diet on discovery results

As mentioned, secondary models also included total energy intake and other dietary factors. After adjustment for dietary factors, the ranking of some metabolites was not substantially affected, whereas for others the difference was larger. The relevant results are shown in **Appendix 1** (**Table A.32** when all 174 metabolites were included and **Table A.33** when only the 82 overlapping metabolites were included in the discovery). In both cases, SM-OH C14:1, which came up as the top metabolite signal for most of dairy types did not change ranking. The ranking of PC aa C18:1, which was one of the top signals for milk and yoghurt in the set of 174 metabolites did not change much either (Table A.32). Cis-5-OH Pro, which was the top metabolite signal for yoghurt in the set of 174 metabolites, dropped by 14 positions after adjustment for diet (Table A.32). The same happened with trans-4-OH Pro which was the second top signal for yoghurt and cheese (drop by 26 positions) in the set of the 82 metabolites (Table A.33). Most of the metabolites for butter and total dairy products did not change ranking in either set of metabolites. Specifically for total dairy products, only SM C16:1 dropped by 161 positions and LPC a C18:0 dropped by 32 positions in both cases.

6.10.5 Internal validation

After fitting logistic regression models with the metabolite scores (total or top metabolites) in the discovery and internal validation sets, the AUCs in the internal validation set were lower than AUCs in the discovery set for all the models and for most of the dairy products (**Table A.28**). Specifically, AUCs of the main models, which included the scores, and socio-demographic and lifestyle (except for diet) factors, were lower by 0.01-0.02 points and ranged from 0.68 for milk to 0.81 for total dairy products (**Figure 6.4**). Likewise for linear prediction models of milk and total dairy products, in most cases R^2 was lower by a range of 0.01-0.04 in the internal validation set compared to the discovery set ranging from 0.06 for the total metabolite score of milk to 0.19 for the total metabolite score of total dairy products (**Table 6.8**). R^2 in the discovery and internal validation set were similar when OCSFAs were added to the discovery models.

The largest predictive value of the metabolite score in addition to socio-demographic and lifestyle variables in the internal validation set was observed for total dairy products (AUC increased from 0.59 to 0.81; Figure 6.4E) and butter (AUC increased from 0.59 to 0.73; Figure 6.4D). For yoghurt the AUC increased from 0.67 to 0.7 (Figure 6.4B).

The loss of information due to the use of only the overlapping metabolites between the targeted and the untargeted platforms did not influence the predictive ability of the

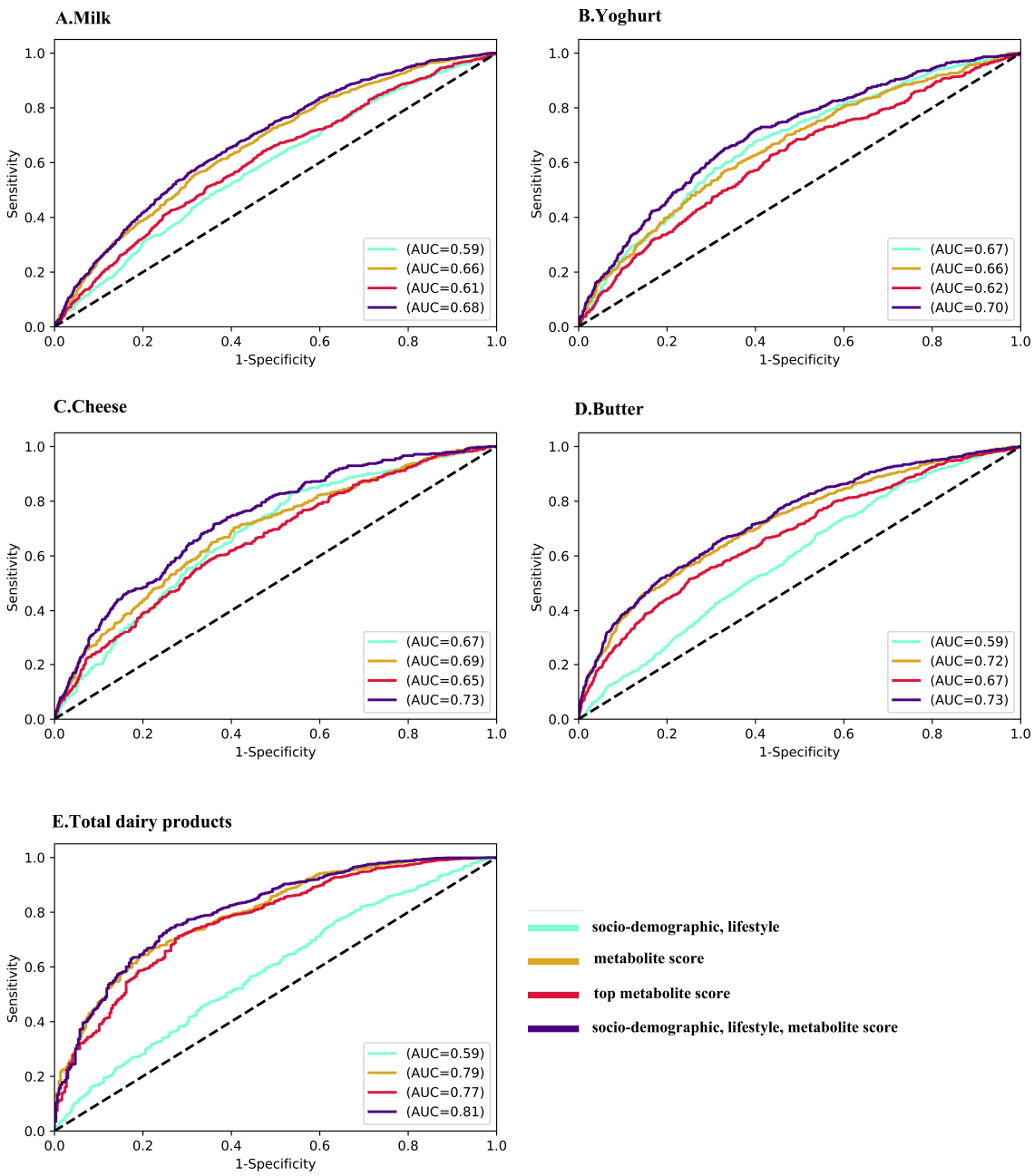


Fig. 6.4 Receiver operating characteristic (ROC) curves and corresponding areas under the curve (AUC) for prediction models of metabolite scores reflecting **A)** milk, **B)** yoghurt, **C)** cheese, **D)** butter and **E)** total dairy consumption in the internal validation set (n=4,246) using the 82 overlapping metabolites. Socio-demographic, lifestyle model: age, sex, test site, smoking status, physical activity, lipid lowering medication, hormone-replacement therapy and body mass index; Top metabolite scores include metabolites with coefficients with absolute values >mean+2*SD

Table 6.8 R^2 of elastic net linear regression models of continuous dairy outcomes in the internal validation set of the Fenland study

Dairy products	Models	Discovery set N=6,035	Internal validation set N=4,246
Total milk	Socio-demographic, lifestyle*	0.01	0.01
	Metabolite score †	0.09	0.06
	Top metabolite score ‡	0.03	0.02
	Socio-demographic, lifestyle, metabolite score*	0.10	0.06
	Socio-demographic, lifestyle, top metabolite score* ‡	0.04	0.03
Total dairy products	Socio-demographic, lifestyle*	0.02	0.03
	Metabolite score †	0.20	0.18
	Top metabolite score ‡	0.14	0.14
	Socio-demographic, lifestyle, metabolite score*	0.21	0.19
	Socio-demographic, lifestyle, top metabolite score* ‡	0.15	0.16

*Socio-demographic and lifestyle factors include age, sex, test site, smoking status, physical activity, lipid lowering medication, hormone-replacement therapy and body mass index
†Metabolite score created from the 82 metabolites overlapping with the metabolites from the external validation set
‡Top metabolite scores were derived from the top mean+2*SD of the absolute values of the elastic net coefficients from the main models, which included age, sex, test site, smoking status, physical activity, lipid lowering medication, hormone-replacement therapy, body mass index and the metabolites

models, as the AUCs for the models with the metabolite score, the socio-demographic and the lifestyle factors in the internal validation set were similar to those derived from the total set of 174 metabolites (**Figure A.1**).

6.10.6 External validation

As expected, AUCs in the external validation set were overall lower than those observed in the internal validation set (**Table 6.9**). Among the metabolite scores derived from the different approaches, the metabolite scores with the highest AUCs overall with also the most significant contributions to the multivariable AUCs, were those derived from the top signals when OCSFAs were also included in the models for milk (AUC=0.65, p=0.005), butter (AUC=0.66, p=0.001) and total dairy products (AUC=0.77, p=0.001) (Table 6.9). Although the contribution of the total and top metabolite scores for milk and the top metabolite score for butter from the approach without OCSFAs were significant, the increase in AUCs from the base model (socio-demographic and lifestyle factors) was of small magnitude, by 0.01 point. For total dairy products, contribution to AUCs were significant for all the approaches apart from the one corresponding to the total metabolite score derived from the 82 overlapping metabolites.

Metabolite scores for yoghurt and cheese were not validated in the external validation set with non-significant contribution to multivariable AUCs of 0.69 and 0.66 respectively (p>0.05) when the 82 overlapping metabolites were used. The contribution of the metabolite scores were not significant even when OCSFAs were included in the scores and produced increases in AUCs up to 0.72 for yoghurt and 0.7 for cheese scores (Table 6.9). Similar results were observed when the AUCs for all the models were estimated in the OCSFA subsample of the external validation set (**Table A.29**).

Table 6.9 Areas under the curve (AUC) for logistic regression models of the dairy outcomes against the metabolite scores adjusted for socio-demographic and lifestyle factors* in the internal and external validation sets

Dairy outcome	Metabolite set ‡	Number of metabolites	Internal validation set			External validation set			
				Metabolite score †			Metabolite score †		
				No	Yes		No	Yes	
			N	AUC	AUC	N	AUC	AUC	*p* §
Milk	overlapping ‖	82	2,758	0.59	0.68	900	0.59	0.60	0.03
	top signals from overlapping #	3		0.59	0.63	900	0.59	0.60	0.02
	overlapping, OCSFAs	84		0.59	0.72	362	0.61	0.62	0.14
	top signals from overlapping, OCSFAs**	2		0.59	0.66	362	0.61	0.65	0.005
Yoghurt	overlapping ‖	82	2,089	0.67	0.70	875	0.69	0.69	0.98
	top signals from overlapping #	2		0.67	0.68	875	0.69	0.70	0.37
	overlapping, OCSFAs	84		0.67	0.75	391	0.72	0.72	0.94
	top signals from overlapping, OCSFAs**	5		0.67	0.68	391	0.72	0.72	0.98
Cheese	overlapping ‖	82	1,574	0.67	0.73	475	0.65	0.66	0.23
	top signals from overlapping #	4		0.67	0.71	475	0.65	0.66	0.18
	overlapping, OCSFAs	84		0.67	0.79	194	0.70	0.70	0.94
	top signals from overlapping, OCSFAs**	3		0.67	0.72	194	0.70	0.70	0.36
Butter	overlapping ‖	82	3,066	0.59	0.73	1,193	0.61	0.62	0.10
	top signals from overlapping #	5		0.59	0.69	1,193	0.61	0.62	0.03
	overlapping, OCSFAs	84		0.59	0.76	515	0.62	0.64	0.05
	top signals from overlapping, OCSFAs**	5		0.59	0.72	515	0.62	0.66	0.001
Total dairy products	overlapping ‖	82	1,969	0.59	0.81	687	0.59	0.64	0.07
	top signals from overlapping #	4		0.59	0.78	687	0.59	0.68	0.003
	overlapping, OCSFAs	84		0.59	0.84	282	0.61	0.74	0.01
	top signals from overlapping, OCSFAs**	3		0.59	0.80	282	0.61	0.77	0.001

*Socio-demographic and lifestyle factors include age, sex, test site, smoking status, physical activity, lipid lowering medication, hormone-replacement therapy and body mass index
†"Metabolite score-No" includes the socio-demographic and lifestyle factors; "Metabolite score-Yes" includes the socio-demographic and lifestyle factors as well as the metabolite score
‡The metabolite set used in the discovery analysis to derive the metabolite scores
§*p* values of the coefficients of the metabolite score
‖Metabolites of the discovery (targeted) set approximately matched with metabolites from the external validation (untargeted) set
#Top signals defined as metabolites with absolute values of coefficients >mean+2*SD
**Top signals derived from the discovery models which included the overlapping metabolites and OCSFAs
Abbreviations: OCSFA: Odd-chain saturated fatty acid

6.10.7 Predictive value of the metabolite scores in addition to C15:0, C17:0 and trans-16:1n-7

Table 6.10 shows the change in AUCs when we added fatty acids to the base model (No, No -> No, Yes), the metabolite score (total or top) to the base model (No, No -> Yes, No), fatty acids to the model with the metabolite score (Yes, No -> Yes, Yes) and the metabolite score (total or top) to the model with fatty acids (No, Yes -> Yes, Yes). When the metabolite score was added to the models with the fatty acids in the internal validation set, socio-demographic and lifestyle factors, the AUC was increased by 0.02 for yoghurt and cheese and 0.06 for milk and total dairy products. Lower increases in the AUCs were observed when the three fatty acids were added to the models with the metabolite score, socio-demographic and lifestyle factors, ranging from 0 for milk and total dairy products to 0.01 for the other dairy types. Multivariable AUCs were higher by a range of 0.02 for yoghurt and cheese to 0.05 for milk, when the top metabolite score was used in place of the

total metabolite score. Results from the analyses including the total set of 174 metabolites were similar (**Table A.30**). The likelihood ratio tests were highly significant apart from the case where the three fatty acids were added to models with the metabolite score and the socio-demographic and lifestyle factors, where the test was highly significant only for butter (p=5.6 x 10^{-5}; **Table 6.11**).

Overall, coefficients of the metabolite scores were highly significant both in the models with and without the fatty acids (**Table 6.12**). In the models with the metabolite scores, C15:0 was significant for milk and yoghurt only when the top metabolite score was included, but for cheese, butter and total dairy products, it was significant both for the total and the top metabolite scores (Table 6.12). In the same models, C17:0 was significant only for total dairy products (continuous) when the top metabolite score was included, while trans-16:1n-7 was significant for yoghurt (total and top metabolite score), butter (top metabolite score) and total dairy products (binary; total and top metabolite score; Table 6.12). These results were not replicated when the same analysis was performed in the external validation set with no obvious additive value of the metabolite scores to dairy prediction models including OCSFAs (**Table A.31**).

Table 6.10 Areas under the curve statistics of the logistic models* with or without the fatty acids C15:0, C17:0, trans-16:1n-7 and the metabolite score (82 metabolites †) for total and types of dairy products in the internal validation set of the Fenland study

Metabolite score				**Top metabolite score ‡**			
Milk		Fatty acids		**Milk**		Fatty acids	
		No	Yes			No	Yes
Metabolite score	No	0.59	0.62	Metabolite score	No	0.59	0.62
	Yes	0.68	0.68		Yes	0.63	0.64
Yoghurt		Fatty acids		**Yoghurt**		Fatty acids	
		No	Yes			No	Yes
Metabolite score	No	0.67	0.69	Metabolite score	No	0.67	0.69
	Yes	0.70	0.71		Yes	0.68	0.69
Cheese		Fatty acids		**Cheese**		Fatty acids	
		No	Yes			No	Yes
Metabolite score	No	0.67	0.72	Metabolite score	No	0.67	0.72
	Yes	0.73	0.74		Yes	0.71	0.73
Butter		Fatty acids		**Butter**		Fatty acids	
		No	Yes			No	Yes
Metabolite score	No	0.59	0.70	Metabolite score	No	0.59	0.70
	Yes	0.73	0.74		Yes	0.69	0.72
Total dairy products		Fatty acids		**Total dairy products**		Fatty acids	
		No	Yes			No	Yes
Metabolite score	No	0.59	0.75	Metabolite score	No	0.59	0.75
	Yes	0.81	0.81		Yes	0.78	0.80

*All models include age, sex, test site, smoking status, physical activity, lipid lowering medication, hormone-replacement therapy and body mass index
†Metabolites approximately matched between the discovery (targeted) set and the external validation (untargeted) set
‡The top metabolite score includes only the top metabolite signals defined as the metabolites with absolute values of coefficients >mean+2*SD

Table 6.11 Likelihood ratio tests for the models with or without the fatty acids* and the metabolite score in the fatty acid set (internal validation set, n=4,246) of the Fenland study

		Milk		Yoghurt		Cheese		Butter		Total dairy products	
Baseline model	**Predictor(s) added**	**LR**	***p***	**LR**	***p***	**LR**	***p***	**LR**	***p***	**LR**	***p***
Socio-demographic, lifestyle †	Fatty acids*	60.71	4.2×10^{-13}	26.62	7.1×10^{-6}	62.78	1.5×10^{-13}	263.14	9.4×10^{-57}	198.56	8.6×10^{-43}
Socio-demographic, lifestyle †	Metabolite score	219.05	1.5×10^{-49}	61.60	4.2×10^{-15}	89.89	2.5×10^{-21}	362.59	7.7×10^{-81}	331.08	5.6×10^{-74}
Socio-demographic, lifestyle †	Top metabolite score ‡	80.30	3.2×10^{-19}	21.67	3.2×10^{-6}	53.18	3.1×10^{-13}	211.86	5.4×10^{-48}	265.71	9.8×10^{-60}
Socio-demographic, lifestyle, fatty acids* †	Metabolite score	160.79	7.6×10^{-37}	42.67	6.5×10^{-11}	37.47	9.3×10^{-10}	121.78	2.6×10^{-28}	145.50	1.7×10^{-33}
Socio-demographic, lifestyle, fatty acids* †	Top metabolite score ‡	46.75	8.1×10^{-12}	8.78	0.003	19.04	1.3×10^{-5}	54.42	1.6×10^{-13}	103.34	2.8×10^{-24}
Socio-demographic, lifestyle, metabolite score †	Fatty acids*	2.45	0.485	7.70	0.053	10.36	0.016	22.33	5.6×10^{-5}	12.98	0.005
Socio-demographic, lifestyle, Top metabolite score † ‡	Fatty acids*	27.16	5.5×10^{-6}	13.74	0.003	28.64	2.7×10^{-6}	105.71	9.2×10^{-23}	36.19	6.8×10^{-8}

*Fatty acids: C15:0, C17:0, trans-16:1n-7

†Socio-demographic and lifestyle factors: age, sex, test site, smoking status, physical activity, lipid lowering medication, hormone-replacement therapy and body mass index

‡The top metabolite score includes only the top metabolite signals defined as the metabolites with absolute values of coefficients >mean+2*SD

Abbreviations: LR: Likelihood ratio

Table 6.12 Beta coefficients (b), 95% CIs and *p*-values of the metabolite score and the three fatty acids from logistic prediction models of dairy consumption.*

	Metabolite score			C15:0			C17:0			trans-16:1n-7		
Dairy products	**b**	**95% CI**	***p***	**b**	**95% CI**	***p***	**b**	**95% CI**	***p***	**b**	**95% CI**	***p***
Binary outcomes												
Milk												
Fatty acids				0.30	0.2, 0.41	3.2×10^{-8}	0.03	-0.07, 0.13	0.52	0.01	-0.08, 0.1	0.86
Metabolite score	0.63	0.54, 0.72	1.4×10^{-43}									
Fatty acids + metabolite score	0.61	0.51, 0.71	2.1×10^{-33}	0.01	-0.11, 0.13	0.87	0.06	-0.05, 0.16	0.30	0.02	-0.07, 0.11	0.63
Fatty acids + top metabolite score †	0.30	0.21, 0.39	8.2×10^{-11}	0.19	0.08, 0.31	0.001	0.06	-0.04, 0.16	0.25	0.01	-0.08, 0.1	0.86
Yoghurt												
Fatty acids				0.27	0.12, 0.41	0.0002	0.10	-0.02, 0.23	0.11	-0.16	-0.29, -0.04	0.01
Metabolite score	0.46	0.34, 0.57	2.7×10^{-14}									
Fatty acids + metabolite score	0.44	0.31, 0.58	1.5×10^{-10}	0.05	-0.11, 0.21	0.53	0.10	-0.03, 0.23	0.13	-0.15	-0.28, -0.03	0.02
Fatty acids + top metabolite score †	0.20	0.07, 0.33	0.003	0.18	0.02, 0.33	0.02	0.10	-0.03, 0.22	0.14	-0.17	-0.29, -0.04	0.01
Cheese												
Fatty acids				0.63	0.46, 0.8	1.1×10^{-12}	-0.21	-0.39, -0.03	0.02	-0.06	-0.2, 0.08	0.39
Metabolite score	0.64	0.5, 0.78	3.6×10^{-19}									
Fatty acids + metabolite score	0.51	0.34, 0.68	2.6×10^{-9}	0.32	0.12, 0.52	0.001	-0.15	-0.33,0.02	0.09	-0.07	-0.21, 0.08	0.36
Fatty acids + top metabolite score †	0.34	0.19, 0.5	1.5×10^{-5}	0.49	0.3, 0.67	2.1×10^{-7}	-0.18	-0.36, 0.00	0.05	-0.09	-0.23, 0.06	0.23
Butter												
Fatty acids				0.70	0.58, 0.82	3×10^{-30}	-0.04	-0.14, 0.06	0.42	0.12	0.02, 0.22	0.02
Metabolite score	0.91	0.81, 1.01	4.4×10^{-66}									
Fatty acids + metabolite score	0.71	0.58, 0.84	2.1×10^{-26}	0.27	0.13, 0.41	0.0002	-0.03	-0.14, 0.07	0.52	0.07	-0.03, 0.17	0.17
Fatty acids + top metabolite score †	0.42	0.3, 0.53	1.8×10^{-12}	0.50	0.37, 0.63	1.1×10^{-13}	-0.03	-0.14, 0.07	0.55	0.12	0.02, 0.22	0.02
Total dairy products												
Fatty acids				1.12	0.92, 1.32	8.2×10^{-28}	-0.12	-0.25, 0.02	0.10	-0.06	-0.22, 0.1	0.46
Metabolite score	1.31	1.14, 1.48	2.6×10^{-51}									
Fatty acids + metabolite score	1.16	0.95, 1.37	1×10^{-27}	0.40	0.17, 0.63	0.001	-0.06	-0.21, 0.09	0.42	-0.18	-0.35, -0.02	0.03
Fatty acids + top metabolite score †	0.92	0.73, 1.11	5.8×10^{-22}	0.64	0.43, 0.86	4.8×10^{-9}	-0.12	-0.24, 0.01	0.08	-0.21	-0.37, -0.04	0.01

Table 6.12 (continued)

Dairy products	Metabolite score b	95% CI	*p*	15:0 b	95% CI	*p*	17:0 b	95% CI	*p*	trans-16:1n7 b	95% CI	*p*
Continuous outcomes												
Total milk												
Fatty acids				0.04	0.03, 0.06	2.1×10^{-7}	0.00	-0.01, 0.02	0.67	0.00	-0.01, 0.02	0.77
Metabolite score	0.10	0.08, 0.11	4.2×10^{-58}									
Fatty acids + metabolite score	0.10	0.08, 0.11	4.1×10^{-47}	-0.004	-0.02, 0.01	0.6	0.01	-0.01, 0.02	0.44	0.00	-0.01, 0.01	0.96
Fatty acids + top metabolite score †	0.05	0.04, 0.07	1.5×10^{-15}	0.02	0.003, 0.04	0.02	0.01	-0.01, 0.02	0.35	0.001	-0.01, 0.01	0.87
Total dairy products												
Fatty acids				0.12	0.1, 0.13	2.2×10^{-57}	-0.01	-0.02, 0.01	0.25	0.00	-0.01, 0.02	0.6
Metabolite score	0.15	0.14, 0.16	3×10^{-176}									
Fatty acids + metabolite score	0.14	0.13, 0.16	1.7×10^{-89}	0.02	0.01, 0.04	0.004	-0.01	-0.02, 0.01	0.31	-0.01	-0.02, 0.005	0.24
Fatty acids + top metabolite score †	0.11	0.1, 0.13	2.1×10^{-59}	0.06	0.04, 0.07	6×10^{-13}	-0.02	-0.03, -0.002	0.02	-0.01	-0.02, 0.001	0.07

*Age, sex, test site, smoking status, physical activity, lipid lowering medication, hormone-replacement therapy and body mass index are included in the models as separate covariates

†The top metabolite score includes only the top metabolite signals defined as the metabolites with absolute values of coefficients >mean+2*SD

6.11 Discussion

6.11.1 Summary of results

In this study, we developed metabolite scores for total and types of dairy products with multivariable AUCs ranging from 0.72 for yoghurt to 0.83 for total dairy products in the discovery set and from 0.68 for milk to 0.81 for total dairy products in the internal validation set. We did not identify metabolites discriminating dairy types from each other, but we consider the metabolite scores to specifically predict individual dairy types. The metabolite SM-OH C14:1 was identified as top signal across different dairy types and across different set of metabolites or analyses. SM C16:1 was one of the top signals for milk and total dairy products across the different sets of metabolites, but adjustment for other dietary factors affected its ranking substantially. LPC a C17:0 was one of the top signals for cheese, butter and total dairy products (most consistently for butter). In secondary analyses, C15:0 was one of the top signals for cheese, butter and total dairy products when OCSFAs were added to the models, whereas C17:0 did not reach the top signals for any dairy type.

In the external validation set, metabolite scores for milk, butter and total dairy products contributed significantly to the prediction of dairy consumption. The contribution of scores to yoghurt and cheese prediction models was not significant. The metabolite scores had predictive value for dairy consumption in addition to the fatty acids for all the dairy types. The fatty acids had predictive value in addition to the metabolite scores only for butter in the internal validation set.

6.11.2 Findings in context of previous evidence

Metabolomic profiling is useful to identify potential novel biomarkers of dairy consumption and potentially address the limitations of OCSFAs and trans-16:1n-7 as candidate dairy biomarkers (section 6.4). In our study, we showed that metabolite scores for total and types of dairy products had predictive value for dairy consumption in addition to OCSFAs and trans-16:1n-7, while fatty acids had predictive value in addition to the metabolite scores only for butter consumption. Of the dairy products, butter has the highest proportion of saturated fat, so prediction of butter consumption with OCSFAs might be stronger than of other dairy types. This was the case especially for C15:0, which was one of the top signals for the prediction of butter consumption in the discovery analysis, which included OCSFAs. On the contrary, C17:0 was not a top metabolite signal for any dairy type. Considering that our method accounted for multicollinearity between metabolites and that C17:0 has been proposed to be also produced endogenously by alternative pathways[208], it might be assumed that failure to identify C17:0 as a top metabolite for dairy prediction was due to the inclusion of other metabolites that were highly correlated with C17:0 and contributed

more to the prediction of dairy consumption than C17:0. One of these metabolites might be LPC a C17:0 as one of the top signals for cheese, total dairy products and more consistently across different analytical methods for butter.

Concerning previous studies on the use of metabolomics for the discovery of potential dairy biomarkers, the study by Pallister et al was the most comparable with our study in terms of the biological sample (blood), the study design (cross-sectional), the dairy products of interest (milk, butter, total dairy products) and the set of metabolites analysed (Biocrates metabolite set)[261]. In this study, they also identified SM-OH C14:1 as one of the top metabolites for milk, butter and total dairy products along with PC aa C28:1 for butter and total dairy products in a UK population and they were able to replicate these findings in an Estonian and a German population[261]. They did not report any results for yoghurt or cheese consumption. The replication of the top signal for the part of the analysis that was common between our study and the study by Pallister et al provides us with an additional external validation of this part of our results.

Comparison of our results with results from other studies on discovery of dairy biomarkers with metabolomics is limited due to methodological differences including different biological samples (urine[253, 256–258, 260, 262], faeces[257]), application of H-NMR[254], and use of a different set of blood metabolites[227, 229, 255, 259].

6.11.3 Biological interpretation

Interpretation of the top metabolite signals of dairy consumption depends on the discovery set of metabolites used. The metabolites in this study included 21 amino acids, 12 biogenic amines, 40 acylcarnitines, 88 PCs, 11 SMs and 1 hexose. Most of the metabolites we identified as top signals for dairy prediction were lipids. When all the 174 metabolites were included in the discovery analysis, the non-lipid top (as we defined them) metabolite signals were interleucine for milk, cis-4-OH Pro and taurine for yoghurt and Trp for cheese. Food composition information from the FooDB database (http://foodb.ca/) indicates that these compounds are not specific to these dairy types. For example, taurine is included in meat, fish, legumes among other foods (http://foodb.ca/compounds/FDB003191) and cis-4-OH Pro is included in cheese, processed meat, garlic and several vegetables (http://foodb.ca/compounds/FDB013511). We also observed the non-specificity of these metabolites to dairy consumption in our data from the drop in the ranking of these metabolites after adjustment for dietary factors.

Thus, the set of metabolites that we have used seems to reflect predominantly the fat in dairy products. As expected, some of the top signals might not be specific to dairy types such as SM C16:1, which was not a top metabolite after adjustment for dietary factors. However, the top metabolite SM-OH C14:1 was more consistent across dairy products and different analyses. This molecule represents the isobaric molecules of SM(d18:1/ 15:0) and SM(d16:1/ 17:0), which contain OCSFAs. Information on these SMs from metabolite

databases is limited. SMs and PCs as top signals for prediction of dairy consumption might partly reflect the fat in the milk fat globule membrane and partly the free fat. It has been reported that 60-70% of the polar lipids in milk i.e. phospholipids and SMs are located in the milk fat globule membrane[34]. The majority of the phospholipids found are PCs (35%), while SMs comprise 25% of the polar lipids of the membrane[34]. Thus, with the metabolomics platform we used, we covered approximately 60% of these lipids. In addition, with the inclusion of PCs and SMs we covered most of the saturated fat of the membrane, as the majority of both lipid classes are saturated, while other lipid classes of the membrane such as phophatidylinositols or phophatidyethanolamines are mainly unsaturated[34].

Milk, yoghurt and cheese contain higher amounts of the milk fat globule membrane, while butter contains lower amounts due to homogenisation[34]. However, considering the complicated metabolic pathways and processing that the dairy lipid molecules go through after their uptake in the organism, it is not possible to disentangle the association we observed by identifying the specific origin of the blood lipids. For example, we observed SM-OH C14:1 as the most consistent top signal, which contains the OCSFAs. It is not possible with this observation alone to conclude the specific source molecules and the extent that each of them contributes to blood SM-OH C14:1. Considering the different forms, in which lipids are contained in dairy products (section 1.4), potential sources are the uptake and processing of the free OCSFAs; the metabolism of dairy triglycerides; the dairy SMs, which might be metabolised to produce free OCSFAs and then form SMs; dairy phospholipids that contain these fatty acids or endogenous production of these fatty acids, especially C17:0, from alternative pathways as previously described[233, 244, 245].

6.11.4 Strengths and limitations

The key strength of our study is the development of metabolite scores predictive of dairy consumption instead of the single metabolite approach, thus increasing the probability that the metabolite scores are more specific to dairy products, which have a complex food matrix and share many nutrients with other foods[207]. While replication of results of a nutritional metabolomics study in an independent study is not very common due to the unavailability of resources and is identified as a limitation of such studies[288], we were able not only to perform internal validation, but also external validation by using an independent cohort despite some discrepancies due to the study design. We also applied a machine learning method (elastic net regression), which was shown to be a flexible and appropriate approach for our purpose leading to lower prediction errors as reported from previous metabolomics studies[280], but has not previously been applied in other nutritional metabolomics studies to our knowledge.

This study has also limitations. Although our primary aim was the prediction of dairy consumption, it is also of interest to unravel the unconfounded and causal associations

between dairy consumption and metabolites. The observational nature and the cross-sectional design of our study imposed difficulties in identifying such associations, but it would still be important to generate hypotheses to be examined in RCTs, which are considered the gold standard. We have used a subjective method of dietary assessment to investigate potential biomarkers of dairy consumption in order to use them for the objective assessment of dairy consumption, which might introduce the measurement errors in self-reported dietary assessment (section 6.1). In addition, following about a decade of the use of metabolomics in nutrition, the need for studies which characterise dose response relationships between dietary intake and biomarkers has been emphasised, since lack of assessment of the dose response relationship is common in most nutritional metabolomics studies[288]. However, since the availability of exploratory metabolomics studies for the identification of potential blood biomarkers of dairy consumption is limited, we consider our study informative for hypothesis generation for future assessment of the dose-response relationship in feeding trials. Finally, the set of metabolites we used was limited, missing proteins (only targeted amino acids) and lactose metabolites. Nevertheless, with the metabolite scores we generated, we could approximate the associations between an important part of the dairy lipid profile with part of the human lipid profile, which could be an important predictor of cardio-metabolic disease.

6.11.5 Conclusion

To our knowledge, this is the first study to develop metabolite scores predictive of both total and types of dairy consumption, assess their validity (internally and externally), and show that such scores have predictive value for dairy consumption in addition to OCSFAs and trans-16:1n-7 as candidate dairy biomarkers. External validation was successful for milk, butter and total dairy products, but not for yoghurt and cheese. Thus, we were able to create metabolite scores specific for discriminating between consumers and non-consumers of milk, butter and total dairy products, but not yoghurt or cheese.

Future RCTs can further advance our findings and examine the hypotheses we generated. Studies with different sets of metabolites e.g. untargeted metabolomics and metabolites related to gut microbiome, application of more machine learning methods with high flexibility and prediction accuracy, but also the combination of metabolomics, transcriptomics, epigenomics, and genomics could further contribute to the discovery of biomarkers and a greater understanding of biological mechanisms[265]. Metabolomics might be a useful tool for the discovery of novel nutritional biomarkers, which could potentially elucidate metabolic pathways relating food consumption to disease. The novelty of our study can contribute new evidence to the area of dairy biomarkers with the use of metabolomics and constitute a useful approach for related future research.

Chapter 7

Biomarkers of dairy consumption: Part II. Associations with type 2 diabetes

Summary

Background and aims: Meta-analyses of prospective cohort studies indicated heterogeneity in associations between different types of dairy products and type 2 diabetes (T2D) risk. The majority of the studies use self-reported dairy consumption, with accompanying measurement error. To improve precision in dietary assessment and hence improve epidemiological understanding of the links between dairy products and T2D, we aimed to use objectively measured metabolite scores related with dairy consumption to investigate associations with T2D risk.

Methods: We evaluated 1,440 participants of the diabetes case-cohort set nested within the EPIC Norfolk study (641 T2D cases and 799 non-cases). Prentice-weighted Cox proportional hazard models were used to investigate associations between standardised metabolite scores generated in an independent cohort and T2D risk adjusting for socio-demographic, lifestyle and clinical factors.

Results: During 16,360 person-years of follow-up, 641 T2D cases were identified. Significant inverse associations with T2D risk were observed for the metabolite score for milk, butter and total dairy products, which significantly predicted dairy consumption in this cohort. For example, total metabolite scores for total dairy products, milk and butter were associated with a 24% [HR=0.76 (0.69, 0.83)], 11% [HR=0.89 (0.82, 0.97)] and 43% [HR=0.57 (0.52, 0.62)] lower risk of T2D in multivariable models (without adjustment for BMI) respectively. Metabolite scores for yoghurt and cheese were not successfully externally validated.

Conclusions: We demonstrated that metabolite scores, which could significantly predict milk, butter and total dairy consumption, were associated with lower T2D risk. Metabolite scores for yoghurt and cheese, which were not externally validated, were associated with a higher or no risk. These findings indicate that our novel approach

of the development of metabolite scores for the prediction of dairy consumption using metabolomics is also a promising approach for the identification of objective markers of dairy consumption, specific to individual dairy types and can open up possibilities for understanding the pathways to disease aetiology. Further research is necessary for the discovery of scores, which better predict certain types of dairy products such as yoghurt and cheese, which are of great interest for their associations with cardio-metabolic disease.

What is already known

- Different types of dairy products have been associated differently with type 2 diabetes (T2D) risk in meta-analyses of prospective cohort studies
- The majority of the studies on associations between dairy consumption and T2D risk use self-reported methods of dietary assessment, which entail measurement error

What this research adds

- Metabolite scores, which significantly predicted total dairy, milk and butter consumption were significantly associated with a lower risk of T2D by 24%, 11% and 43% respectively in multivariable adjusted models.
- Metabolomics opens up the possibility for the elucidation of potential pathways linking dairy consumption to T2D risk.

Publication

Trichia E, Imamura F, Koulman A, Brage S, Griffin SJ, Langenberg C, Khaw KT, Wareham NJ, Forouhi N G. Development and validation of dairy prediction models using metabolomics in two UK cohort and the associations of derived metabolite scores with type 2 diabetes risk (Manuscript under preparation)

7.1 Previous evidence

Evidence from meta-analyses of prospective cohort studies indicates a heterogeneity of associations between different types of dairy products and type 2 diabetes (T2D) risk (Chapter 1). For example, no significant associations were reported between full-fat milk[12, 18, 19, 21], cream[12], fermented[19] or high-fat[12, 16, 18, 19, 21] dairy products and T2D risk. On the other hand, significantly inverse associations were consistently reported overall between yoghurt[12, 16, 18, 19, 21] or butter[14] and T2D risk. Finally, mixed associations were reported between milk (low-fat[12, 18, 19] or total[12, 18, 19, 22]), cheese[12, 18, 19], or dairy products (low-fat[12, 16, 18, 19, 21] or total[8, 12, 16, 18, 19, 21, 22]) and T2D risk.

To a certain extent, this heterogeneity of associations might be attributed to true variation due to different characteristics of different dairy types, which raises questions about what contributed mainly to the observed associations with T2D and through which pathways if causal effects were present. In addition to true variation, the heterogeneity of associations might also be partly a result of differential measurement error from the self-reported methods of dietary assessment. The use of nutritional biomarkers has the potential to contribute towards a better understanding of the relevant pathways underpinning such associations by reflecting specific sets of characteristics of dairy products, but also to overcome the limitations of subjective methods of dietary assessment.

The fatty acids C15:0, C17:0 and trans-16:1n-7 are the candidate biomarkers of dairy consumption proposed so far. By investigating associations of these fatty acids with T2D or related endpoints, we may obtain a deeper insight into pathways linking dairy consumption (specifically dairy fat in this case) to T2D. Results from individual prospective cohort studies on the association of the three fatty acids with T2D risk were mixed[218, 221, 223, 226, 228, 232, 289–291] indicating inverse or null associations. The sources of this heterogeneity of associations are not clear, but they are probably related to methodological differences between the studies. Results of a pooling project among prospective cohort studies within the Fatty Acids and Outcomes Research (FORCE) consortium, which used a harmonised protocol, showed inverse associations per 10^{th} to 90^{th} percentile range of C15:0 (n=16, RR=0.80, 95% CI:0.73, 0.87), C17:0 (n=13, RR=0.65, 95% CI:0.59, 0.72) and trans-16:1n-7 (n=8, RR=0.82, 95% CI: 0.70, 0.96) after adjustment of various covariates including adiposity measures[292]. No study has reported an increase in risk with consumption or circulating levels of these fatty acids. In summary, there is some evidence indicating a decrease in risk of T2D with consumption or circulating levels of odd-chain fatty acids (OCSFAs) and trans-16:1n-7, but due to the heterogeneity of results, further investigation is needed. There is therefore an interest in the discovery and use of more potential biomarkers of dairy consumption.

7.2 Study aims

The primary aim of this study was to assess the associations of metabolite scores, which were generated from metabolomics data and reflect consumption of different types of dairy products, with T2D risk. A secondary aim of this study was to assess the associations of metabolite scores generated from both metabolomics and OCSFAs, with T2D risk.

7.3 Methods

7.3.1 Study design and population

We evaluated 1,440 adults with available metabolomics profiles (641 incident T2D cases and 799 non-cases, 16,360 person-years of follow-up) from the incident diabetes case-cohort study nested within the EPIC Norfolk study, UK. This sample was derived after exclusion of participants who were missing dietary data (n=44), men with energy intake less than 800 kcal or more than 4,000 kcal or women with energy intake less than 500 kcal or more than 3,500 kcal (n=37) and participants with more than 50% of the metabolites missing (n=2). In a secondary analysis, where we generated metabolite scores using OCSFAs in an independent cohort, participants with no fatty acid measurements were further excluded (n=848) leaving 592 participants (356 incident T2D cases and 236 non-cases). Details about the case-cohort study were reported in section 6.9.1 and details about the EPIC Norfolk study were reported in section 5.3. An overview of the study design was presented in Figure 6.2.

7.3.2 Development of metabolite scores and metabolomics analyses

Scores to predict consumption of each dairy product were based on algorithms developed in the Fenland study. Metabolomics profiling and validation of the metabolite scores in the EPIC Norfolk study were presented in detail in sections 6.9.2 and 6.9.5 respectively. Briefly, to assess success of external validation we used the Area under the curve (AUC) statistics and the significance of the contribution of metabolite scores to logistic prediction models for dairy consumption. Metabolite scores for yoghurt and cheese failed to validate externally as shown in Table 6.9, which means that they do not reflect consumption of these dairy foods in the EPIC Norfolk diabetes case-cohort study.

7.3.3 Diabetes case ascertainment

Cases of incident T2D were ascertained until 31 July 2006. Two sources of ascertainment were used combining information from self-report and participant health records. Within the EPIC Norfolk cohort, information was collected prospectively on incident self-reported

diagnosed diabetes or use of diabetes medications. Data were also collected from record linkage with general practices or local hospital records, hospital admissions, and diabetes codes from the Office of National Statistics mortality data. To minimise misclassification of case status, only verified cases were included. Thus, self-report of physician diagnosis alone was not considered as a confirmed case unless verified by another source internal (e.g. medicine use) or external (record linkage) to the study.

7.3.4 Statistical analysis

The primary analysis included associations of metabolite scores generated from the 82 overlapping metabolites between the Fenland study and the case-cohort study nested within the EPIC Norfolk study, with T2D risk. The secondary analysis included associations of metabolite scores generated from the 82 metabolites and the OCSFAs with T2D risk in the subset of participants with OCSFA measurements. Metabolite scores were standardised by dividing them by their standard deviation. Top metabolite scores refer to the scores, which included only the top metabolite signals with absolute coefficients higher than mean+2*SD.

We used Prentice-weighted Cox proportional hazards models to estimate hazard ratios (HRs) and their robust standard errors for the associations between the standardised metabolite scores and T2D risk. Associations were adjusted for previously established potential confounders. Specifically, model 1 included age, sex, educational level, socio-economic status and family history of T2D. Model 2 was additionally adjusted for smoking, physical activity, lipid-lowering medication, anti-hypertensive medication and hormone-replacement therapy (HRT). Model 3 was additionally adjusted for total energy intake and dietary factors and model 4 additionally included body mass index (BMI). The proportional-hazards assumption was tested based on Schoenfeld residuals.

7.4 Results

7.4.1 Descriptive characteristics

Descriptive characteristics of cases and non-cases are shown in **Table 7.1**. On average, cases were three years older than non-cases, included 16.4% more women, had a lower educational level and socio-economic status, were more frequently former smokers than non-smokers, were less active, had higher BMI, and were more frequently on lipid-lowering and anti-hypertensive medication. At baseline, future T2D case participants were also more frequently non-consumers of yoghurt and consumed more processed meat, sugar sweetened beverages and alcoholic beverages than non-cases.

Table 7.1 Descriptive characteristics of socio-demographic, lifestyle, clinical and dietary factors in the EPIC Norfolk diabetes case cohort study*

		Total		**T2D cases**		**Non-cases**	
Participants (N)		1,440		641		799	
Socio-demographic factors							
Age (years)		60.1	9.0	61.8	8.2	58.8	9.4
Sex	Women	48.9		58.0		41.6	
Educational level (ref. Low)	Medium	37.6		34.9		39.8	
	High	12.3		10.1		14	
Socio-economic status (ref. Low)	Medium	16.0		13.4		18.0	
	High	42.1		39.3		44.3	
Lifestyle factors							
Smoking status (ref. Never)	Former smoker	43.0		48.8		38.3	
	Current smoker	12.7		12.0		13.3	
Physical activity (ref. Inactive)	Moderately inactive	26.9		24.2		29.2	
	Moderately active	21.9		18.7		24.5	
	Active	16.7		13.9		19.0	
Energy intake (kcal/d)		2,023	578	2,035	608	2,013	554
BMI (kg/m^2)		27.6	4.5	29.6	4.6	26.0	3.6
Medications / Supplements							
Lipid-lowering medication	Yes	1.9		2.7		1.4	
Anti-hypertensive medication	Yes	23.6		34.2		15.1	
Hormone replacement therapy	Yes	16.0		11.9		19.4	
Dietary supplements †	Yes	54.2		58.7		50.6	
Types of dairy products							
Milk	Energy density †	1.8	0.9	1.8	0.9	1.8	1.0
	< 1 serving/d	21.5		20.4		22.4	
	≥ 2 serving/d	41.0		41.5		40.6	
Yoghurt	< 1 serving/wk	53.5		59.4		48.7	
	≥ 1 serving/d	7.3		6.1		8.3	
Cheese	< 1 serving/wk	24.8		27.9		22.3	
	≥ 1 serving/d	8.2		7.5		8.8	
Butter	< 1 serving/wk	63.4		65.2		62.0	
	≥ 1 serving/d	19.4		20.6		18.5	
Total dairy products	Energy density †	3.1	1.4	3.1	1.4	3.1	1.4
	< 1 servings/d	3.9		3.9		3.9	
	≥ 3 servings/d	43.8		43.1		44.4	
Non-dairy dietary factors (g/d)							
Fruits		241.1	198.9	231.1	198.0	249.2	199.5
Vegetables		233.1	120.3	224.2	121.8	240.2	118.6
Cereals		154.4	85.5	151.7	89.3	156.7	82.3
Red meat		64.3	46.0	65.1	45.7	63.6	46.2
Processed meat		31.4	25.5	34.5	27.0	28.9	24.0
Fish		37.2	25.5	36.0	25.4	38.1	25.6
Margarines		17.2	16.7	18.1	18.2	16.5	15.4
Sweet snacks		115.9	85.9	113.9	86.3	117.5	85.6
Sugar-sweetened beverages		41.0	84.8	49.5	96.8	34.1	73.2
Coffee		385.1	337.4	364.9	317.5	401.4	351.9
Tea		636.2	370.0	630.0	373.6	641.2	367.3
Alcoholic beverages		120.7	229.9	134.6	257.2	109.5	205.0

*The mean and SD are presented for continuous variables and column percentages are presented for categorical variables. Missing values for each variable were < 2% with total non-overlapping missing values of 3.8% across all variables

†Energy from milk or total dairy products in 2,000 kcal of total energy intake /day

Abbreviations: T2D: type 2 diabetes

7.4.2 Associations of metabolite scores with type 2 diabetes risk

HRs from the Cox proportional hazard models and their 95% CIs for associations of total and top (**Table 7.2**) metabolite scores and T2D risk are presented separately for the dairy types which showed evidence of external validity (milk, butter, total dairy products; **Tables 7.3, 7.4**) and for those which did not (yoghurt, cheese; **Tables 7.5, 7.6**).

Primary results: Total and top metabolite scores without odd-chain fatty acids

The metabolite score for the binary variable of milk consumption was inversely associated with T2D in multivariable models without adjustment for BMI [HR=0.89 (95% CI: 0.82, 0.97) per 1 SD of the total metabolite score and HR=0.77 (0.70, 0.83) per 1 SD of the top metabolite score, Table 7.3]. Results for the metabolite scores of the continuous variable of milk were similar, but stronger [HR=0.78 (0.72, 0.85) per 1 SD of the total metabolite score and HR=0.75 (0.69, 0.82) per 1 SD of the top metabolite score].

The top metabolite score for butter was associated with a significant lower risk of T2D by 28% per 1 SD of the metabolite score [HR=0.72 (0.66, 0.78)]. Although the total metabolite score did not significantly predict butter consumption, it also showed a significant inverse association with T2D risk.

Metabolite scores for total dairy consumption (binary and continuous variables) were overall significantly inversely associated with T2D risk. For the scores reflecting binary consumption, 1 SD of the top metabolite score, which significantly predicted total dairy consumption was associated with 26% lower T2D risk [HR=0.74 (0.68, 0.80)]. Similar associations were observed for the scores of the continuous variable of total dairy consumption.

After adjustment for BMI, all associations were attenuated, but were still significant apart from the association between the total metabolite score for the binary variable of milk consumption and T2D risk.

Associations for metabolite scores reflecting yoghurt and cheese consumption were either positive or null (Table 7.4).

Secondary results: Total and top metabolite scores including odd-chain saturated fatty acids

After inclusion of OCSFAs in the discovery analysis for the metabolite scores, results for the scores, which significantly predicted dairy consumption were similar with results without the OCSFAs (Table 7.5). One SD of the total and top metabolite scores for the binary variable of milk was associated with 27% and 26% lower risk of T2D respectively [HR=0.73 (0.65, 0.82) and 0.74 (0.66, 0.83)] in multivariable models without adjustment for BMI. Results for the metabolite scores of the continuous variable of milk were similar. Metabolite scores for butter were associated with 24% lower T2D risk per 1 SD of the

top score [0.76 (0.68, 0.86)]. Metabolite scores for the binary variable of total dairy consumption were associated with a 25% T2D risk reduction per 1 SD of the total score [HR=0.75 (0.66, 0.84)] and 20% risk reduction per 1 SD of the top score [HR=0.80 (0.71, 0.89)]. Similar associations were observed for the scores reflecting the continuous variable of total dairy consumption.

After adjustment for BMI associations were attenuated, but still significant.

Associations for metabolite scores reflecting yoghurt and cheese consumption were either positive or null (Table 7.6)

Table 7.2 Metabolites included in the top metabolite scores for total and types of dairy consumption from elastic net regression models using metabolomics data in the discovery set of the Fenland study

Dairy type related to metabolite score **Metabolite set: 82 metabolites***	**Top metabolites †**
Milk (binary)	SM-OH C14:1, LPC a C20:3, SM C16:1
Milk (continuous-energy density)	SM-OH C14:1, SM C16:1, LPC a C20:3
Yoghurt	SM-OH C14:1, trans-OH-Pro
Cheese	SM-OH C14:1, trans-OH-Pro, LPC a C18:0, C0
Butter	SM-OH C14:1, PC ae C34:1, LPC a C17:0, LPC a C14:0, PC aa C32:0
Total dairy products (binary)	SM-OH C14:1, SM C16:1, LPC a C18:0, LPC a C17:0
Total dairy products (continuous-energy density)	SM-OH C14:1, LPC a C18:0, LPC a C17:0, SM C16:1
Metabolite set: 82 metabolites and OCSFAs‡	
Milk (binary)	SM-OH C14:1, SM C16:1
Milk (continuous-energy density)	SM-OH C14:1, SM C16:1, LPC a C16:1
Yoghurt	LPC a C18:2, SM-OH C14:1, LPC a C20:4, C10, C8
Cheese	trans-OH-Pro, C15:0, C3
Butter	SM-OH C14:1, C15:0, LPC a C17:0, PC ae C34:1, SM C16:1
Total dairy products (binary)	SM-OH C14:1, SM C16:1, C15:0
Total dairy products (continuous-energy density)	SM-OH C14:1, SM C16:1

*82 metabolites overlapping between the Fenland study and the incident diabetes case-cohort sub-study within the EPIC Norfolk study, UK

†Top metabolites were defined as the metabolites with absolute values of the elastic net coefficients higher than their mean*2SD and are presented in decreasing order of magnitude of the absolute value of the elastic net coefficients

‡82 metabolites overlapping between the Fenland study and the incident diabetes case-cohort sub-study within the EPIC Norfolk study, UK (section 6.10.2) and OCSFAs included in the discovery

Abbreviations: LPC: Lysophosphatidylcholines; OCSFA: Odd-chain fatty acid; PC: Phosphatidylcholine; SM: Sphyngomyeline

Table 7.3 Hazard ratios of type 2 diabetes per 1 SD of metabolite scores significantly predicting dairy consumption in the incident diabetes case-cohort set nested within the EPIC Norfolk study

	Model	HR	95% CI	*p*
Total metabolite scores*				
Milk (binary) score	Socio-demographic + family history of T2D †	0.86	0.79, 0.93	< 0.001
	+Smoking, physical activity, medications ‡	0.88	0.81, 0.95	0.002
	+ Diet §	0.89	0.82, 0.97	0.009
	+ BMI	0.95	0.87, 1.04	0.27
Milk (continuous-energy density) score	Socio-demographic + family history of T2D †	0.76	0.70, 0.82	< 0.001
	+Smoking, physical activity, medications ‡	0.78	0.72, 0.84	< 0.001
	+ Diet §	0.78	0.72, 0.85	< 0.001
	+ BMI	0.86	0.78, 0.94	0.001
Butter score	Socio-demographic + family history of T2D †	0.56	0.52, 0.60	< 0.001
	+Smoking, physical activity, medications ‡	0.58	0.53, 0.62	< 0.001
	+ Diet §	0.57	0.52, 0.62	< 0.001
	+ BMI	0.62	0.57, 0.68	< 0.001
Total dairy products (binary) score	Socio-demographic + family history of T2D †	0.74	0.69, 0.80	< 0.001
	+Smoking, physical activity, medications ‡	0.76	0.71, 0.83	< 0.001
	+ Diet §	0.76	0.69, 0.83	< 0.001
	+ BMI	0.84	0.77, 0.92	< 0.001
Total dairy products (continuous-energy density) score	Socio-demographic + family history of T2D †	0.71	0.66, 0.77	< 0.001
	+Smoking, physical activity, medications ‡	0.73	0.68, 0.79	< 0.001
	+ Diet §	0.72	0.66, 0.79	< 0.001
	+ BMI	0.81	0.74, 0.89	< 0.001
Top metabolite scores* ‖				
Milk (binary) score	Socio-demographic + family history of T2D †	0.75	0.70, 0.81	< 0.001
	+Smoking, physical activity, medications ‡	0.77	0.71, 0.83	< 0.001
	+ Diet §	0.77	0.70, 0.83	< 0.001
	+ BMI	0.83	0.76, 0.91	< 0.001
Milk (continuous-energy density) score	Socio-demographic + family history of T2D †	0.74	0.68, 0.79	< 0.001
	+Smoking, physical activity, medications ‡	0.75	0.70, 0.81	< 0.001
	+ Diet §	0.75	0.69, 0.82	< 0.001
	+ BMI	0.83	0.76, 0.90	< 0.001
Butter score	Socio-demographic + family history of T2D †	0.69	0.64, 0.75	< 0.001
	+Smoking, physical activity, medications ‡	0.71	0.66, 0.77	< 0.001
	+ Diet §	0.72	0.66, 0.78	< 0.001
	+ BMI	0.77	0.70, 0.84	< 0.001
Total dairy products (binary) score	Socio-demographic + family history of T2D †	0.72	0.67, 0.78	< 0.001
	+Smoking, physical activity, medications ‡	0.74	0.68, 0.80	< 0.001
	+ Diet §	0.74	0.68, 0.80	< 0.001
	+ BMI	0.83	0.76, 0.90	< 0.001
Total dairy products (continuous-energy density) score	Socio-demographic + family history of T2D †	0.70	0.65, 0.76	< 0.001
	+Smoking, physical activity, medications ‡	0.72	0.66, 0.78	< 0.001
	+ Diet §	0.71	0.65, 0.78	< 0.001
	+ BMI	0.80	0.73, 0.87	< 0.001

*Metabolite scores generated from the total sample of the incident diabetes case-cohort study (N=1,440) with 641 incident diabetes cases during 16,30 person-years of follow-up

†Models adjusted for age (continuous in years), sex, educational level (3 categories: low, medium, high), socio-economic status (3 categories: low, medium, high) and family history of T2D (2 categories: Yes, No)

‡Models additionally adjusted for smoking (3 categories: never, former, current smoker), physical activity (4 categories: inactive, moderately inactive, moderately active, active), lipid-lowering medication (2 categories: yes, no), anti-hypertensive medication (2 categories: yes, no), hormone-replacement therapy (3 categories: yes, no, men)

§Models additionally adjusted for total energy intake (continuous in kcal/day), dietary supplement use (2 categories: yes, no), consumption (g/day) of fruit, vegetables, total cereals, red meat, processed meat, fish, margarine, sweet snacks, sugar-sweetened beverages, coffee, tea, alcoholic beverages

‖Top metabolites are defined as the metabolites with coefficients above mean+2*SD. The top metabolite scores for each dairy type are listed in Table 7.2

Abbreviations: BMI: Body mass index; CI: Confidence interval; HR: Hormone-replacement therapy; T2D: Type 2 diabetes

Table 7.4 Hazard ratios of type 2 diabetes per 1 SD of metabolite scores not significantly predicting dairy consumption in the incident diabetes case-cohort set nested within the EPIC Norfolk study

	Model	HR	95% CI	*p*
Total metabolite scores*				
Yoghurt score	Socio-demographic + family history of T2D †	1.29	1.19, 1.40	< 0.001
	+Smoking, physical activity, medications ‡	1.32	1.22, 1.43	< 0.001
	+ Diet §	1.32	1.22, 1.43	< 0.001
	+ BMI	1.30	1.19, 1.41	< 0.001
Cheese score	Socio-demographic + family history of T2D †	1.08	0.99, 1.17	0.068
	+Smoking, physical activity, medications ‡	1.11	1.02, 1.20	0.012
	+ Diet §	1.13	1.04, 1.23	0.004
	+ BMI	1.11	1.02, 1.21	0.014
Top metabolite scores* ‖				
Yoghurt score	Socio-demographic + family history of T2D †	0.94	0.87, 1.02	0.146
	+Smoking, physical activity, medications ‡	0.98	0.90, 1.06	0.556
	+ Diet §	1.00	0.91, 1.09	0.915
	+ BMI	1.04	0.96, 1.14	0.341
Cheese score	Socio-demographic + family history of T2D †	0.87	0.80, 0.94	< 0.001
	+Smoking, physical activity, medications ‡	0.91	0.84, 0.99	0.024
	+ Diet §	0.93	0.86, 1.02	0.121
	+ BMI	1.01	0.93, 1.10	0.832

*Metabolite scores generated from the total sample of the incident diabetes case-cohort study (N=1,440) with 641 incident diabetes cases during 16,30 person-years of follow-up

†Models adjusted for age (continuous in years), sex, educational level (3 categories: low, medium, high), socio-economic status (3 categories: low, medium, high) and family history of T2D (2 categories: Yes, No)

‡Models additionally adjusted for smoking (3 categories: never, former, current smoker), physical activity (4 categories: inactive, moderately inactive, moderately active, active), lipid-lowering medication (2 categories: yes, no), anti-hypertensive medication (2 categories: yes, no), hormone-replacement therapy (3 categories: yes, no, men)

§Models additionally adjusted for total energy intake (continuous in kcal/day), dietary supplement use (2 categories: yes, no), consumption (g/day) of fruit, vegetables, total cereals, red meat, processed meat, fish, margarine, sweet snacks, sugar-sweetened beverages, coffee, tea, alcoholic beverages

‖Top metabolites are defined as the metabolites with coefficients above mean+2*SD. The top metabolite scores for each dairy type are listed in Table 7.2

Abbreviations: BMI: Body mass index; CI: Confidence interval; HR: Hormone-replacement therapy; T2D: Type 2 diabetes

Table 7.5 Hazard ratios of type 2 diabetes per 1 SD of metabolite scores significantly predicting dairy consumption in the OCSFA subset of the incident diabetes case-cohort set nested within the EPIC Norfolk study

	Model	HR	95% CI	*p*
Total metabolite scores *				
Milk (binary) score	Socio-demographic + family history of T2D †	0.69	0.62, 0.77	< 0.001
	+Smoking, physical activity, medications ‡	0.73	0.65, 0.82	< 0.001
	+ Diet §	0.73	0.65, 0.83	< 0.001
	+ BMI	0.79	0.70, 0.90	< 0.001
Milk (continuous-energy density) score	Socio-demographic + family history of T2D †	0.68	0.62, 0.76	< 0.001
	+Smoking, physical activity, medications ‡	0.71	0.64, 0.79	< 0.001
	+ Diet §	0.70	0.62, 0.78	< 0.001
	+ BMI	0.75	0.67, 0.85	< 0.001
Butter score	Socio-demographic + family history of T2D †	0.58	0.52, 0.65	< 0.001
	+Smoking, physical activity, medications ‡	0.60	0.53, 0.67	< 0.001
	+ Diet §	0.58	0.50, 0.66	< 0.001
	+ BMI	0.64	0.56, 0.73	< 0.001
Total dairy products (binary) score	Socio-demographic + family history of T2D †	0.74	0.66, 0.82	< 0.001
	+Smoking, physical activity, medications ‡	0.77	0.69, 0.86	< 0.001
	+ Diet §	0.75	0.66, 0.84	< 0.001
	+ BMI	0.81	0.72, 0.92	0.001
Total dairy products (continuous-energy density) score	Socio-demographic + family history of T2D †	0.73	0.66, 0.81	< 0.001
	+Smoking, physical activity, medications ‡	0.76	0.69, 0.85	< 0.001
	+ Diet §	0.75	0.67, 0.84	< 0.001
	+ BMI	0.81	0.72, 0.91	0.001
Top metabolite scores * ‖				
Milk (binary) score	Socio-demographic + family history of T2D †	0.73	0.66, 0.81	< 0.001
	+Smoking, physical activity, medications ‡	0.76	0.69, 0.84	< 0.001
	+ Diet §	0.74	0.66, 0.83	< 0.001
	+ BMI	0.79	0.71, 0.89	< 0.001
Milk (continuous-energy density) score	Socio-demographic + family history of T2D †	0.77	0.69, 0.85	< 0.001
	+Smoking, physical activity, medications ‡	0.79	0.71, 0.88	< 0.001
	+ Diet §	0.77	0.69, 0.87	< 0.001
	+ BMI	0.81	0.72, 0.91	< 0.001
Butter score	Socio-demographic + family history of T2D †	0.75	0.68, 0.83	< 0.001
	+Smoking, physical activity, medications ‡	0.77	0.70, 0.86	< 0.001
	+ Diet §	0.76	0.68, 0.86	< 0.001
	+ BMI	0.81	0.72, 0.91	0.001
Total dairy products (binary) score	Socio-demographic + family history of T2D †	0.78	0.71, 0.87	< 0.001
	+Smoking, physical activity, medications ‡	0.81	0.73, 0.90	< 0.001
	+ Diet §	0.80	0.71, 0.89	< 0.001
	+ BMI	0.85	0.76, 0.95	0.005
Total dairy products (continuous-energy density) score	Socio-demographic + family history of T2D †	0.78	0.71, 0.86	< 0.001
	+Smoking, physical activity, medications ‡	0.81	0.73, 0.90	< 0.001
	+ Diet §	0.79	0.71, 0.89	< 0.001
	+ BMI	0.85	0.75, 0.95	0.004

*Metabolite scores generated from the subset of the case-cohort study with available measurements of OCSFAs including 592 participants, 356 cases and 5,569 person-years of follow-up

†Models adjusted for age (continuous in years), sex, educational level (3 categories: low, medium, high), socio-economic status (3 categories: low, medium, high) and family history of T2D (2 categories: Yes, No)

‡Models additionally adjusted for smoking (3 categories: never, former, current smoker), physical activity (4 categories: inactive, moderately inactive, moderately active, active), lipid-lowering medication (2 categories: yes, no), anti-hypertensive medication (2 categories: yes, no), hormone-replacement therapy (3 categories: yes, no, men)

§Models additionally adjusted for total energy intake (continuous in kcal/day), dietary supplement use (2 categories: yes, no), consumption (g/day) of fruit, vegetables, total cereals, red meat, processed meat, fish, margarine, sweet snacks, sugar-sweetened beverages, coffee, tea, alcoholic beverages

‖Top metabolites are defined as the metabolites with coefficients above mean+2*SD. The top metabolite scores for each dairy type are listed in Table 7.2

Abbreviations: BMI: Body mass index; CI: Confidence interval; HR: Hormone-replacement therapy; OCSFA: Odd-chain fatty acid; T2D: Type 2 diabetes

Table 7.6 Hazard ratios of type 2 diabetes per 1 SD of metabolite scores not significantly predicting dairy consumption in the OCSFA subset of the incident diabetes case-cohort set nested within the EPIC Norfolk study

	Model	HR	95% CI	*p*
Total metabolite scores *				
Yoghurt score	Socio-demographic + family history of T2D †	1.37	1.23, 1.52	< 0.001
	+Smoking, physical activity, medications ‡	1.35	1.21, 1.50	< 0.001
	+ Diet §	1.35	1.21, 1.51	< 0.001
	+ BMI	1.27	1.14, 1.43	< 0.001
Cheese score	Socio-demographic + family history of T2D †	0.95	0.85, 1.06	0.354
	+Smoking, physical activity, medications ‡	1.03	0.92, 1.16	0.59
	+ Diet §	1.02	0.91, 1.14	0.778
	+ BMI	1.10	0.98, 1.24	0.112
Top metabolite scores * ‖				
Yoghurt score	Socio-demographic + family history of T2D †	1.22	1.10, 1.35	< 0.001
	+Smoking, physical activity, medications ‡	1.19	1.07, 1.32	0.001
	+ Diet §	1.18	1.06, 1.32	0.004
	+ BMI	1.11	0.99, 1.24	0.071
Cheese score	Socio-demographic + family history of T2D †	0.92	0.83, 1.03	0.16
	+Smoking, physical activity, medications ‡	0.99	0.89, 1.11	0.866
	+ Diet §	0.99	0.89, 1.11	0.927
	+ BMI	1.03	0.92, 1.16	0.555

*Metabolite scores generated from the subset of the case-cohort study with available measurements of OCSFAs including 592 participants, 356 cases and 5,569 person-years of follow-up

†Models adjusted for age (continuous in years), sex, educational level (3 categories: low, medium, high), socio-economic status (3 categories: low, medium, high) and family history of T2D (2 categories: Yes, No)

‡Models additionally adjusted for smoking (3 categories: never, former, current smoker), physical activity (4 categories: inactive, moderately inactive, moderately active, active), lipid-lowering medication (2 categories: yes, no), anti-hypertensive medication (2 categories: yes, no), hormone-replacement therapy (3 categories: yes, no, men)

§Models additionally adjusted for total energy intake (continuous in kcal/day), dietary supplement use (2 categories: yes, no), consumption (g/day) of fruit, vegetables, total cereals, red meat, processed meat, fish, margarine, sweet snacks, sugar-sweetened beverages, coffee, tea, alcoholic beverages

‖Top metabolites are defined as the metabolites with coefficients above mean+2*SD. The top metabolite scores for each dairy type are listed in Table 7.2

Abbreviations: BMI: Body mass index; CI: Confidence interval; HR: Hormone-replacement therapy; OCSFA: Odd-chain fatty acid; T2D: Type 2 diabetes

7.5 Discussion

7.5.1 Summary of results and overall interpretation

Metabolite scores, which significantly predicted milk, butter and total dairy consumption were associated with a lower risk of T2D. In contrast, metabolite scores, which predicted yoghurt and cheese in the Fenland study, but not in the EPIC Norfolk study, were associated with either a higher risk or no risk of T2D.

These associations might be indicative of related metabolic pathways linking dairy consumption to the risk of T2D, but the exact interpretation is challenging. The scores did not completely reflect dairy consumption, but only specific aspects of them, which are limited to the set of metabolites that we used. Thus, interpretation of these findings should be done in combination with the results from the external validation. An example of such interpretation for total dairy products is that with an AUC of 0.68 from multivariable prediction models and a significant contribution of the metabolite score to this prediction, an HR of 0.83 can be interpreted as "a decrease in the risk of T2D by 17% per 1 SD of the metabolite score, which was associated with a 68% probability of reflecting consumers of total dairy products after accounting for socio-demographic and lifestyle parameters".

7.5.2 Findings in the context of previous evidence

Distinct from studies on associations between OCSFAs and T2D risk (section 7.1), which reported null or inverse associations, some previous studies explored metabolomic profiles discriminating people by diabetes status for the prediction of T2D. These studies appraised metabolites in the context of T2D prediction and not in relation to dairy consumption.

SM-OH C14:1 identified as one of the top metabolites predicting any of the dairy types we examined, was inversely associated with prevalent T2D in a Korean[293] and a German population[294] and with prevalent impaired glucose tolerance status in a German population[294].

LPC C17:0, which was one of the top signals for butter and total dairy consumption in our study, was associated with a lower risk of T2D in a Swedish nested case-control study[295], with a lower prevalence of T2D and impaired glucose tolerance in a German cross-sectional study[294] and with lower levels of 2-hour glucose in a cross-sectional Korean study[293]. LPC C17:0 contains C17:0, which was previously associated with lower T2D incidence[228, 292].

PC ae C34:1, another top signal for butter consumption, was associated with lower levels of 2-hour glucose in a cross-sectional Korean study[293].

SM C16:1, a top signal for milk and total dairy consumption, but not after adjustment for other dietary factors, was associated with a lower risk of T2D[296], lower prevalence of T2D[294, 297] and impaired glucose tolerance[294] and lower 2-hour glucose levels[293].

However, this metabolite was inversely associated with dairy consumption, but due to the attenuation of the signal after adjustment for other dietary factors, this metabolite might have not contributed much to the association between the metabolite scores and T2D risk after adjustment for dietary factors.

Overall, the inverse associations of metabolite scores for milk, butter and total dairy products with T2D risk that we reported, are supported by previous evidence on associations between metabolites that contributed the most to our scores and T2D risk or related endpoints. These associations are also in accordance with associations between dairy consumption and T2D risk as reported in meta-analyses of prospective cohort studies for butter[14] and total dairy products[8, 12, 18, 19, 21, 22]. Most of the meta-analyses on milk reported null associations[12, 18, 19], which might mean that sets of biological pathways other than the potential pathway related to our results might differentially link milk to T2D.

7.5.3 Interpretation and potential mechanisms

The interpretation of the observed inverse associations between metabolite scores predictive of milk, butter and total dairy consumption and T2D risk is complex and should be formed with caution. The extent to which these associations reflect the association between self-reported dairy consumption and T2D is limited to the characteristics of the dairy food matrix that our scores and the metabolites we used reflect. We considered potential biological mechanisms based on the metabolites most significantly associated with one or more types of dairy consumption. The complexity of the interpretation can be partly reduced if we focus on the associations between the top metabolite signals and T2D risk, as also reported from previous studies, assuming that these metabolites contributed the most to the overall observed associations. Most top signals were PCs and SMs.

The SMs that we identified as top signals i.e. SM-OH C14:1, SM C16:1, were associated also in other studies with a lower T2D risk. It might be helpful to consider the fatty acids that these SMs contain, as they represent isomeric and isobaric compounds that contain the OCSFAs C15:0 and C17:0 or the even-chain fatty acids C16:0 and C18:0. The lipids, which were positively associated with dairy consumption mainly included OCSFAs (LPC a C17:0) or consisted of isobaric or isomeric compounds, which included OCSFAs (SM d16:1/17:0 and SM d18:1/15:0 for SM-OH C14:1; PC 15:0_18:1 and PC 16:1_17:0 for PC ae C34:1). The metabolites inversely associated with dairy consumption (SM C16:1, LPC a C18:0) included mainly even-chain fatty acids. Since usually the lipids are measured in relative amounts, it is difficult to conclude whether the associations observed for a certain fatty acid can be attributed to that particular fatty acid or reflect its competitive relation with other fatty acids, that have an opposite effect. For example, C16:0 and C18:0 were associated with higher insulin resistance[298, 299] and inflammation[300, 301]. In addition to uptake from the diet, C16:0 and C18:0 are produced endogenously from de-novo

lipogenesis, which has been associated with hepatic steatosis and insulin resistance[302]. OCSFAs cannot be produced through this pathway, but there is evidence that they can be produced through alternative pathways from C16:0 and C18:0 (section 6.4)[233, 244, 245]. Thus, not only do OCSFAs have a neutral effect on the de-novo lipogenesis pathway, but since part of them can be produced from C16:0 and C18:0, higher levels of these fatty acids may imply lower levels of C16:0 and C18:0 and thus less harmful effects on β-cell apoptosis and insulin resistance.

We reported null or positive associations of metabolite scores for yoghurt and cheese with T2D risk. These scores failed to validate externally, so they might reflect random noise and are not interpretable. Thus, it is not possible to provide a biological explanation currently and the need for further research especially on these two dairy types is warranted.

7.5.4 Strengths and limitations

Our study on associations of metabolite scores predictive of dairy consumption with risk of type 2 diabetes has several strengths. We used metabolite scores, which reflected dairy consumption, with internal and external (for milk, butter and total dairy consumption) validity in an independent case-cohort set nested within the EPIC Norfolk study. In this way, we could avoid the measurement error of self-reported dietary assessment methods. In addition, although we did not identify single metabolites as specific biomarkers to individual dairy types, the use of scores with metabolite weights specific for each dairy type, is a way of developing biomarkers specific for dairy types.

This study has also several limitations. Metabolite scores for yoghurt and cheese did not show evidence of external validity in this cohort, so the associations between these scores and T2D were not interpretable. The scores were derived from an observational cross-sectional study using a self-reported method of dietary assessment, which might entail more confounding and measurement error than an RCT design. It is not clear how such error could influence our results, but it is possible for example, that due to confounding we identified metabolites that are not causally related to dairy consumption or we did not identify metabolites that are causally related. Nevertheless, it is also important to acknowledge that RCTs of dietary interventions pose their own challenges such as issues related to lack of blinding, adherence to the dietary advice, attrition and complexity of interpretation of total energy intake or intakes of dietary compounds other than the intervention foods.

7.5.5 Conclusion

In this study, we used a novel approach of developing and validating metabolite scores predictive of dairy consumption, which constitutes a more promising method to identify markers specific to individual dairy types. We observed a lower risk of T2D for metabolic

profiles significantly associated with a higher consumption of milk, butter and total dairy products. This study laid the foundations for the development and use of metabolite scores as potential biomarkers of dairy consumption and as tools to unravel pathways underpinning the link between dairy consumption and T2D. Further studies are needed with different sets of metabolites and different study designs to identify more valid scores predictive of dairy consumption especially for yoghurt and cheese, that we were not able to identify and which have shown associations with cardio-metabolic health. The combination of information from both observational and controlled studies could provide a more comprehensive understanding on how metabolomics profiles predictive of dairy consumption are related to T2D risk.

Chapter 8

Genetic predictors of dairy consumption

Summary

Background and aims: Genetic variants of lactase persistence (LP) are well-known genetic predictors of dairy consumption. These variants sufficiently predict milk consumption, but not consumption of yoghurt or cheese, which contain less lactose than milk and also contribute to an enhanced digestion of lactose. We aimed to investigate genetic predictors of total and types of dairy products (milk, yoghurt and cheese) consumption to elucidate the link of genetic predisposition to dairy consumption and develop potential tools for Mendelian randomisation analyses.

Methods: We conducted genome-wide association studies on total and types of dairy consumption in the UK Biobank. The primary dairy outcomes were milk consumed as a drink or added to cereals (n=451,900), yoghurt (n=193,505), cheese (n=445,330) and total dairy consumption (n=193,505). Diet was assessed with a general questionnaire (three repeated measures available) and a questionnaire about diet in the last 24 hours (Oxford WebQ; five repeated measures available). Information from both tools was used to generate variables of dairy consumption. Total dairy consumption was expressed as a continuous variable in servings/day. For dairy types, participants were split into consumers and non-consumers. The associations of the single nucleotide polymorphism (SNP) for LP, rs4988235, with milk and total dairy consumption were used as positive controls. Genome-wide association analyses were conducted among white European-origin participants with linear mixed models adjusted for age and sex.

Results: The majority of the participants were milk and cheese consumers (92.9% and 97.1% respectively). Almost half of the participants were yoghurt consumers (46.7%). Highly significant positive associations were observed between rs4988235 and both milk consumed as a drink and added to cereals (p=2.3 x 10^{-12}) and total dairy consumption (p=1.1 x 10^{-15}) in the positive control analysis among white Europeans. We identified

seven SNPs on chromosomes 1, 2, 4 and 12 predictive of total dairy consumption; four SNPs on chromosome 2 predictive of milk consumption; one SNP on chromosome 1 predictive of yoghurt consumption and four SNPs on chromosomes 1, 2 and 7 predictive of cheese consumption, which reached genome-wide significance.

Conclusion: We identified novel genetic predictors of total and types of dairy consumption in addition to confirming the strong prediction of milk and total dairy consumption by the LP SNP rs4988235. To our knowledge, no genome-wide association study on total and types of dairy consumption has been previously reported. Based on the large sample size of the UK Biobank, enabling the detection of genetic associations of small magnitude, these findings lay the foundations for the use of such predictors to elucidate links of genetic predisposition to dairy consumption. This also opens up possibilities to the use of the identified SNPs as instrumental variables in Mendelian randomisation analyses for the investigation of causal associations with disease endpoints. The replication of our findings in other populations and in studies with more precise methods of dietary assessment will further advance our understanding of the genetic predictors of dairy consumption.

What is already known

- Prior candidate-gene research has shown that single nucleotide polymorphisms (SNPs) for lactase persistence predict milk consumption
- No genetic predictors of yoghurt and cheese consumption have been reported to date
- Mendelian randomisation is a useful tool to investigate causal associations
- SNPs have been used as instrumental variables in Mendelian randomisation studies on associations between milk or total dairy consumption and cardio-metabolic endpoints

What this research adds

- Seven SNPs predicted total dairy consumption, four SNPs predicted milk consumption, four SNPs predicted cheese consumption and one SNP predicted yoghurt consumption in genome-wide association studies of a large sample (n≈450,000) of white European participants in the UK Biobank study.
- The identified SNPs open up the possibility for the elucidation of genetic predisposition to dairy consumption and as instrumental variables in Mendelian randomisation analyses on dairy consumption and disease incidence.

Publication

Trichia E, Day FR, Imamura F, Perry J, Wareham NJ, Forouhi N G. Associations of total and types of dairy consumption with type 2 diabetes risk: a Mendelian Randomisation study (analysis ongoing)

NOTE: Definitions of concepts from genetic epidemiology, which are used throughout this Chapter, are provided in the glossary of **Appendix B**.

8.1 Lactase persistence

The exploration of the genetic predictors of dairy consumption has been limited so far to the investigation of genetic polymorphisms related to lactase persistence (LP). LP has been associated with higher consumption of lactose-containing foods, especially milk[303]. A meta-analysis of 5 observational studies reported that people with LP consumed on average more milk by 55 g/day compared to those with the lactase non-persistence phenotype[304]. Absence of lactase persistence leads to lactose intolerance. Related mechanisms are well-established. Lactase (lactase-phlorizin hydrolase) is an enzyme, which hydrolyses lactose into glucose and galactose in the intestinal epithelial cells[303]. LP into adulthood is the phenotype resulting from the heterozygosity or homozygosity for an autosomal dominant allele, whereas decline in lactase activity after weaning is a result of a genotype homozygous for the recessive allele[303].

The main single nucleotide polymorphisms (SNPs) identified to date in relation to LP, are rs4988235 (C/T-13910) and rs182549 (G/A-22018) in the gene *MCM6* on chromosome 2, which codes for a promoter of the enzyme lactase[303]. These SNPs are located upstream of the *LCT* gene, which codes for lactase[303]. In the absence of this genetic variant, down-regulation of the lactase enzyme activity is observed leading to lower lactose hydrolysis, absorption and higher transfer from the small intestine to the colon[303]. This results in osmotic load and unpleasant symptoms of lactose intolerance including abdominal discomfort, flatulence and diarrhoea and thus lower consumption of milk[303]. In addition, colonic bacteria metabolise lactose and use the metabolic products to produce short-chain fatty acids and hydrogen gas, which increase the aforementioned symptoms[303].

Evidence suggests that the LP phenotype is a recent (5,000-10,000 years) result of natural selection, while the default phenotype has been lactase non-persistence and this selection is one of the strongest observed in the human genome[305]. It has been hypothesised that LP was developed in regions that also developed dairy farming for food supply, especially pastoral regions, i.e. regions where dairy farming was the main economic activity. As can be seen in **Figure 8.1a**, there is a higher LP prevalence in northern Europe (89-96%) compared to southern and eastern Europe (19-54%), western and northern India (63%) compared to southern India (23%) and in pastoral regions of Africa (64% among the pastoralists Sudanese Beni Amir) compared to non-pastoral regions (20% among the non-pastoralists Sudanese Dounglawi), while the lowest prevalence is observed in Asia with about 1% in China[306].

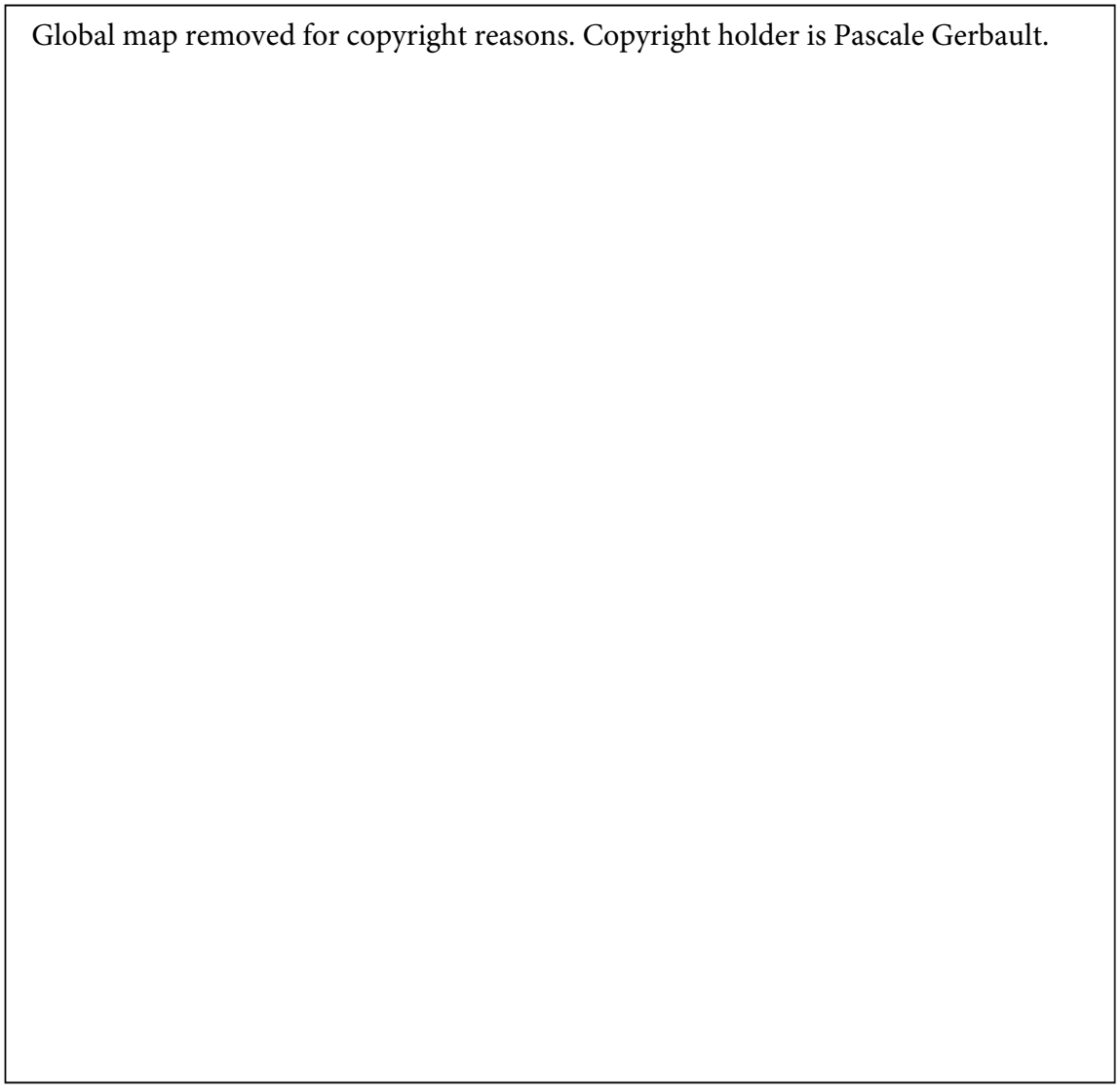


Fig. 8.1 Global distribution of (a) lactase persistence (LP) and (b) the LP-associated -213910*T allele. Colour scale represents magnitude of the LP frequency. Figure adapted from Gerbault et al.[306].

This observation coincides with archaeological evidence on the spread of dairy farming into Europe and certain African regions[306]. It has been assumed that along with the introduction of dairy farming, the low exposure to sunlight in northern Europe and the high drought levels in Africa, contributed to higher consumption of milk and gradual natural selection, a phenomenon known as gene-culture co-evolution[306]. In addition to that, more LP SNPs have been identified for different ethnic groups, which is expected since rs4988235 explains LP mainly in Europe (**Figure 8.1b**)[306]. For example, rs41380347 (T/G-13915) has been identified as an LP SNP in Saudi Arabia and rs41525747 (C/G-13907) in Ethiopia and Sudan[307]. The multiple SNPs identified might also indicate that LP developed simultaneously and independently in different populations[307]. The long haplotype observed in relation to LP increases the possibility that more causal genetic variants of LP are located on it apart from the one identified in each population such as rs4988235 in Europeans [305].

The ethnicity-specific genetic differences observed for LP are also reflected in dairy consumption levels, as shown in the "old world" map (Europe, Africa, Australasia) in **Figure 8.2**, where milk consumption levels in 2010 were higher in northern Europe and some parts of Africa. A more specific example of two countries with very different milk

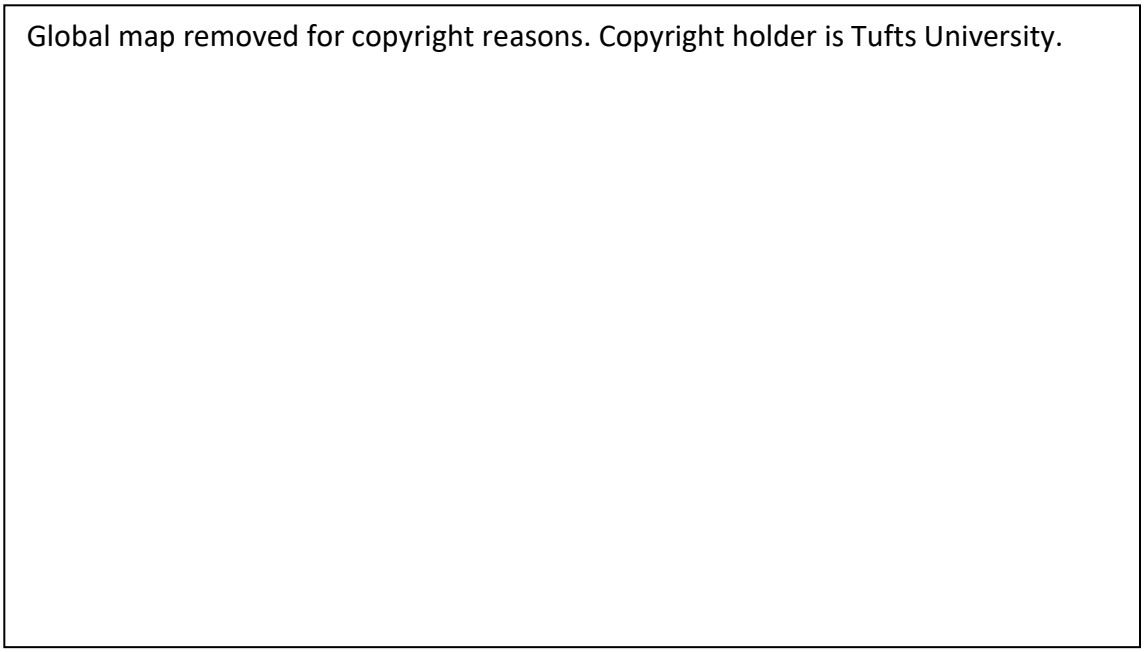


Fig. 8.2 Global distribution of milk consumption in 2010, Global Dietary Database (https://www.globaldietarydatabase.org/country-comparisons.html, date of access: 24 June 2018)

consumption levels such as the UK (260.4 g/day in 2010) and China (17.9 g/day in 2010) is shown in **Figure 8.3**.

Pie charts removed for copyright reasons. Copyright holder is Tufts University.

Fig. 8.3 Milk contribution to the total amount of food consumed in the UK and China in 2010, Global Dietary Database (https://www.globaldietarydatabase.org/ country-comparisons.html, date of access: 24 June 2018)

LP is a strong genetic predictor of dairy consumption, but it has some limitations. First, while the LP genetic variant can predict milk consumption, it is not suitable for the prediction of the other two main dairy types i.e. yoghurt and cheese[308]. Depending on the dairy type, the lactose content might not be much lower than in milk (**Table 8.1**), but fermented dairy products contain bacteria, which contribute to a more efficient metabolism of lactose compared to the non-fermented products[309]. Although it could be expected that the bacterial lactase would be deactivated when passing through the acidic stomach, the buffering effect of yoghurt and the bacterial cell membrane, protect the bacterial lactase. When yoghurt reaches the small intestine, which has a higher pH, the bacterial lactase is activated and contributes to the metabolism of lactose[309]. Second, the LP genetic variants cannot explain the large variability of symptoms of lactose intolerance. Although there is evidence that on average people with lactose intolerance can tolerate up to 12 g lactose daily[303], there is variation by genetic and environmental factors. These include sex (symptoms seem to be worse among women than men)[310], the distribution of lactose consumption throughout the day (24 g of lactose might be tolerated when spread throughout the day compared to acute consumption)[303], the consumption of lactose as part of a meal (better tolerated than consumed alone)[303] and composition of the gut microbiota[311]. It is therefore of interest to explore additional genetic predictors of dairy consumption.

Table 8.1 Lactose content in 100 g of different types of dairy products[35]

Dairy products	Sub-type	Lactose content (%)
Milk	Cow's, any fat content	4.7
	Goat	4.4
	Sheep	5.1
Cheese	Brie	Trace
	Cheddar	0.1
	Spread	4.4
	Cottage	3.1
	Edam / Gouda	Trace
	Feta	1.4
	Goats	0.9
	Mozzarella	Trace
	Parmesan	0.9
Yoghurt	Plain	4.7
	Greek style	3.5

8.2 Mendelian randomisation

Apart from the understanding of potential biological pathways that may lead to consumption of different dairy types, the identification of genetic predictors of dairy products could also contribute to their use as instrumental variables in Mendelian randomisation analyses. Mendelian randomisation is increasingly used in epidemiology to investigate causal associations using observational data and it has some similarities with the design of a randomised controlled trial (**Figure 8.4**)[312, 313]. It is based on Mendel's second law of random inheritance of alleles by using genetic proxies of the exposure (instrumental variables), to assess associations with a lower risk of reverse causation and confounding that are common challenges in observational studies[313]. This method might be particularly useful when studying dietary factors, for which long-term trials are usually not feasible.

There are several studies which used Mendelian randomisation to investigate causal associations of dairy consumption with cardio-metabolic disease outcomes. Results from these studies indicated a non-significant association of genetically predicted milk consumption with type 2 diabetes (T2D)[304, 314], myocardial infarction[315] and ischaemic heart disease[304, 315]. Mendelian randomisation studies of associations of milk consumption with cardio-metabolic markers, reported positive associations with BMI[304, 308], obesity[308] and insulin[304]; null associations with the ratio of waist to hip circumference[304], systolic (SBP) and diastolic blood pressure (DBP)[308], Haemoglobin A1c (HbA1c), fasting and 2-hour glucose, and triglycerides[304]; and inverse associations with low-density lipoprotein (LDL-C) and high-density lipoprotein cholesterol (HDL-C)[304]. Likewise, for total dairy consumption, a meta-analysis of 25

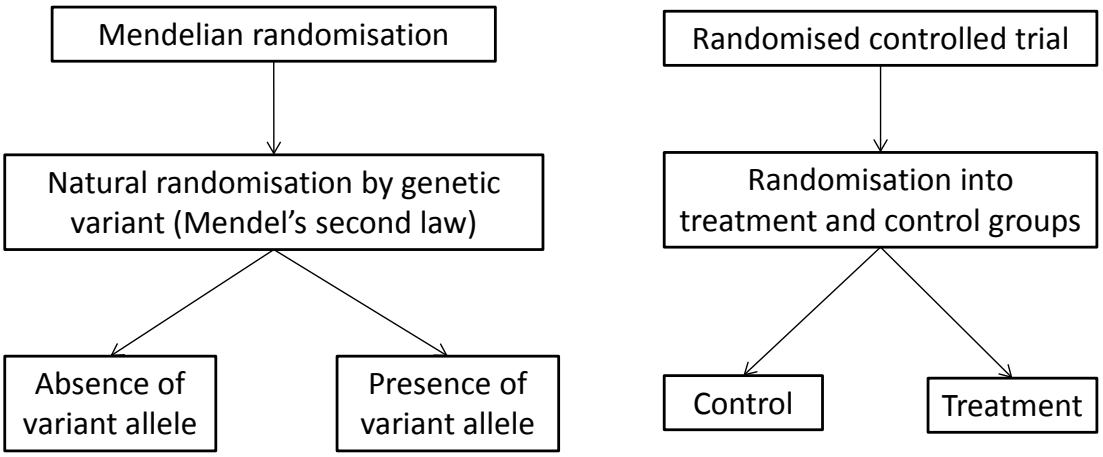


Fig. 8.4 Comparison of Mendelian randomisation with randomised controlled trials. Figure content adapted from Burgess et al[312].

Mendelian randomisation studies reported a positive association with BMI and a Mendelian randomisation analysis of 22 studies from the CHARGE (Cohort for Heart and Ageing Research in Genomic Epidemiology) Consortium reported non-significant associations with SBP or hypertension[84].

All these studies used as instrumental variable the LP SNP rs4988235, as the most established genetic predictor of milk consumption. As described above, the use of the SNP entails several limitations. First, it can be used as a genetic proxy for milk consumption only and not for other dairy types. Second, even for milk consumption it is subject to within individual variability. Third, the LP SNP might not fulfil the requirements of a valid instrumental variable, which should be independent of potential confounders of the association under study[313]. For example, a potential confounder of the association between genetically predicted milk consumption and BMI could be total energy intake if we assume that lactose intolerant people might follow a more restricted diet independent of dairy consumption than LP people. These limitations provide the rationale for further discovery efforts to identify genetic variants for different types of dairy products.

8.3 Study aims

The aim of this study was to investigate genetic predictors of the three main dairy types i.e. milk, yoghurt and cheese using an exploratory approach with genome-wide association studies instead of the candidate gene approach followed in previous studies, and mainly concerned the LP SNPs. This aim serves a dual purpose: a better understanding of the biological pathways related to the profiles of dairy consumers and non-consumers; and the use of the genetic predictors as instrumental variables in Mendelian randomisation analyses.

8.4 Methods

8.4.1 Study design and population

For the genome-wide association study (GWAS), we evaluated data from the UK Biobank, a large scale prospective cohort study of approximately 503,000 participants, aged 40-69 years. Participants were recruited from April 2007 to July 2010 from 22 assessment centres in England, Scotland and Wales through population-based registries and based on living close to an assessment centre with an overall response rate of 5.5%. A subset of 20,345 people living within a 35 km radius from the Stockport study centre from a total of 103,514 participants who were invited, participated in the first repeated assessment in 2012-2013 and a subset of a similar size participated in the second repeated assessment in 2015. Ethical approval was obtained by the National Information Governance Board for Health and Social Care and the NHS Northwest Multicentre Research Ethics Committee (MREC).

8.4.2 Genotyping

Blood samples were stored in racks of 96 microtubes (1.2 ml) at -80°C or -196°C and placed in 96-well plates for DNA extraction. Participants were genotyped with the Affymetrix Axiom array (n≈450,000; Affymetrix GeneTitan® Mutli-Channel Instrument, Affymetrix Research Services Laboratory, Santa Clara, CA, USA) or the UK BiLEVE array (n≈50,000). Approximately the same procedures were followed for genotyping with both arrays with small differences and with an approximately 95% overlap of SNPs included. The 5% non-overlapping SNPs were in the Axiom array, which was an updated version of UK BiLEVE. Two of the 96 wells were used for controls (1,000 Genomes). Each batch included ∼4,700 samples. After quality control (QC) checks, less than 5% (∼38,000) of the SNPs in the Axiom array were set to missing, because they did not meet the QC standards. More information on the genotyping was provided in relevant reports from the UK Biobank[316]. Genotypes not assayed were imputed using the UK10K merged with the 1,000 Genomes Phase III reference panel and the IMPUTE2 programme, as described in more detail in a UK Biobank report[317].

8.4.3 Dietary assessment and data processing

Two different tools were applied for the dietary assessment in the UK Biobank. The first tool was used in the total sample and comprised a collection of frequency questions for main food groups and items included in the general touchscreen questionnaire along with other questions on e.g. socio-demographic factors, administered during the assessment visit. This questionnaire was filled in by the participants in all three repeated assessments (first: n≈500,000, second: n≈20,000, third: n≈8,000). The second tool was the Oxford WebQ, a hybrid of a 24-hour dietary recall and a 200-item food frequency questionnaire

(FFQ)[318] administered in a sub-sample of approximately 200,000 participants. The WebQ was administered in approximately the last 70,000 participants at the assessment centre at baseline. Then, it was sent via e-mail to approximately 320,000 participants who had provided an email address and were asked to fill it in four times distributed across the year 2011-2012 to capture seasonal variation and habitual diet. Participants were asked to complete the WebQ on specific days, so that within day variation and differences between weekdays and weekend days were captured. It was completed by approximately 100,000 participants in the first cycle (February 2011 - April 2011), 80,000 in the second (June 2011 - August 2011), 100,000 in the third (October 2011 - December 2011) and 100,000 participants in the fourth cycle (April 2012 - June 2012). The total number of participants, who completed the WebQ at least once was 200,000 and some of them completed it more than once up to five times. An overview of the administration of the touchscreen questionnaire and the Oxford WebQ is shown in **Figure 8.5** along with the number of participants with one or more WebQs available.

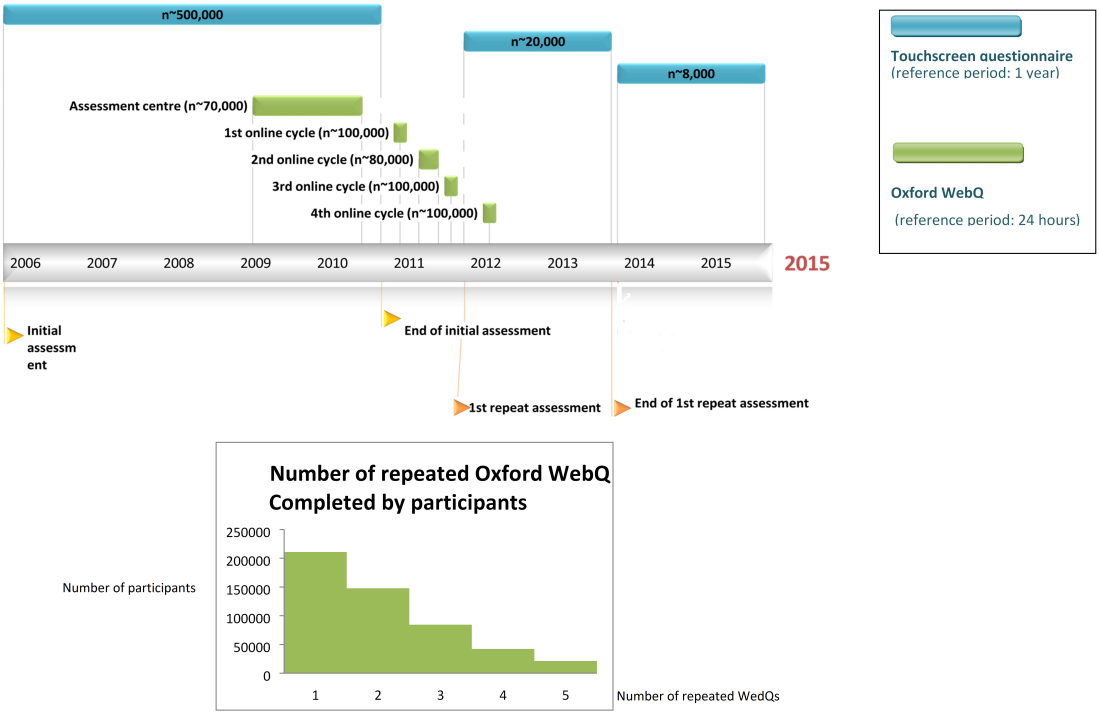


Fig. 8.5 Dietary assessment timeline in the UK Biobank

Dairy consumption from the touchscreen questionnaire

The questions from the touchscreen questionnaire related to dairy consumption were as follows:

- "Which of the following do you never eat?" with multiple choice answers including "Dairy products"
- "How often do you eat cheese? (include cheese in pizzas, quiches, cheese sauce etc)" with eight possible answers: "Never", "Less than once a week", "Once a week", "2-4 times a week", "5-6 times a week", "Once or more daily", "Do not know", "Prefer not to answer"
- "What type of milk do you mainly use?" with eight possible answers: "Full cream", "Semi-skimmed", "Skimmed", "Soya", "Other type of milk", "Never/rarely have milk", "Do not know", "Prefer not to answer"

Yoghurt consumption was not assessed in the touchscreen questionnaire.

Dairy consumption from the Oxford WebQ

Milk consumed with coffee, tea and cereals

Dairy products were assessed in various sections of the Oxford WebQ. For milk, we used milk consumed with coffee, tea and cereals following the algorithms shown in **Figures 8.6-8.8** and under the assumptions listed in **Table 8.2** and also as a beverage. The primary milk phenotype was milk consumed as a drink and added to cereals.

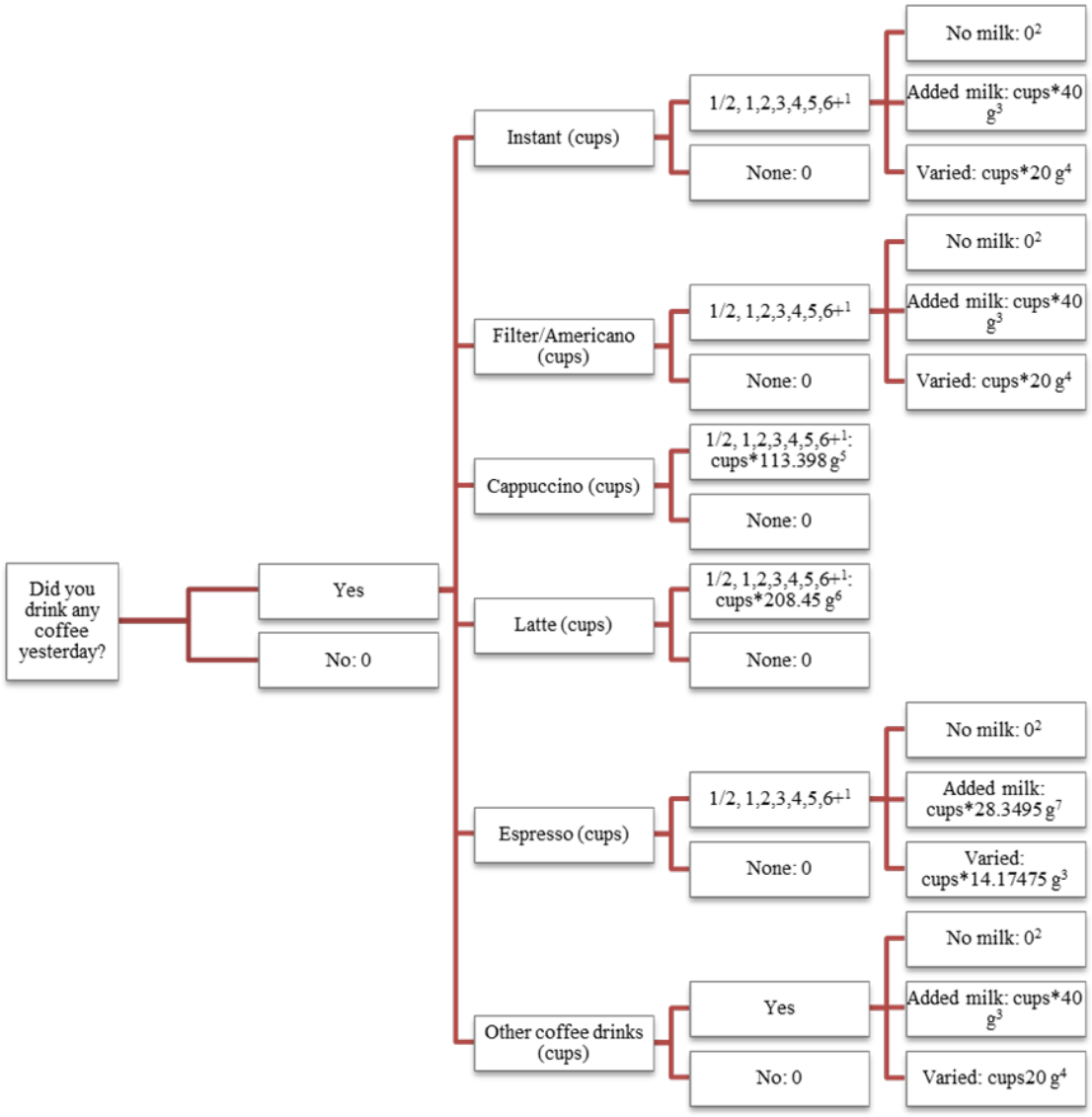


Fig. 8.6 Algorithm for the estimation of milk consumption from coffee as assessed from Oxford WebQ. Superscript numbers refer to numbered assumptions of Table 8.2

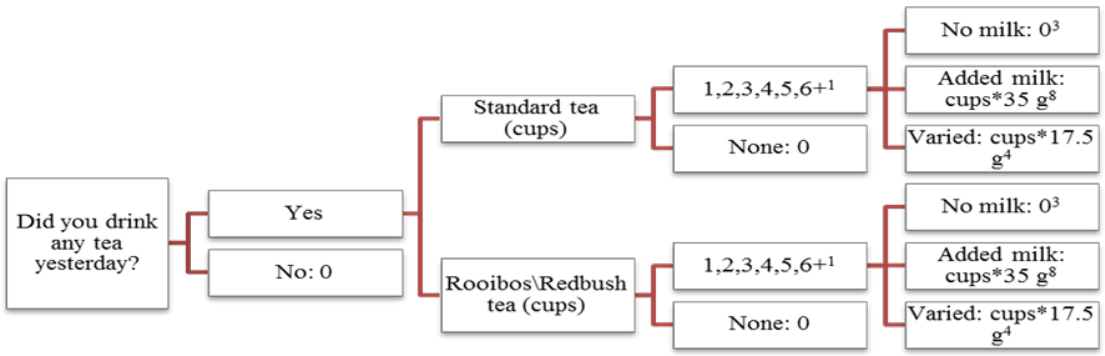


Fig. 8.7 Algorithm for the estimation of milk consumption from tea as assessed from Oxford WebQ. Superscript numbers refer to numbered assumptions of Table 8.2

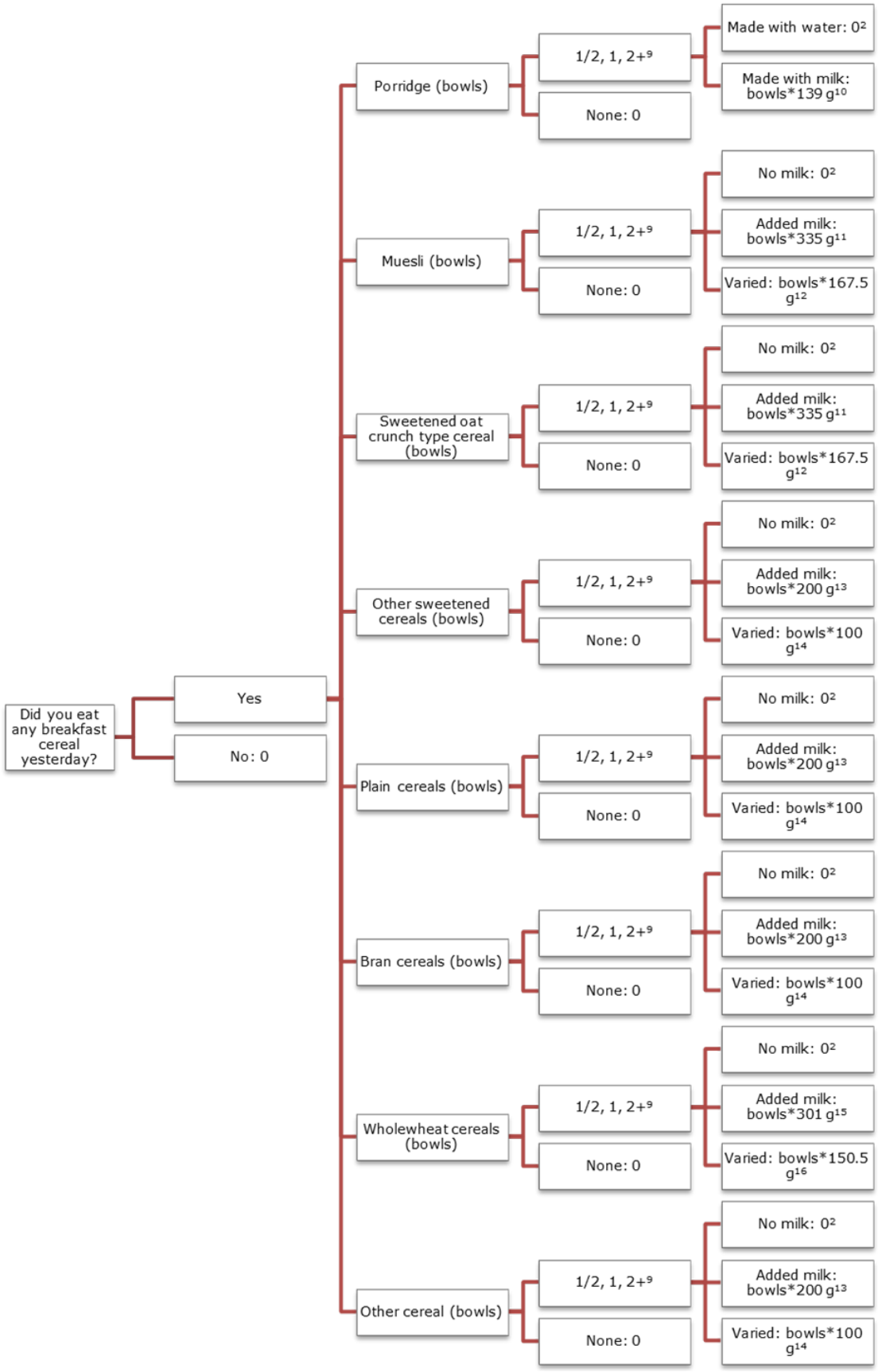


Fig. 8.8 Algorithm for the estimation of milk consumption from cereals as assessed from Oxford WebQ. Superscript numbers refer to numbered assumptions of Table 8.2

Table 8.2 Assumptions for the estimation of milk added to coffee, tea and cereals as assessed from the Oxford WebQ [318] in the UK Biobank

Assumption #	Assumption	Comments
1	6 cups	If the participant chose 6+ cups, it is assumed that 6 cups were consumed
2	0 milk added	If the first question on whether coffee, tea or cereals were consumed yesterday is not missing, but one of the coffee, tea or cereal types are missing, then the milk added to this type is set to 0, because none was coded as missing
3	40 g milk/cup of instant, filter or other coffee drinks	Average amount of milk added in one cup of coffee from 271 participants of the National Diet and Nutrition Survey Rolling Programme (Years 1-4)
4	20 g milk / cup of instant, filter or other coffee drinks	If the participant chose the option "Varied" instead of "Yes" on whether they added milk to their coffee, it is assumed that they added half the amount assumed that they added if they responded "Yes" i.e. 40 g/2 = 20 g
5	113.398 g of milk / cup of cappuccino	Cappuccino is known to have more milk than that added to instant or filtered coffee, so another assumption needed to be made. The relevant information was extracted from the following website: https://nationalcoffeeblog.org/2015/12/08/espresso-field-guide/
6	208.45 g / cup of latte	As for cappuccino, information on the amount of milk added in latte was extracted from the website https://nationalcoffeeblog.org/2015/12/08/espresso-field-guide/ with the additional assumption that one mug of latte is 250 ml, as the 340 ml suggested by the website is probably too big for the UK standards
7	28.3495 g / cup of espresso	Information was extracted from the website https://nationalcoffeeblog.org/2015/12/08/espresso-field-guide/
8	35 g milk/cup of tea	Average amount of milk added in one cup of tea from 1,883 participants of the National Diet and Nutrition Survey Rolling Programme (Years 1-4)
9	2 bowls	If the participant chose 2+ bowls, it is assumed that 2 bowls were consumed
10	139 g milk/ bowl of porridge	According to the porridge recipe from the McCance and Widdowson's food composition tables[35], 500 g of porridge include 60 g oatmeal, 7 g salt and 500 ml milk with a 14% weight loss after processing. Assuming the weight loss affects only the contribution of weight of milk to the total weight: Final milk amount=500 ml - 500 ml*0.14 loss = 500 - 70 ml = 430 ml, so the final contribution of milk to the total weight will be 430 / (430 +60+7) = 430/497 = 0.87. According to the Food Standards Agency (FSA)[197], one serving of porridge is 160 g, so in this serving, the amount of milk will be 160*0.87 = 139.2 g $\approx$ 139 g
11	335 g milk / bowl of muesli or sweetened oat crunch type cereals	Based on the recipe for porridge mentioned in assumption 10, we assumed that the contribution of milk to other types of cereals would be the same, so total amount = amount of cereals + total amount*0.87 => total amount - total amount*0.87 = amount of cereals => total amount*0.13 = amount of cereals => total amount = amount of cereals / 0.13, so the amount of milk will be: milk = 0.87*amount of cereals/0.13 => amount of milk = 6.69*amount of cereals. According to the FSA[197], one serving of muesli is 50 g, so the milk included in one serving of muesli is 50 g*6.69 = 334.5 g $\approx$ 335 g. It is assumed that one serving of sweetened oat crunch type cereals has the same weight as one serving of muesli
12	167.5 g milk / bowl of muesli or sweetened oat crunch type cereals	If the participant chose the option "Varied" instead of "Yes" on whether they added milk to their cereals it is assumed that they added half the amount assumed that they added if they responded "Yes" i.e. 335 g/2 = 167.5 g
13	200 g milk / bowl of other sweetened cereals or plain cereals or bran flakes or other cereals	Based on the derived equation amount of milk = 6.69*amount of cereals described in assumption 11 and on the described serving of 30 g of other sweetened cereals or plain cereals or bran flakes or other cereals according to the FSA[197], the amount of milk in one serving of these cereals should be 30 g * 6.69 = 200.7 g $\approx$ 200 g
14	100 g milk / bowl of other sweetened cereals or plain cereals or bran flakes or other cereals	If the participant chose the option "Varied" instead of "Yes" on whether they added milk to their cereals it is assumed that they added half the amount assumed that they added if they responded "Yes" i.e. 200 g/2 = 100 g
15	301 g milk / bowl of whole-wheat cereals	Based on the derived equation amount of milk = 6.69*amount of cereals described in assumption 11 and on the described serving of 45 g of whole-wheat cereals according to the FSA[197], the amount of milk in one serving of whole-wheat cereals should be 45 g * 6.69 = 301.05 g $\approx$ 301 g
16	150.5 g milk / bowl of whole-wheat cereals	If the participant chose the option "Varied" instead of "Yes" on whether they added milk to their cereals it is assumed that they added half the amount assumed that they added if they responded "Yes" i.e. 301 g/2 = 1505 g

Milk consumed as a beverage

For milk consumed as a beverage, participants could choose one of the options ranging from None, $\frac{1}{2}$,1,2,3,4,5 to 6+ glasses of milk. If the participant chose the option 6+, it was assumed that they consumed six glasses of milk. In a separate question, participants were asked "Which type of milk did you use most frequently yesterday?". They had the option to choose one of the following:

- I did not have any type of milk or milk substitute yesterday
- Semi-skimmed cow's milk
- Skimmed cow's milk
- Whole (full cream) cow's milk
- Cholesterol lowering milk e.g. Flora pro.active
- Soya milk with added calcium
- Soya milk without added calcium
- Goat's or sheep's milk
- Rice, oat milk or other vegetable milk e.g. Rice Dream, Plamil
- Powdered milk
- I do not know which type of milk I used the most
- Other type of milk

If the participant chose any type of milk not compatible with the biological definition provided by Codex Alimentarius as described in Chapter 1 such as soya milk with or without calcium, rice, oat milk or other vegetable milk or "other type of milk", the milk consumption was replaced with zero. This was done so that the associations between SNPs related to LP and milk consumption would not be obscured. "Other type of milk" was also replaced with zero, because processed milk low in lactose was not included in the options, so it was assumed that there is a higher possibility that such a type of milk would be reported as "Other type". In this way, we tried to minimise missclassification, since currently the data do not make it possible to examine any specific answers of the participants, when they would choose "other type of milk".

Yoghurt consumption

Yoghurt was assessed with the question "Did you eat any yoghurt or ice-cream yesterday?". If the participant replied "Yes", they could choose one of the options ranging from None, $\frac{1}{2}$, 1, 2 to 3+ servings/individual pots of yoghurt. If the participant chose the option 3+, it was assumed that they consumed three servings of yoghurt. One serving of yoghurt was defined as 125 g according to the Food Standards Agency (FSA)[197].

Cheese consumption

Cheese consumption was assessed with the question "Did you eat any cheese yesterday?". If the participant replied "Yes", they could choose the type of cheese they consumed from the following:

- Low fat hard cheese (e.g Edam, reduced fat Cheddar)
- Hard cheese (e.g. Cheddar, Parmesan)
- Soft cheese (e.g. Brie)
- Blue cheese (e.g. Stilton)
- Low fat spreadable cheese
- Spreadable cheese (e.g. cream cheese)
- Cottage cheese
- Feta
- Mozzarella
- Goat's cheese
- Other cheese

and the quantity options ranging from None, $\frac{1}{2}$, 1, 2 to 3+ servings. If the participants chose the option 3+, it was assumed that they consumed three servings. One serving of cheese was assumed to correspond to 40 g as suggested by the FSA[197].

8.4.4 Statistical analysis

Derivation of dairy outcomes

In order to make the most of the available data, we derived the dairy variables by combining information from both tools when possible. In this way, we combined the advantages of both tools by using the larger sample size that the touchscreen questionnaire was administered to and the greater detail available in the Oxford WebQ.

Continuous dairy variables from the Oxford WebQ were skewed. Log-transformation and Box-Cox transformation were applied to adjust the distribution of the variables closer to the normal. Distributions of milk, yoghurt and cheese variables did not approximate the normal even after the transformations, so the binary variables of consumers and non-consumers were used.

We created three binary variables of milk by splitting the participants into consumers and non-consumers. The first variable included only milk consumed as a drink from the Oxford WebQ. The second variable was derived from dichotomisation of the sum of milk consumed as a drink and milk from coffee, tea or cereals from the Oxford WebQ. In both cases, participants were considered consumers if they reported milk consumption

in at least one repeated assessment. The third variable was derived from the touchscreen questionnaire classifying as consumers those who chose either "Full cream", "Semi-skimmed" or "Skimmed" from the relevant question in at least one repeated assessment. This variable was updated with the answers from the Oxford WebQ. Specifically, we re-classified the touchscreen-defined non-consumers into consumers if they had reported milk consumption in the Oxford WebQ.

Yoghurt consumers were defined as those who reported yoghurt consumption at least in one repeated assessment with the Oxford WebQ.

A binary variable for cheese consumption was created from the information provided with the touchscreen questionnaire classifying as consumers those who reported cheese consumption in at least one repeated assessment with either the touchscreen questionnaire or the Oxford WebQ.

Total dairy consumption was estimated as the weighted average of the sum of milk, yoghurt and cheese from the repeated measures of Oxford WebQ accounting for weekdays and weekends. The weighted average was then winsorised to the 5^{th} and 95^{th} percentiles and was Box-Cox transformed. Due to the lack of information on yoghurt, the touchscreen questionnaire was not considered for the estimation of total dairy consumption.

The primary dairy outcomes were milk consumed as a drink and added to cereals (Oxford WebQ), yoghurt (Oxford WebQ), cheese (touchscreen questionnaire and Oxford WebQ) and total dairy consumption (Oxford WebQ).

Genetic analyses

Continuous variables for SNPs were used with values ranging from 0 to 2 depending on heterozygosity (value of 1) or homozygosity for the effect (value of 2) or the alternative allele (value of 0). Decimal points indicated genotype expectations for imputed SNPs.

Positive control analysis

Before conducting the GWAS, the association of the SNP rs4988235 with milk and total dairy consumption was assessed as a positive control, because rs4988235 is the most well characterised causal SNP for LP in white populations. In this positive-control candidate-gene analysis, we fitted logistic regression models adjusted for age, sex and the first 10 principal components to account for population stratification both in the total sample and after exclusion of non-white participants. Data processing and positive control analyses were performed in Stata version 14.2 (College Station, TX: StataCorp LP, 2015).

Genome-wide association studies

GWASs were performed on the six dairy phenotypes using an established pipeline developed by members of the MRC Epidemiology Unit genetics group. Due to the large data size, high performance computing (HPC) was used. Linear mixed regression models were developed for the GWAS of all the phenotypes in BOLT-LMM version 2.3.2 software[319], which does not include logistic regression models. Models were adjusted

for age and sex. Non-white participants were excluded. BOLT-LMM uses a relationship matrix that controls for both population stratification and relatedness.

Biological interpretation

SNP GWAS hits with $p<0.01$ were presented in Manhattan plots generated with the qqman package in R. Information on the annotation and related genes of the top SNP hits was extracted from the HaploReg v4.1 database[320]. Top hits were defined as the SNPs which reached genome-wide significance ($p<5x10^{-8}$) after excluding variants with minor allele frequency less than 0.01 and imputation quality less than 0.3 and pruning within 250 kb from the SNP with the strongest association. Information on the relevant genes was searched for in the Gene NCBI (National Center for Biotechnology Information) database (https://www.ncbi.nlm.nih.gov/gene/). Phenotypes, which have previously been associated with our top signals in the UK Biobank, were identified from the Oxford Brain Imaging Genetics (BIG) database (http://big.stats.ox.ac.uk/). Information on linkage disequilibrium between the lactase persistence SNP rs4988235 and other SNPs of chromosome 2 was extracted from SNiPA v3.3[321].

8.5 Results

8.5.1 Description of dairy outcomes

Starting from 502,628 participants, after exclusion of participants with no genetic data (n=15,219), no overall dietary data (n=14), non-white participants (n=35,095) and participants with missing phenotypic information, the number of participants used in the analyses of data from the touchscreen questionnaire were 451,900 for milk consumed as a drink and added to cereals and 445,330 for cheese. After further excluding participants with no dairy information from the Oxford WebQ, the number of participants further reduced to 193,505 for analyses on total dairy products, milk consumed as a drink, milk consumed as a drink or added to coffee, tea or cereals and yoghurt (**Table 8.3**). The majority of the participants were consumers of milk added to coffee, tea or cereals (90.2%), and cheese (97.1%), whereas for milk consumed as a beverage only 6.9% were consumers and for yoghurt almost half of the participants were consumers (46.7%; Table 8.3).

8.5.2 Positive control analysis

Results from the positive control analysis indicated highly significant positive associations between the LP SNP rs4988235 and both the milk phenotypes and total dairy consumption (**Table 8.4**). The strongest associations were observed from the logistic regression model on milk consumed as a drink. In this case, β=0.11 could be interpreted as a higher likelihood of being a milk consumer by 0.11 times for each 1% higher probability of having the effect allele for LP. The likelihood was reduced to 0.06 times for milk consumed

Table 8.3 Number of consumers and non-consumers for binary dairy phenotypes and mean±SD for the continuous dairy phenotypes † in the UK Biobank subsets used for the genome-wide association studies (GWAS)

Dairy phenotypes	Total set N	Subsets Consumers	Non-consumers	Data set
Total dairy products	193,505	1.9±0.7		Oxford WebQ
Milk as a drink and added to cereals	451,900	92.9	7.1	Total ‡
Milk as drink and added to coffee, tea and cereals	193,505	90.2	9.8	Oxford WebQ
Milk as drink	193,505	6.9	93.1	Oxford WebQ
Yoghurt	193,505	46.7	53.3	Oxford WebQ
Cheese	445,330	97.1	2.9	Total ‡

† Only total dairy consumption was used as a continuous phenotype, so the mean±SD is reported for that, whereas for the rest of the phenotypes the percentage of participants within each category of consumers and non-consumers is reported
‡ Information combined from the touchscreen questionnaire and the Oxford WebQ

as a drink and added to coffee, tea and cereals. For total dairy products, β=0.03 for each effect allele of lactase persistence, which is not directly interpretable due to the Box-Cox transformation (Table 8.4).

Table 8.4 Positive control results for the associations of the lactase persistence single nucleotide polymorphism (rs4988235) with milk and total dairy consumption in subsets of all ethnicities or white people only in the UK Biobank

Dairy phenotype	Model	N	b	se	*p*
Total dairy products †	All ethnicities	206,381	0.03	$1.3x10^{-5}$	$9.0x10^{-20}$
	White	193,505	0.03	$1.4x10^{-5}$	$1.1x10^{-15}$
Milk as a drink and added to cereals ‡	All ethnicities	486,045	0.08	$8.5x10^{-5}$	$2.9x10^{-17}$
	White	451,833	0.07	$9x10^{-5}$	$2.3x10^{-12}$
Milk as drink and added to coffee, tea and cereals ‡	All ethnicities	206,381	0.07	0.0001	$1.3x10^{-8}$
	White	193,505	0.06	0.0002	$1.8x10^{-6}$
Milk as drink ‡	All ethnicities	206,381	0.11	0.0002	$4.6x10^{-14}$
	White	193,505	0.10	0.0002	$1.2x10^{-11}$

† Total dairy products were included in the analysis as a continuous variable in servings/day, so the effect estimates are derived from linear regression models. Models were adjusted for age, sex and 10 principal components
‡ The different phenotypes for milk were included in the analysis as binary variables of consumers and non-consumers of milk, so the effect estimates are derived from logistic regression models. Models were adjusted for age, sex and 10 principal components.

8.5.3 Genome-wide association studies

Manhattan plots for total dairy products (Oxford WebQ) and the two dairy outcomes derived from the combination of the touchscreen and Oxford WebQ questionnaires are presented in **Figure 8.9**. Manhattan plots for the rest of the dairy outcomes derived from the Oxford WebQ are presented in **Figure 8.10**. The most significant hits from the GWAS ($p<5x10^{-8}$) were observed for total dairy consumption within the subset of 193,505 participants with WebQ data, located on chromosomes 1, 2, 4 and 12 (p range: 8.6×10^{-16} - 1×10^{-8}; Figure 8.9). For the main milk phenotype (consumed as a drink or added to

cereals) significant hits were located on chromosome 2 (p range: 3.8 x 10^{-13} - 2 x 10^{-10}; Figure 8.9). The same pattern was observed for milk consumed as drink (p range: 5.7 x 10^{-12} - 2.3 x 10^{-10}; Figure 8.10). SNPs for milk added to coffee, tea or cereals did not reach genome-wide significance (Figure 8.10). Significant hits for cheese were located on chromosomes 1, 2 and 7 (p range: 4.4 x 10^{-9} - 1.2 x 10^{-8};Figure 8.9), while for yoghurt only one SNP on chromosome 1 reached genome-wide significance (p= 1.4 x 10^{-8}; Figure 8.10).

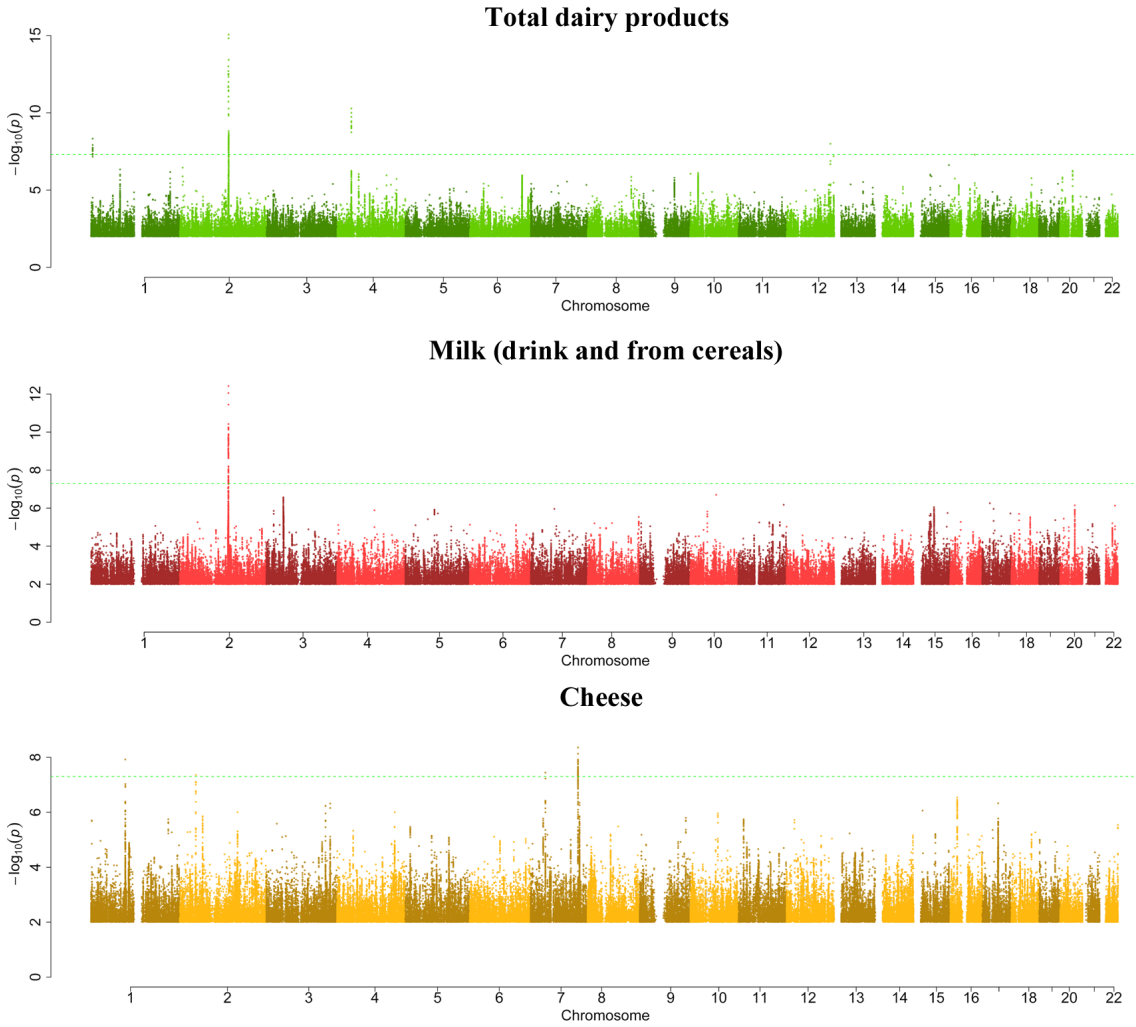


Fig. 8.9 Manhattan plots of the negative logarithm of the genome-wide association studies p values for SNPs with p<0.01 for total dairy products (continuous phenotype; Oxford WebQ subset, N=193,505; green plot), milk consumed as drink or added to cereals (binary phenotype; N=451,900; red plot) and cheese (binary phenotype; N=445,330; yellow plot) in the UK Biobank. The dotted line indicates the genome-wide significance threshold of 5x10^{-8}. GWAS: Genome-wide association study; SNP: Single nucleotide polymorphism

Information of SNPs producing p<5x10^{-8} is presented in **Table 8.5**. Seven SNPs were identified as top signals for total dairy products, of which five were intronic; four intronic SNPs were identified for milk consumed as a drink or added to cereals; four intronic SNPs

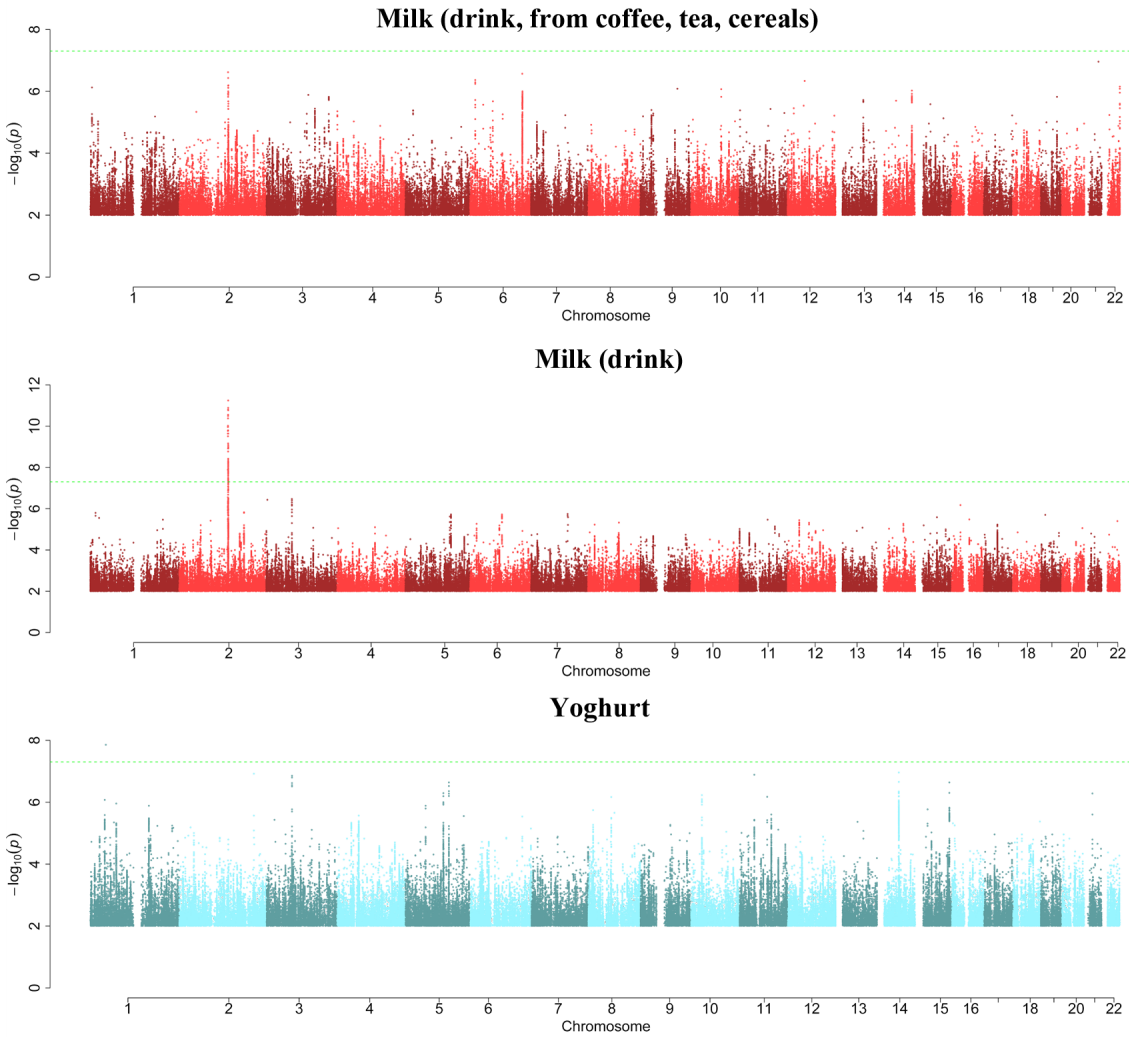


Fig. 8.10 Manhattan plots of the negative logarithm of the genome-wide association studies p values for SNPs with p<0.01 for milk consumed as drink or added to coffee, tea or cereals (binary phenotype; N=193,505; top red plot), milk consumed as a drink only (binary phenotype; N=193,505; bottom red plot) and yoghurt (binary phenotype; N=193,505; blue plot) in the UK Biobank Oxford WebQ subset. The dotted line indicates the genome-wide significance threshold of $5x10^{-8}$. GWAS: Genome-wide association study; SNP: Single nucleotide polymorphism

were identified for milk consumed as a drink; one SNP was identified for yoghurt; and four SNPs were identified for cheese, of which one was intronic.

To examine the specificity of the SNPs for each of the main dairy types, the associations between SNPs, which were identified as top hits for at least one dairy type, and all the dairy types are presented in **Table A.34**. Although some of the top hits for the milk phenotypes overlap with those for total dairy products and other milk phenotypes, the top hits of the three main dairy types i.e. milk, yoghurt and cheese do not overlap with each other, which makes them specific for the corresponding dairy types.

Table 8.5 Single nucleotide polymorphisms (SNPs) significantly associated with milk, yoghurt, cheese and total dairy consumption from genome-wide association studies in the UK Biobank †

Dairy phenotype	SNP	Chromosome	Functional annotation ‡	Nearest gene ‡	Base position	Effect allele	Minor allele	EAF	INFO §	b ‖	se	*p*
Total dairy products #	rs12409187	1	intronic	*AJAP1*	4733793	G	A	0.95	0.96	0.04	0.01	$4.7x10^{-9}$
	rs7570971	2	intronic	*RAB3GAP1*	135837906	C	A	0.72	1.00	0.03	0.004	$9.8x10^{-14}$
	rs3940549	2	intronic	*ZRANB3*	136138627	A	G	0.72	0.98	0.03	0.004	$2x10^{-13}$
	rs4988235	2	intronic	*MCM6*	136608646	G	A	0.26	0.97	-0.03	0.004	$8.6x10^{-16}$
	rs28815269	2	none	104kb 5' of *CXCR4*	136979530	G	A	0.35	0.96	-0.02	0.003	$2x10^{-9}$
	rs11940694	4	intronic	*KLB*	39414993	A	G	0.40	0.98	0.02	0.003	$5.2x10^{-11}$
	rs35754956	12	none	4.6kb 3' of *AC156455.1*	122511116	A	G	0.87	0.99	0.03	0.005	$1x10^{-8}$
Milk as a drink and added to cereals *	rs7570971	2	intronic	*DARS*	135837906	C	A	0.73	1.00	0.004	0.001	$2x10^{-10}$
	rs3940549	2	intronic	*R3HDM1*	136138627	A	G	0.73	0.98	0.004	0.001	$9.3x10^{-10}$
	rs62168795	2	intronic	*RAB3GAP1*	136429366	T	C	0.75	0.95	0.004	0.001	$6x10^{-11}$
	rs6754311	2	intronic	*ZRANB3*	136707982	T	C	0.75	0.97	0.005	0.001	$3.8x10^{-13}$
Milk as drink *	rs7570971	2	intronic	*MCM6*	135837906	C	A	0.72	1.00	0.01	0.001	$2.3x10^{-10}$
	rs151022760	2	intronic	*R3HDM1*	136098560	T	TATTG	0.73	0.98	0.01	0.001	$9.6x10^{-11}$
	rs74775210	2	intronic	*ZRANB3*	136352327	T	TTC	0.73	0.98	0.01	0.001	$2.8x10^{-11}$
	rs182549	2	intronic	*RAB3GAP1*	136616754	C	T	0.26	0.98	-0.01	0.001	$5.7x10^{-12}$
Yoghurt *	rs35025768	1	none	5.5kb 3' of U6	43483707	G	A	0.98	0.86	0.04	0.01	$1.4x10^{-8}$
Cheese *	rs1222762	1	none	19kb 5' of Metazoa_SRP	97029325	G	A	0.43	1.00	-0.002	0.0004	$1.2x10^{-8}$
	rs504764	2	none	*RP11*-89K21.1	45154662	C	T	0.67	0.99	-0.002	0.0004	$4.4x10^{-8}$
	rs10245608	7	intronic	*CHCHD3*	41241212	C	T	0.51	0.99	-0.002	0.0004	$3.6x10^{-8}$
	rs10264126	7	none	68kb 3' of *AC005022.1*	132620098	C	T	0.74	1.00	0.002	0.0004	$4.4x10^{-9}$

†Selection of top hits was based on a statistical significance passing the genome-wide significance threshold of $p=5x10^{-8}$ and after pruning of hits within 500 kb from the SNP with the strongest associations

‡Information was extracted from HaploReg v4.1[320]

§Imputation score for the assessment of the quality of imputation done in IMPUTE2 programme. SNPs with a score <0.4 were dropped from the analyses

‖Beta coefficient and standard error as derived from BOLT-LMM software, which uses linear regression models

#A continuous phenotype was used for total dairy products in servings/day

*Categorical phenotypes were used for milk, yoghurt and cheese splitting the sample into consumers and non-consumers

Abbreviations: SNP: single nucleotide polymorphisms, EAF: effect allele frequency

Information extracted for top SNP hits and the corresponding genes is provided in **Table 8.6**. Most of the SNPs are described as intronic, which means that they are involved in protein expression and that they might be of regulatory importance. In addition, genes located on chromosome 2, which included top SNP hits for total dairy consumption and milk consumption (as a drink only or as a drink and added to cereals) are in close proximity with each other, so there is a higher probability that they are in linkage disequilibrium and are highly correlated. Indeed, for example, the LP SNP rs4988235 is in high linkage disequilibrium ($r^2 \geq 0.8$) with SNPs from the genes *DARS*, *MCM6* and *R3HDM1* (**Figure 8.11**).

Table 8.6 Functional annotation and corresponding genes for the genome-wide association study (GWAS) top single nucleotide polymorphism (SNP) hits of dairy phenotypes in the UK Biobank †

Dairy phenotype	SNP	Chromosome	Gene position	Functional annotation	Nearest gene
Total dairy products	rs12409187	1	0.0979031	intronic	*AJAP1*
	rs4988235	2	1.58582	intronic	*MCM6*
	rs7570971	2	1.58509	intronic	*RAB3GAP1*
	rs3940549	2	1.58514	intronic	*ZRANB3*
	rs28815269	2	1.58915	none	104kb 5' of *CXCR4*
	rs11940694	4	0.613896	intronic	*KLB*
	rs35754956	12	1.46556	none	4.6kb 3' of *AC156455.1*
Milk as drink and added to cereals	rs6754311	2	1.58595	intronic	*DARS*
	rs62168795	2	1.58551	intronic	*R3HDM1*
	rs7570971	2	1.58509	intronic	*RAB3GAP1*
	rs3940549	2	1.58515	intronic	*ZRANB3*
Milk as drink	rs182549	2	1.58582	intronic	*MCM6*
	rs74775210	2	1.58521	intronic	*R3HDM1*
	rs151022760	2	1.58514	intronic	*ZRANB3*
	rs7570971	2	1.58509	intronic	*RAB3GAP1*
Yoghurt	rs35025768	1	0.707581	none	5.5kb 3' of U6
Cheese	rs1222762	1	1.26566	none	19kb 5' of Metazoa_SRP
	rs504764	2	0.697037	none	*RP11-89K21.1*
	rs10264126	7	1.44451	intronic	*CHCHD3*
	rs10245608	7	0.657781	none	68kb 3' of AC005022.1

†Information was extracted from HaploReg v4.1[320]
Abbreviations: GWAS: Genome-wide association study; SNP: Single nucleotide polymorphism

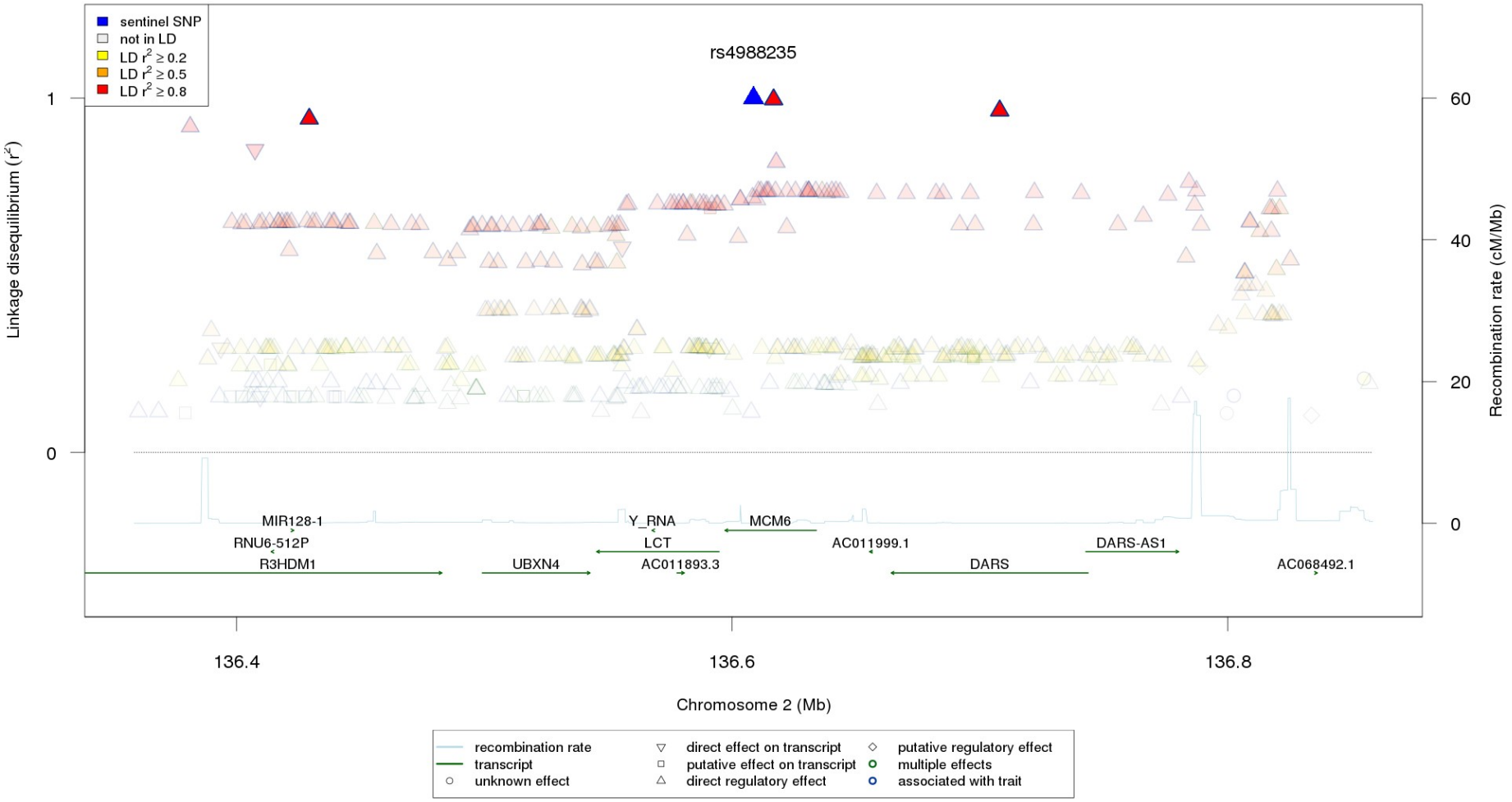


Fig. 8.11 Linkage disequilibrium (LD) of single nucleotide polymorphisms (SNPs) with the lead SNP, rs4988235, known to be associated with lactase persistence

8.5.4 Exploratory work on biological interpretation

Information on the genes that carry the top SNP hits for dairy types is presented in **Table 8.7**. The functional role of some of the identified genes (e.g. *AJAP1*, *ZRANB3*, *R3HDM1*, *RP11-89K21.1*, *KLB*) is not clear. The most relevant gene is *MCM6*, which contains the LP SNP rs4988235. No known information related to dairy consumption or dairy nutrients was available for the genes *RAB3GAP1*, *DARS*, *CHCHD3*.

Table 8.7 Information on genes, which include the top single nucleotide polymorphism (SNP) hits from the genome-wide association studies (GWAS) on the dairy phenotypes in the UK Biobank†

Nearest gene	Chromosome	SNPs	Dairy phenotypes	Gene summary	Gene expression
AJAP1	1	rs12409187	Total dairy products		Biased expression in brain (RPKM 4.3), ovary (RPKM 2.3) and 5 other tissues
5.5kb 3' of U6	1	rs35025768	Yoghurt		
19kb 5' of Metazoa_SRP	1	rs1222762	Cheese		
MCM6	2	rs4988235; rs182549	Total dairy products; Milk as drink	Lactase persistence	
RAB3GAP1	2	rs7570971	Total dairy products; Milk as drink and added to cereals; Milk as drink	Encodes the catalytic subunit of a Rab GTPase activating protein. Mutations in this gene are associated with Warburg micro syndrome. Alternate splicing results in multiple transcript variants.	Ubiquitous expression in brain (RPKM 16.8), thyroid (RPKM 15.9) and 25 other tissues
ZRANB3	2	rs3940549; rs151022760	Total dairy products; Milk as drink and added to cereals; Milk as drink		Broad expression in testis (RPKM 1.8), thyroid (RPKM 0.8) and 24 other tissues
DARS	2	rs6754311	Milk as drink and added to cereals	Next to *MCM6* gene; Encodes a member of a multienzyme complex that functions in mediating the attachment of amino acids to their cognate tRNAs. The encoded protein ligates L-aspartate to tRNA(Asp). Mutations in this gene have been found in patients showing hypomyelination with brainstem and spinal cord involvement and leg spasticity.	Ubiquitous expression in thyroid (RPKM 24.7), lymph node (RPKM 20.9) and 25 other tissues
R3HDM1	2	rs62168795; rs74775210	Milk as drink and added to cereals; Milk as drink		broad expression in brain (RPKM 20.1), testis (RPKM 10.5) and 23 other tissues
RP11-89K21.1	2	rs504764	Cheese		
104kb 5' of CXCR4	2	rs28815269	Total dairy products	Encodes a CXC chemokine receptor specific for stromal cell-derived factor-1. It acts with the CD4 protein to support HIV entry into cells and is also highly expressed in breast cancer cells. Mutations in this gene have been associated with WHIM (warts, hypogammaglobulinemia, infections, and myelokathexis) syndrome.	Biased expression in bone marrow (RPKM 479.2), lymph node (RPKM 201.0) and 9 other tissues
KLB	4	rs11940694	Total dairy products		Biased expression in fat (RPKM 16.3), liver (RPKM 5.1) and 2 other tissues

Table 8.7 (continued)

Nearest gene	Chromosome	SNPs	Dairy phenotypes	Gene summary	Gene expression
CHCHD3	7	rs10264126	Cheese	Encodes an inner mitochondrial membrane scaffold protein. Absence of the encoded protein affects the structural integrity of mitochondrial cristae and leads to reductions in ATP production, cell growth, and oxygen consumption. Several transcript variants encoding different isoforms have been found for this gene.	Ubiquitous expression in heart (RPKM 24.9), duodenum (RPKM 14.8) and 25 other tissues
68kb 3' of AC005022.1	7	rs10245608	Cheese		Low expression observed in reference dataset
4.6kb 3' of AC156455.1	12	rs35754956	Total dairy products		

†Infromation extracted from the Gene NCBI database (https://www.ncbi.nlm.nih.gov/gene/)

As shown in **Table 8.8**, phenome-wide association data from the UK Biobank extracted from the Oxford BIG database, indicated that body composition traits were associated with top signals of total dairy products (rs4988235, rs7570971), milk consumed as a drink or added to cereals (rs6754311, rs7570971), milk consumed as a drink (rs182549, rs7570971) and cheese (rs1222762). Sleep duration was associated with top signals for total dairy products (rs4988235, rs7570971), milk consumed as a drink or added to cereals (rs6754311, rs7570971), milk consumed as a drink (rs182549, rs7570971) and cheese (rs10264126). Alcohol consumption was associated with top signals of total dairy products (rs11940694) and cheese (rs504764), while smoking was associated with top signals of cheese (rs504764).

Table 8.8 Traits associated with the top SNP hits from phenome-wide association studies (PheWAS) in the UK Biobank†

Dairy phenotype	SNP	Chromosome	Nearest gene	PheWAS traits				
				Anthropometry / Body composition	Respiratory traits	Physical activity / sleep	Alcohol	Smoking	Other
Total dairy products	rs12409187	1	*AJAP1*						
	rs4988235	2	*MCM6*	Weight, Body fat percentage, Leg fat mass (left), Trunk fat percentage, Leg fat percentage (right), Leg fat mass (right), Leg fat percentage (left), Trunk fat mass, Whole body fat mass, Arm fat percentage (right), Arm fat mass (right), Arm fat mass (left), Body mass index (BMI), Arm fat percentage (left), Hand grip strength (left)	Forced vital capacity (FVC), Forced expiratory volume in 1-second (FEV1)	Sleep duration			Potassium in urine
	rs7570971	2	*RAB3GAP1*	Body fat percentage, Trunk fat percentage, Trunk fat mass, Leg fat mass (left), Leg fat mass (right), Whole body fat mass, Arm fat percentage (right), Arm fat mass (right), Arm fat mass (left), Leg fat percentage (right), Leg fat percentage (left), Arm fat percentage (left), BMI, Hand grip strength (left), Hand grip strength (right)	Forced vital capacity (FVC), Forced expiratory volume in 1-second (FEV1)	Sleep duration			
	rs3940549	2	*ZRANB3*						
	rs28815269	2	104kb 5' of *CXCR4*		Forced vital capacity (FVC), Forced expiratory volume in 1-second (FEV1)				
	rs11940694	4	KLB				Alcohol intake frequency, Reason for reducing amount of alcohol drunk: Health precaution, Alcohol intake versus 10 years previously, Average weekly beer plus cider intake		
	rs35754956	12	4.6kb 3' of *AC156455.1*	Standing height					
Milk consumed as a drink and added to cereals	rs6754311	2	DARS	Body fat percentage, Leg fat mass (left), Leg fat mass (right), Leg fat percentage (right), Trunk fat percentage, Trunk fat mass, Arm fat mass (right), Whole body fat mass, Leg fat percentage (left), Arm fat mass (left), Body mass index (BMI), Arm fat percentage (right), Arm fat percentage (left), Weight, Hand grip strength (left)	Forced vital capacity (FVC), Forced expiratory volume in 1-second (FEV1)	Sleep duration			potassium in urine
	rs62168795	2	*R3HDM1*						
	rs7570971	2	*RAB3GAP1*	Body fat percentage, Trunk fat percentage, Trunk fat mass, Leg fat mass (left), Leg fat mass (right), Whole body fat mass, Arm fat percentage (right), Arm fat mass (right), Arm fat mass (left), Leg fat percentage (right), Hand grip strength (left), Leg fat percentage (left), Arm fat percentage (left), BMI, Hand grip strength (right)	Forced vital capacity (FVC), Forced expiratory volume in 1-second (FEV1)	Sleep duration			
	rs3940549	2	*ZRANB3*						
Milk consumed as a drink	rs182549	2	*MCM6*	Body fat percentage, Leg fat mass (left), Trunk fat percentage, Leg fat mass (right), Leg fat percentage (right), Trunk fat mass, Whole body fat mass, Leg fat percentage (left), Arm fat mass (right), Arm fat mass (left), Arm fat percentage (right), Body mass index (BMI), Arm fat percentage (left), weight, Hand grip strength (left)	Forced vital capacity (FVC), Forced expiratory volume in 1-second (FEV1)	Sleep duration, Time spent watching television (TV)			Potassium in urine
	rs74775210	2	*R3HDM1*						
	rs151022760	2	*ZRANB3*						
	rs7570971	2	*RAB3GAP1*	Body fat percentage, Trunk fat percentage, Trunk fat mass, Leg fat mass (left), Leg fat mass (right), Whole body fat mass, Arm fat percentage (right), Arm fat mass (right), Arm fat mass (left), Leg fat percentage (right), Hand grip strength (left), Leg fat percentage (left), Arm fat percentage (left), BMI, Hand grip strength (right)	Forced vital capacity (FVC), Forced expiratory volume in 1-second (FEV1)	Sleep duration,			
Yoghurt	rs35025768	1	5.5kb 3' of U6						Treatment with prednisolone

Table 8.8

Dairy phenotype	SNP	Chromosome	Nearest gene	PheWAS traits					
				Anthropometry / Body composition	Respiratory traits	Physical activity / sleep	Alcohol	Smoking	Other
Cheese	rs1222762	1	19kb 5' of Metazoa_SRP	arm predicted mass (left), BMR, arm fat-free mass (left), whole body fat-free mass, whole body water mass, trunk fat-free mass, trunk predicted mass, arm fat-free mass (right), arm predicted mass (right), leg predicted mass (left), hip circumference, weight, leg fat-free mass (right), leg predicted mass (right),					treatment with olmesartan
	rs504764	2	*RP11-89K21.1*				alcohol	ever smoked, past tobacco smoking, pack years of smoking, current smoking	sodium in urine
	rs10264126	7	*CHCHD3*			sleep duration			
	rs10245608	7	68kb 3' of *AC005022.1*						

†Information extracted from the Oxford Brain Imaging Genetics (BIG) database (http://big.stats.ox.ac.uk/)

8.6 Discussion

8.6.1 Summary of results

In addition to the LP SNP rs4988235, in our GWASs we identified six SNPs in chromosomes 1, 2, 4 and 12 predictive of total dairy consumption, four SNPs in chromosome 2 predictive of milk consumed as a drink or added to cereals, four SNPs in chromosome 2 predictive of milk consumed as a drink, one SNP in chromosome 1 predictive of yoghurt consumption and four SNPs in chromosomes 1, 2 and 7 predictive of cheese consumption. Most of the SNPs identified are intronic, so probably of regulatory importance. Some of the SNPs have been associated with body composition traits, sleep duration, alcohol consumption and smoking in a PheWAS conducted previously in the UK Biobank.

8.6.2 Interpretation and previous evidence

Many of the SNPs that we identified as predictors of dairy consumption in our GWAS are in close proximity to the LP SNP and fall within the genetic region that has been causally associated with LP. Further steps that are required to make any causal inferences such as fine mapping of the genetic loci, replication in other studies and further exploration of potential confounders, were outside the scope and the timeline of this PhD project. However, as previously reported, the possibility of multiple causal genetic variants is in agreement with the long haplotype observed in relation to LP[305]. For example, the SNP rs182549, which we identified as one of the top signals for milk consumption, was previously associated with LP[305].

We did not identify any other studies reporting results of GWAS on dairy consumption. However, there are some studies, which investigated associations between selected SNPs and dairy consumption. Data from the PREDIMED-Valencia study showed that the SNP rs1466113 in the somatostatic receptor 2 (*SSTR2*) gene, which was strongly associated with BMI, was also associated with lower dairy consumption by 44 g/day for the homozygous compared with the heterozygous genotype[322]. The association between rs1466113 and dairy consumption was not significant in our study. Associations were also significant for meat and total protein intake and the authors suggested that it is difficult to disentangle the specific pathways, because somatostatin is involved in many pathways related to the regulation of appetite[322]. The SNP rs4788099 in the gene *SH2B1* located in chromosome 16, which was associated with the risk of obesity, was also associated with dairy consumption in the Look AHEAD trial[323]. The association between rs4788099 and dairy consumption was also significant in our study, although it did not reach genome-wide significance (p=0.003). Pirastu et al. investigated associations between genes related to taste and food preferences in populations from five countries from Eastern Europe and Western Asia[324]. In this study, rs2229642 located in the gene *ITPR3* of chromosome

6 was significantly associated with liking of sheep-origin cheese[324]. The association between rs2229642 and cheese consumption was not significant in our study. A more recent study including data from the study by Pirastu et al, but also data from an Italian and a Dutch population, reported significant associations of rs12994253 on chromosome 2 with blue cheese liking and rs4239891 of the *IGL* gene on chromosome 22 with plain yoghurt liking[325]. Such associations were not significant in our GWAS on cheese and yoghurt consumption respectively.

Genetic predictors of candidate biomarkers of dairy consumption i.e. the odd-chain fatty acids C15:0 and C17:0, could also elucidate potential biological mechanisms related to dairy consumption. In a GWAS on odd-chain fatty acids conducted from the CHARGE consortium, no significant hits were observed for C15:0, whereas rs13361131 in the gene *MYO10* of chromosome 5 was significantly associated with C17:0[326]. The link between this gene and C17:0 is not clear, but C17:0 can also be produced from alternative endogenous pathways[208], apart from the uptake from dairy consumption. The association between rs13361131 and dairy consumption was not significant in our study, so assuming that the association observed in the study by the CHARGE consortium is causal, this SNP might be linked to the alternative endogenous pathway of C17:0 production.

The failure to replicate most of the associations reported above in our study might be partly due to differences in the research questions. For example, food preference is not an identical phenotype to dairy consumption, even though it is one of the potential pathways of genetic predisposition to dairy consumption apart from lactose intolerance. In addition, the populations used in these studies included Western Asians, who might have a different genotype compared to our population of white Europeans. Furthermore, results from the GWAS on C17:0 might reflect pathways of endogenous production of C17:0.

The expression of the gene *R3HDM1*, which includes two of the SNPs identified as top genetic predictors of milk consumption in our GWAS, was lower among patients with coeliac disease[327]. Secondary lactose intolerance may result from coeliac disease, so the two conditions might share some common mechanisms[328] and our observation of associations between two SNPs in the gene *R3HDM1* and milk consumption might be related to this pathway, under the assumption that the associations were causal.

8.6.3 Strengths and limitations

To our knowledge, this is the largest and only (reported) GWAS on total and the main types of dairy consumption. The key strength of our study was its large sample size, enabling us to detect associations of small magnitude. While a large sample size can also be achieved through consortia data or meta-analyses, we evaluated data from a single study of participants with similar ancestry residing in the same country and thus exposed to the same culture. In addition, all the data went through the same procedures of collection and processing. These are two sources of heterogeneity often identified in meta-analyses

of different studies and could create more noise in the effect estimates. We overcame this problem by evaluating data from the UK Biobank. Finally, we used the main types of dairy products i.e. milk, yoghurt and cheese as phenotypes in the GWAS to attempt an identification of genetic predictors specific to each dairy type. Such predictors could be used as instrumental variables in Mendelian randomisation studies on associations between specific dairy types and disease, which are currently not available apart from milk consumption, genetically predicted by the LP SNP.

The main limitation of this work relates to dietary assessment. The dairy phenotypes were derived from information from the touchscreen questionnaire in the total sample (n≈450,000), which did not assess milk consumption quantitatively and did not assess yoghurt consumption at all. For cheese, though the question included reporting of average frequency of consumption, it could not capture different forms of cheese with different serving sizes. The use of the Oxford WebQ in the subset (n≈200,000) provided greater detail compared with the touchscreen questionnaire, but the dietary assessment was based on a single 24-hour recall in the form of a questionnaire with an a priori list of foods consumed. Due to difficulty in adjusting the dairy distributions to approximate the normal, we used binary phenotypes for dairy types, which might have resulted in loss of information. Although the measurement error of self-reported methods of dietary assessment is a common limitation in nutritional epidemiology and in our study, these methods are still an important source of information in observational studies. For reasons of feasibility, cost and participant burden, there is usually a trade-off between the large sample size and the amount of detail that the methods used provide. For example, although 7-day diaries might be more accurate and provide a greater level of detail, they are not feasible to be used in so large studies as the UK Biobank. The use of the Oxford WebQ, which has been validated against a single interviewer-administered 24-hour recall[318] is an example of this trade-off.

8.6.4 Implications and recommendations for future research

The identification of genetic predictors of dairy consumption is important for three reasons: to elucidate any genetic predisposition to dairy consumption, which might potentially further inform clinical and public health interventions; to disentangle potential genetic pathways linked to biological pathways underpinning associations between dairy consumption and disease; and to investigate causal associations between dairy consumption and disease in Mendelian randomisation analyses.

The replication of our results in other studies with more detailed methods of dietary assessment e.g. 7-day food diaries and other populations is important to confirm the prediction of dairy consumption from the SNPs we identified. The next step would be the investigation of the SNPs that are causally related to dairy consumption, which can be done with fine mapping and further exploration of sources of confounding and pleiotropy

of effects. Specifically, some of the SNPs we identified as predictors of dairy consumption were previously associated with traits related to anthropometry, respiratory function, physical activity, alcohol and sleep in the UK Biobank. It is not clear whether the interplay between these SNPs, dairy consumption and the traits identified in the phenome-wide association study is a result of pleiotropy of effects, confounding or causal associations. These possibilities should be addressed and further investigated in future studies to assess the suitability of the identified SNPs as instrumental variables in Mendelian randomisation analyses. Finally, the suitable SNPs can be used individually or as parts of genetic risk scores in Mendelian randomisation analyses to investigate causal associations between dairy consumption and disease.

8.6.5 Conclusion

We identified novel SNPs predictive of total and types of dairy products in a large UK cohort, the UK Biobank. These SNPs might be involved in biological mechanisms related to dairy consumption and be useful as instrumental variables in Mendelian randomisation analyses. The replication of these results in independent cohorts and different ethnic groups and the exploration to identify causal genetic variants will further contribute towards the elucidation of the mechanisms of genetic prediction of dairy consumption.

Chapter 9

Discussion

9.1 Summary of results and links to PhD aims

The overall aim of this PhD project was to contribute towards a deeper understanding of the associations between dairy products and cardio-metabolic health. To achieve this aim, I adopted an interdisciplinary approach including aspects from nutritional, molecular and genetic epidemiology.

In **Chapter 2**, I aimed to describe the dairy consumption patterns across time from a UK representative sample and the contribution of dairy consumption to nutrient intakes. Following trends previously reported over the last decades, I observed that full-fat milk consumption continued decreasing over the years from 2008-2016, whereas yoghurt consumption increased among elderly individuals. I also observed a high contribution of dairy products to nutrient intakes including a contribution of more than 25% to vitamin A, vitamin B_{12}, calcium, phosphorus, saturated fat and trans fat.

In **Chapters 4 and 5**, I aimed to investigate associations of total and types of dairy consumption with cardio-metabolic markers including markers of adiposity, lipidaemia, glycaemia, inflammation, hepatic function and blood pressure to elucidate potential pathways linking dairy consumption to cardio-metabolic disease. Results from cross-sectional (Chapter 4) and prospective (Chapter 5) analyses indicated that the most relevant pathways underpinning the association between dairy consumption and cardio-metabolic disease would involve the development of adiposity, especially central adiposity and dyslipidaemia. The cross-sectional findings specifically showed that low-fat dairy consumption was associated with lower visceral adipose tissue (VAT) and lower ratio of VAT to subcutaneous adipose tissue (SCAT), which has been associated with a higher cardio-metabolic risk independent of body mass index (BMI) and total body fat mass[174]. The prospective findings showed that low-fat fermented dairy products are associated with lower levels of adiposity markers compared to other dairy types.

In **Chapter 6**, I aimed to develop and validate metabolite scores predictive of dairy consumption as potential dairy biomarkers to overcome the measurement error of the self-

reported methods of dietary assessment. I developed such metabolite scores in a discovery set and showed significant associations of the scores with total dairy products, milk and butter in the internal validation dataset. When I used an independent set for external validation, only the metabolite score for total dairy products produced a significant result for the association with total dairy consumption. For the scores for dairy subtypes, I did not obtain significant findings, except for milk and butter, when the scores were developed with additional inclusion of odd-chain saturated fatty acids (OCSFAs) as candidate biomarkers of dairy consumption.

In **Chapter 7**, I aimed to investigate associations of the metabolite scores developed and validated in Chapter 6 with type 2 diabetes (T2D) risk. Metabolite scores for total dairy products, milk and butter were consistently associated with a lower risk of T2D after accounting for socio-demographic and lifestyle factors.

Finally, in **Chapter 8**, I aimed to identify genetic predictors of dairy consumption in addition to the lactase persistence variant. The discovery analysis in a large UK cohort of approximately 500,000 people indicated single nucleotide polymorphisms predictive of total and types of dairy products.

9.2 Interpretation of findings

Interpreting research findings involves consideration of potential biological mechanisms supported by the observed associations under the assumption that they are true. It also involves consideration of methodological challenges, which might influence the accuracy of the observed associations, especially for aetiological research. Examples of aetiological research in this PhD are the analyses in Chapters 4, 5 and 7 investigating associations of dairy consumption with cardio-metabolic markers, and of metabolite scores predictive of dairy consumption with T2D risk. For research related to prediction as in Chapter 6 on exploration of metabolomic predictors of dairy consumption and in Chapter 8 on exploration of genetic predictors of dairy consumption, it is more important to consider methodological challenges, although potential biological mechanisms can provide insight to causal associations and underlying pathways.

9.2.1 Potential biological mechanisms

As described in Chapter 1, total and types of dairy products have been associated either with lower or no risk of cardio-vascular disease and T2D. In order to understand the underlying mechanisms of these associations, it is important to consider the balance of effects of the nutrients, which constitute the dairy food matrix, on cardio-metabolic disease and its intermediate endpoints. For example, yoghurt has been more consistently associated with a lower risk of T2D[12, 16, 18, 19, 21] and it is of interest to understand what are

the nutrients and features of the food matrix of yoghurt that contribute to the favourable associations observed. Suggested features of importance are related to the fermentation process and include the higher bioavailability of nutrients due to the acidity of yoghurt, which decreases gastric pH[32] and the higher content of vitamin K2, which has been associated with a lower risk of T2D[329] and coronary heart disease[330] in prospective cohort studies.

A.Body weight and body composition

Evidence from randomised controlled trials (RCTs) supports favourable effects of total dairy consumption on body weight and composition under conditions of energy restriction. Specifically, total dairy consumption decreased body weight[58, 59, 61, 62], total fat mass[58–62], waist circumference[58, 61], abdominal fat[77, 78] and visceral adipose tissue (VAT)[78], and increased lean mass[58, 59, 61] in RCTs with energy restriction. In accordance with this line of evidence, we observed inverse associations of low-fat dairy consumption with VAT and the ratio of VAT to SCAT (Chapter 4) and of yoghurt and low-fat cheese with body weight and BMI (Chapter 5) and positive associations of milk (total, full-fat and low-fat) with body lean mass (Chapter 4). However, we also observed positive associations between increasing high-fat dairy consumption and increasing body weight and BMI (Chapter 5). There are several mechanisms reported in the literature, which support beneficial effects of dairy products on body composition and they are related to dairy protein, calcium, conjugated linoleic acid (CLA), vitamin K2 and growth hormone.

Dairy proteins, i.e. whey and casein, might be involved in pathways related to body weight and composition. In a meta-analysis of 14 RCTs, although whey protein supplementation decreased body weight and body fat mass from baseline to follow-up, no statistically significant differences were observed with supplementation with other types of protein or carbohydrates[331]. In a 20-week RCT in 48 obese Japanese participants, an intervention of 21 g of milk protein per day decreased body weight, BMI, VAT, SCAT and adiponectin compared with an intervention of 12 g of soy protein and 9 g of milk protein, but it had no effects on lipid and glycaemic markers[332]. The two types of dairy protein i.e. casein and whey might have differential effects on the metabolic regulation of body weight, since their kinetics are different with whey being absorbed and raising the blood amino acid levels faster than casein[333].

A mechanism, which has been extensively studied, is the regulation of appetite as a proxy of food intake. From studies which assess appetite subjectively with questions on hunger, satiety and fullness or with the prospective food intake after the protein preload, there are indications that whey preload decreases appetite compared with casein preload, but results are inconclusive[178]. Veldhorst et al[179] and Alfenas et al[180], proposed some sources of the heterogeneity of effects. In one RCT there was a decrease in appetite with a simultaneous increase in blood amino acids after an intervention with 10% of

energy from whey compared with 10% of energy from casein or soy protein, whereas no differential effects were observed after an intervention with 25% of energy from whey compared with an intervention with equal amounts of casein or soy protein[179]. The authors suggested that a more favourable effect is observed for whey in the lower amounts only, because in these amounts, the whey profile of amino acids, which reach the circulation, includes higher amounts from amino acids such as leucine, isoleucine and tryptophan, which have been associated with appetite regulation, than casein or soy protein[179]. On the contrary, at higher intakes of the proteins, amino acids from all proteins reach the threshold, which is necessary for the effects on appetite to become obvious[179]. The second RCT by Alfenas et al. suggested that intervention with casein led to a lower energy intake after seven days compared with whey, which might be an indication that whey exerts effects on appetite in the short-term, but casein exerts such effects in the longer-term[180]. Apart from the effects of the increasing plasma amino acid levels on appetite regulation, there is some evidence that whey protein intake is associated with gut hormone levels, which might contribute towards the regulation of appetite and include higher levels of cholecystokinin (CCK)[178, 334], gastric inhibitory polypeptide (GIP)[334], glucagon-like peptide 1 (GLP-1)[178, 179, 334], glucagon[334] and peptide YY (PYY)[334] and lower ghrelin levels[334].

Despite any differential effects between the two proteins, an RCT showed that intake of their combination in the form of low-fat milk resulted in lower energy intake than a whey drink or a casein drink alone[335]. In addition to these results, a cross-over RCT showed that the type of dairy product also makes a difference in the regulation of appetite, as consumption of low-fat milk resulted in lower appetite compared with water, but consumption of yoghurt or cheese resulted in even lower appetite[336]. The authors suggest that the decrease in appetite after yoghurt consumption compared with milk consumption might be due to the higher protein content of yoghurt[336]. This was not the case for cheese, but the authors suggested that the requirement of chewing cheese as a solid food compared to a semi-solid food like yoghurt or a liquid food like milk, might have led to lower appetite compared with milk consumption[336].

Calcium is another nutrient that has been extensively studied for its involvement in adiposity-related pathways. A meta-analysis of 13 RCTs did not identify a significant effect of calcium supplementation on body weight, but in subgroup analyses, the authors identified a small decreasing effect when calcium was allocated in the form of supplements rather than as part of dairy consumption[337]. The authors also identified flaws in trial randomisation with imbalances at baseline weights, which might have obscured the true associations[337]. A mechanism for the potential effect of calcium on decreasing body weight and adiposity is through the increase of faecal fat excretion. In a meta-analysis of 13 RCTs, the estimated faecal fat excretion was 5.2 (95% CI: 1.6-8.8) g per 1,241 mg dairy calcium per day [338]. Another mechanism proposed by studies of Zemel et

al, is that dietary calcium leads to an increase in the extracellular calcium, which results in a decrease of parathyroid hormone and 1,25-dihydroxyvitamin D and a subsequent drop in intracellular calcium[339]. The drop in intracellular calcium levels was shown to decrease the expression of the fatty acid synthase gene and increase lipolysis resulting in lower triglyceride storage in adipocytes[177, 340]. Calcium may exert synergistic effects with the gut microbiome for the regulation of body weight and body fat. Related mechanisms include (1) the increase in gastric acid secretion by calcium leading to a decrease in the gastric pH and the decrease of bacterial populations that are not tolerant to acidity (e.g.Salmonella spp) but subsequent increase of those that are more tolerant (e.g.Lactobacillus)[341], (2) maintaining intestinal integrity through the inhibition of the production of cytotoxic substances from bile and fatty acids from colonic bacteria[342] due to an increase in faecal fat excretion, and (3) the decrease in bacterial translocation due to higher intestinal integrity[341].

CLA has also been associated with a lower fat mass by 0.024 kg / g of intake / week in a meta-analysis of 18 RCTs[343]. *In vitro* studies on human adipocytes showed that the decreasing effect on fat mass might be due to the suppression of transcription factors involved in adipogenesis[344, 345]. An example of such a transcription factor is the peroxisome proliferator-activated receptor γ (PPARγ), which is involved in the differentiation of preadipocytes to adipocytes[344, 345].

Supplementation with Vitamin K2 decreased abdominal fat, waist circumference and VAT, and increased adiponectin levels among participants who were good respondents to the supplementation, as assessed by the circulating levels of carboxylated osteocalcin as a marker of Vitamin K2[346].

The observed positive associations between milk consumption and body lean mass could be partly explained by the effect of milk on increasing growth hormone[182, 183], which increases bone mineral density and muscle mass[184]. The bioactive compounds in milk that increase growth hormone in humans are not well-characterised, but an assumption is that one relevant compound might be bovine growth hormone[182]. Milk protein also increased growth hormone in a cross-over RCT compared to placebo[183].

B.Lipidaemia

As described in detail in section 3.2, butter increases total, low-density lipoprotein (LDL-C) and high-density lipoprotein cholesterol (HDL-C) when compared with other types of fat, whereas null effects on blood lipids were reported for total, low-fat or high-fat dairy products[63]. Cheese decreased LDL-C and HDL-C when compared with butter[83]. The current PhD analyses extended previous evidence by investigating associations between dairy types and lipid markers. In line with these results, we also reported positive associations of high-fat dairy consumption and especially butter with total and LDL-C (Chapters 4-5). On the contrary, in our analyses low-fat dairy consumption, milk and

yoghurt (total and low-fat; Chapter 5) were associated with lower total, LDL-C (Chapter 5) and HDL-C (Chapter 4).

Nutrients which have been related to increases in blood lipids are saturated and trans fat. While controlled trials have shown an increasing effect of saturated fat on total, LDL and HDL-C, when it substitutes carbohydrates, mono- or poly-unsaturated fatty acids[200], results from trials using different dairy types with the same saturated fat content have shown differential effects on blood lipids[83, 347]. This implies that constituents and matrices of different dairy products might play a role in the overall effects of these foods on lipidaemia. These differences might be partly explained by differences in nutrient content, but also differential processes such as fermentation or homogenisation. In addition, there is an indication that saturated fat from dairy products contributes to a more favourable lipid profile in terms of the LDL-C particle size, being associated with a decrease in small size LDL-C both when measured as part of the diet and as fatty acid biomarkers in blood (especially C15:0 and C17:0)[348].

A cross-over RCT reported that a diet with 1.5% of energy from ruminant trans fat had a neutral effect on the lipid profiles compared with a diet with 0.8% energy from total trans fat, whereas two diets with 3% energy from ruminant or industrial trans fat resulted in less favourable lipid profiles[349]. Even 1.5% of energy from ruminant trans fatty acids is high for a habitual consumption, which has been estimated to be around 0.5%, so the authors concluded that habitual consumption of ruminant trans fatty acids should not be harmful in relation to lipidaemia[349].

The absence of positive associations with lipids for total and other dairy types (null associations as reported from RCTs or inverse associations as observed in this PhD) might be explained after considering the balance of effects between saturated or trans fat with other nutrients e.g. calcium, probiotics or the food matrix of different dairy types. For example, calcium has been shown to decrease lipolysis and increase lipogenesis[177]. The presence of the milk fat globule membrane has also been related to a more favourable lipidaemic profile. In a parallel group RCT in free living participants, after an intervention of a whipping cream-based snack for eight weeks total and LDL-C did not change, in contrast to the control group, who received a butter oil-based snack with the same amount of energy, total fat, saturated fat, carbohydrates, protein and calcium as the intervention and who increased their total cholesterol by 0.3 ± 0.49 mmol/l and their LDL-C by 0.36 ± 0.5 mmol/l[347]. The main difference in the two snacks was that the milk fat globule membrane was intact in the whipping cream, but was absent in the butter oil due to homogenisation[347]. No differential effects were observed on other lipids, glycaemic and inflammatory markers[347]. In the same study, they also observed down-regulation of gene expression of peripheral blood mononuclear cells in the whipping cream group which was correlated with lower circulating cholesterol[347]. Complementary to these results, in a cross-over RCT participants received buttermilk for four weeks and a placebo product with

the same nutrient profile as buttermilk apart from the milk fat globule membrane nutrient profile[350]. The intervention with buttermilk resulted in 3.1% lower total cholesterol and 10.7% lower triglycerides compared to the intervention with the placebo, whereas the decrease in LDL-C was significantly lower only among participants with higher LDL-C at baseline[350]. The proposed underlying mechanism was the decrease in cholesterol absorption concluded from the correlation of the proxy markers of cholesterol absorption i.e. phytosterol and b-sitosterol with the changes in LDL-C induced by the intervention[350]. Probiotics administered through yoghurt or capsules contributed to a decrease in total and LDL-C, but with no effects on HDL-C or triglycerides as indicated by a meta-analysis of 13 RCTs with durations of interventions ranging from 4-10 weeks[351].

C.Glycaemia and insulinaemia

RCTs have failed to report any effects of total, low-fat or high-fat dairy products on glycaemic markers including fasting blood glucose, insulin, haemoglobin A1c (HbA1c) or the homoestasis model assessment for insulin resistance (HOMA-IR). In agreement with this evidence, the majority of associations we observed between total or types of dairy consumption and glycaemic markers were null (Chapters 4-5) apart from full-fat milk, which was associated with an increase in HbA1c (Chapter 5). Some nutrients contained in dairy products have been associated with lower risk of T2D and with beneficial effects on insulin resistance, while others might exert harmful effects.

Observational studies have reported inverse associations between vitamin D intake and T2D risk[352], but evidence from RCTs has not confirmed beneficial effects of vitamin D supplementation on glycaemic regulation, which poses the question of causality of observational associations[353]. Probiotics also seem to decrease fasting glucose, insulin and HbA1c both among patients with T2D and among people with metabolic syndrome or some of its components as reported from a meta-analysis of 18 RCTs[354]. *In vitro* studies in human adipocytes have shown an increase in insulin resistance after supplementation with CLA[344, 355]. On the contrary an animal and an *in vitro* study showed that vaccenic acid (trans-18:1n-7) decreases insulin resistance by promoting insulin secretion and islet growth through alterations in expression of mRNA[356].

D.Inflammation

Fermented dairy consumption decreased inflammatory markers overall in RCTs[134], but low-fat or high-fat dairy consumption did not have an effect on C-reactive protein (CRP). In this PhD I did not observe any significant associations between dairy products and inflammatory markers. The evidence on potential mechanisms linking dairy consumption to inflammation is not clear. CLA is one nutrient that has been reported to increase the production of inflammatory cytokines[344, 345, 355, 357], but decrease the expression of adiponectin[345] in *in vitro* studies of human tissues.

E.Blood pressure

Evidence from RCTs indicates a beneficial effect of fermented milk on systolic (SBP) and diastolic (DBP) blood pressure[85], but no effects for total[84], low-fat or high-fat dairy products[63]. Milk, total and low-fat dairy products have been associated with lower risk of hypertension in prospective cohort studies[86]. This PhD's results of null associations between dairy consumption and blood pressure agreed with the previously reported associations and extended such findings to individual dairy types. Some nutrients such as protein, calcium and potassium have been linked to beneficial effects on blood pressure especially among hypertensive people and might thus be involved in related mechanisms.

A meta-analysis of 33 RCTs with an average duration of 7.8 weeks showed that interventions including casein-derived lactotripeptides (valine-proline-proline and isoleucine-proline-proline), which are mainly released during fermentation of dairy products by lactobacteria, decreased both systolic (SBP) and diastolic blood pressure (DBP)[358]. The effects remained significant for SBP also after identification of publication bias[358]. In subgroup analyses, stronger effects were observed among people with higher blood pressure levels at baseline and in Japanese populations compared with European populations[358]. *In vitro* studies have shown that these peptides act on blood pressure by inhibiting the angiotensin I converting enzyme by approximately 50% leading to lower vasoconstriction[359].

A meta-analysis of 40 RCTs with a median duration of 2.4 months reported decreases in SBP and DBP by 1.86 mmHg (95% CI: -2.91, -0.81) and 0.99 mmHg (-1.61, -0.37) respectively after an average of 1,200 mg calcium supplementation daily compared with control[360]. Effects were attenuated, but still significant when only double blind trials were included, while effects were stronger among people with baseline calcium intakes lower than 800 mg /day[360]. Increase in potassium intake was also shown to decrease SBP and DBP in a meta-analysis of 21 RCTs with stronger effects observed when potassium intake reached the range 90-120 mmol/day and significant effects were only among hypertensive people[361]. Also, for every 10 mmol/day of magnesium intake the SBP decreased by 4.3 mmHg (-6.3, -2.2) in a meta-analysis of 20 RCTs[362].

Overall, results from clinical trials investigating dairy nutrient effects on blood pressure support the observed beneficial associations of dairy consumption with blood pressure among hypertensive people and weaker or null associations among normotensive people.

F.Dairy metabolite scores and type 2 diabetes risk

Although the use of metabolite scores predictive of dairy consumption might contribute to the elucidation of pathways linking dairy consumption to T2D risk, the specific biological mechanisms involved are not clear and are hard to disentangle. However, some assumptions for relevant mechanisms were discussed in more detail in section 7.5.3. Briefly, the inverse associations observed in this PhD might reflect a set of pathways related to the

metabolites I used to create the score and especially the top metabolites, which consist of phosphatidylcholines (PCs) and sphingomyelins (SMs). Most of the PCs and SMs comprised sum of isobaric and isomeric compounds, which contain the odd-chain saturated fatty acids (OCSFAs) C15:0 and C17:0, or the even-chain fatty acids C16:0 and C18:0. The observed associations might be a result of the balance of effects between OCSFAs and the even-chain fatty acids. Blood C16:0 and C18:0 have been previously associated with higher insulin resistance[298, 299], inflammation[300, 301] and T2D[228]. Considering that fatty acids are expressed in relative amounts, and that OCSFAs can be alternatively produced by even-chain fatty acids[208], higher concentrations of OCSFAs might reflect lower concentrations of C16:0 and C18:0 and thus overall lipid profile, which favours insulin sensitivity.

9.2.2 Epidemiological considerations

Accuracy of observed associations

The accuracy of an observed association can be assessed by evaluating its validity i.e. whether the observed association reflects the true association and its precision i.e. how close the observed association is to the true association[363]. Systematic error or bias can be a threat to validity and random error can be a threat to precision. Validity can be understood in terms of: internal validity that is validity of associations within the source population of the study, and external validity or generalisability that is validity of associations in the general population[363]. Different sources of bias, which can influence the validity of the observed associations are selection bias, information bias and confounding.

Selection bias

Types of selection bias relevant to the studies of this PhD are sampling bias, attrition bias, non-response bias, and collider bias[202]. From the cohorts I used in this project, the National Diet and Nutrition Survey (NDNS) has a lower probability of involving sampling bias compared with the other cohorts. NDNS sampling was random and representative of the general population. On the contrary, eligibility of participants in the other three cohorts (Fenland study, EPIC Norfolk study and UK Biobank) was not based on the representativeness of the sample. While such bias might not have an impact on the internal validity of the findings, it might limit the external validity. For example, more health-conscious people of a higher educational level might have a higher probability of agreeing to participate, thus resulting in a non-representative sample of the general population, which might influence the external validity (generalisability)[202]. Thus, generalisability of our findings from the cohorts other than NDNS was made with caution. Replication of these findings in other studies will provide more confidence to generalise our results from the source population to the general population.

Attrition bias might appear in longitudinal studies if participants lost to follow-up have systematically different characteristics from those who continue. A more specific effect of attrition bias might be the healthy survivor's effect, when participants who were healthier were more likely to stay in the study for follow-up than those who might have developed a disease or died. In the prospective analysis on associations between dairy consumption and cardio-metabolic markers in the EPIC Norfolk study (Chapter 5), such bias could lead to more optimistic results e.g. more favourable associations between dairy consumption and cardio-metabolic markers than the true associations, which could compromise internal validity, but also the inclusion of a healthier sample than the general population, which could compromise the external validity. To handle this possibility, I applied weights based on the inverse of the probability that a participant would be censored with weights generated from prediction models for censoring and results were not materially different from the main approach. In addition, I performed multiple imputation, which has been shown to give equally valid results with complete-case analysis under the assumption of data missing completely at random, but more valid results under the assumption of data missing at random in simulation studies[168]. Multiple imputation is also a useful tool when there are missing data due to non-response.

Non-response bias is another type of bias, which may appear if participants who do not respond to some questions have different characteristics from participants who respond and these characteristics might be related to the risk of disease[202]. Collider bias might also appear in some cases when we adjust for a variable, which affects the outcome and it is also affected by the exposure[202]. An example of collider bias that could appear in this PhD is the one described in Chapter 5 on the bias caused by the adjustment for the baseline outcome in longitudinal observational studies, where the baseline outcome (e.g.baseline weight) affects the follow-up outcome (e.g follow-up weight) and might be affected by the exposure (e.g. yoghurt) or other factors related to the exposure (e.g. an overall healthy lifestyle)[194]. For this reason, we did not adjust for the baseline outcome in the longitudinal analysis, but we accounted for it by using as outcome the change between the baseline and follow-up in the analysis of change.

Information bias

Information bias might also influence the associations observed in this PhD and stems from the error in the information acquired during the data collection[363]. Recall bias is a form of information bias and is related to the accuracy of information provided by participants or their proxies when they have to recall it. For example, recall bias might be greater among elderly people due to memory issues. Most of the cohorts used in this project consisted of adults with a proportion of them being elderly, so we cannot exclude the possibility of recall bias. It is difficult to predict how associations might be influenced by the presence of recall bias. Another form of information bias is social desirability bias, which appears when participants provide the kind of information, which is more socially

acceptable, even if it is not true[364]. In nutritional epidemiology this bias is common. For example, many studies have observed under-reporting of total energy intake due to under-reporting of foods high in fat and sugar among participants who are overweight or obese[145–152]. In this PhD's cross-sectional analysis (Chapter 4), I observed an inverse association of high-fat dairy consumption with BMI and a positive association of low-fat dairy consumption with BMI. This observation might be the result of social desirability bias with participants with higher BMI reporting higher consumption of low-fat dairy products. However, this might also be the result of participants of higher BMI actually consuming higher amounts of low-fat dairy products. The true explanation is not certain, but these are plausible scenarios. In any case, in subsequent cross-sectional analyses of outcomes other than BMI, I adjusted for BMI to partly account for that observation. When classifying participants into categories of a variable, for which information has been obtained from self-reporting, information bias might lead to misclassification of participants, which then can lead to spurious associations[363]. To avoid this kind of bias when classifying participants into categories, I chose to compare extreme categories, as the probability of misclassification between extreme categories is lower than consecutive categories. For example, in the analysis of Chapter 6, where the aim was to identify biomarkers of total and types of dairy consumption, I described metabolic profiles of very low consumers and high consumers. In the genetic analyses of Chapter 8, I did not use the extreme categories only, but I specified consumers and non-consumers of total and types of dairy products. However, in this study, there were two dietary tools available (touchscreen questionnaire and the Oxford WebQ 24-hour recall) and for each tool we had multiple repeated measures of dietary assessment available. The multiple reports of consumption or no consumption by the same participant provided the possibility to confirm the consistency of reporting of dairy consumption. Since genetics would predict consumption over a lifetime, genetically predicted non-consumption should be consistent across the repeated measurements, so with at least one report of dairy consumption, participants were classified as consumers and in this way, we theoretically decreased the effect of misclassification bias on this analysis. Misclassification can be differential or non-differential depending on whether it is systematically related to a participant characteristic and thus it happens towards a specific direction or it is random and it happens in both directions[363]. An example of differential misclassification in this PhD was that mentioned on overweight or obese participants over-reporting total energy intake. Non-differential misclassification will lead to attenuation of true associations[363].

Confounding

Confounding is another phenomenon that can influence internal validity of the results. It is related to the observation of spurious associations due to failure to account for factors, which are related to both the exposure and the outcome, but are not in the causal pathway of the association between the exposure and the outcome and can thus obscure this

association[363]. In RCTs, if the randomisation is successful, factors generally confounding an observational association should have a similar distribution between the intervention and the control group and should not obscure any observed effects of the intervention. However, in observational research confounding is a major point of consideration. All the studies used in this PhD were observational, so I tried to account for confounding by adjusting the models for potential confounders including socio-demographic factors, lifestyle factors and clinical factors.

In nutritional epidemiology, adjustment for energy intake is of special relevance and several ways have been described depending on the purpose of the study[204]. For example, if the aim of a study is to examine substitution of a nutrient/food with other nutrients/foods, then we can adjust for total energy intake and all the nutrients/foods apart from the one that we want to substitute our exposure with[204]. In the studies of this PhD, I was interested in investigating the associations of total and types of dairy consumption with markers of metabolic risk and adiposity accounting for total energy intake and other food groups, so I adjusted for them. In this way, the comparison was for lower dairy consumption by 1 serving/day without any substitution. Although I adjusted for several potential confounders, the possibility of residual confounding cannot be excluded due to either unmeasured confounders or confounders measured with some degree of error. For example, dietary assessment with the food frequency questionnaire gives a rough estimate of habitual consumption and is suitable for the ranking of participants, but it is prone to the types of information bias mentioned above. Estimation of food consumption with a degree of error leads to estimation of nutrient intakes including total energy intake with a higher degree of error due to the additional level of assumptions and sources of error included in the step of dietary data processing to generate nutrient intakes. Confounding can also influence the precision of the observed associations, but it is difficult to identify its exact impact. Precision of effect estimates is mainly influenced by random error, can have the form of sampling error or measurement error and usually leads to attenuation of associations[363].

Effect-measure modification is also a point to consider, which means differential associations for different levels of a factor and which can be approximated by statistical interaction under the assumption of no bias[363]. In the two analyses of aetiological research of Chapters 4 and 5, I investigated potential interactions by age, sex and BMI and when significant, I also presented results from analyses stratified by these factors. However, no major differences were observed, but in the presence of bias, failure to observe statistically significant interaction is not equivalent to the absence of biological interactions. This was not something that I could test within the studies.

Causality of observed associations

Causality is a major consideration in epidemiological studies of aetiological research. Three of the studies included in this PhD were of aetiological character (Chapters 4, 5 and 7), while the other studies were of a descriptive (Chapter 2) or predictive (Chapters 6 and 8) character. There have been several approaches to describe and assess causality in a study, but one approach, which is widely used in epidemiological research is the Bradford-Hill criteria[363]. These criteria include strength of association, consistency, specificity, temporality, biological gradient, plausibility, coherence, experimental evidence and analogy[363]. The greater the number of these criteria that are fulfilled in a study, the higher the probability that the associations observed are causal. However, although these criteria can give an overall good guidance to make causal inferences, they also have limitations.

Strength of association

According to the Bradford-Hill criteria, the stronger an association, the more probable that it is causal. However, it is also possible that the true biological association is weak, especially when it is about a multi-factorial outcome, such as cardio-metabolic disease, which might have many risk factors of small effects rather than only a few with large affects. Nevertheless, the scenario that a weak association is the result of the attenuation due to confounding is still possible. In two of the analyses of aetiological character, where there were multiple exposures and outcomes, I presented results in forest plots with their 95% confidence intervals, so that we could compare the magnitude of estimates and their confidence intervals across different associations assessed under the same conditions. Such conditions might refer to the same study design or adjustment for the same confounders. Forest plots provided us with an overall picture of the most convincing associations out of all that we observed, especially since it is difficult to judge in absolute terms which associations are considered strong. In our investigation of associations between dairy metabolite scores and T2D risk, significant associations were considered on average strong with a risk reduction ranging from 17% to 43%.

Consistency

Consistency of associations observed under different circumstances and different populations might be an indication of causality. The lack of consistency, though, is not equivalent to lack of causality, because some associations might be specific to certain populations e.g. certain ethnic groups. In this PhD analysis, I used as a positive control the association between butter consumption and LDL-C, which has been established from evidence of RCTs and seems to be replicated across different populations and circumstances. The observation of a positive association, which was consistent with previous evidence provided more confidence for the observed associations. In addition, in both the cross-sectional (Chapter 4) and the prospective (Chapter 5) analyses, we overall observed significant associations of dairy consumption with markers of adiposity

or lipidaemia. This might further confirm that adiposity and lipidaemia might be more strongly related to pathways underpinning associations between dairy consumption and cardio-metabolic disease. In the survival analysis on associations between dairy metabolite scores and T2D risk, we confirmed that top metabolites associated in a similar way with T2D in other observational studies.

Specificity

The criterion of specificity refers to specific associations observed between an exposure and an outcome. When the exposure is a dietary factor such as dairy products, it is often not possible to identify an outcome that the exposure will be specific for. Likewise, when a disease is multi-factorial i.e. multiple risk factors (causes) can lead to disease, such as cardio-metabolic disease, it is often not possible to identify exposures specific to this outcome.

Temporality

For the temporality assumption to be fulfilled, the exposure should precede the outcome. In our prospective analysis (Chapter 5), this assumption was fulfilled, but not in our cross-sectional analysis (Chapter 4), where the possibility of reverse causation i.e. that the outcome affects the exposure is high. Although it is not possible to know the direction of association and eliminate the possibility for reverse causation in a cross-sectional analysis, it is reasonable to assume that more phenotypically obvious outcomes such as BMI might influence the behaviour of the participants than less obvious outcomes such as visceral fat, when it is assessed independent of BMI. To do this, we adjusted the rest of associations for BMI, especially after the observation that low-fat dairy consumption was associated with a higher BMI. Nevertheless, these findings are important for hypothesis generation, as for example the association between low-fat dairy consumption and the VAT/SCAT ratio has not been previously examined.

Biological gradient and analogy

The criteria of biological gradient, i.e. a linear dose-response association and analogy, have been criticised. Non-linear associations can also be causal. In addition, the assumption of analogy e.g. assumption of harmful effects of an exposure on outcome B, because they were observed on outcome A, might contradict the assumption of specificity, that the exposure would be specific to one outcome.

Biological plausibility, coherence and experimental evidence

Biological plausibility, coherence and experimental evidence concerning the associations we observed were discussed in the previous paragraph about evidence on potential mechanisms.

Overall, our analyses that involve aetiological research fulfil partly or fully some of the Bradford-Hill criteria including strength of association, consistency, temporality, biological gradient, plausibility, coherence and experimental evidence. Based on that, although we can be more confident of the associations we observed in terms of causality, we have to

acknowledge the limitations of our research and we cannot be certain about causality. This work adds to the existing evidence, but is only a part of the evidence synthesis and we should draw conclusions based on the total body of evidence rather than individual studies.

9.3 Strengths and limitations

The overall strength of this PhD was the application of a multi-disciplinary approach investigating aspects of nutritional, molecular and genetic epidemiology that might contribute towards a greater understanding of the link between dairy consumption and cardio-metabolic disease. To describe dairy consumption patterns and their contribution to nutrient intakes in the UK, I used a representative sample of the UK population and contributed to increased precision of estimates by incorporating dairy consumption as part of composite foods, which made a difference for some of the estimates.

Although RCTs are the gold standard in the ranking of evidence quality, they also have some limitations including low sample size, short duration, questionable compliance, impossible blinding and diversity in comparison groups as mentioned in section 3.8[63], which makes observational studies a valuable source of research. Another limitation I identified in the literature related to RCTs and especially in pooled results reported in meta-analyses was that the majority of the available evidence concerned total dairy products or total low- or high-fat dairy products rather than specific types[58–61, 63, 84]. Although this is still informative, due to the heterogeneity of dairy products as a food group, it is of greater interest to investigate specific dairy types. Therefore, across all the studies in this PhD, I consistently investigated both total dairy products, but also dairy types with the highest level of detail possible based on the limitations of the dietary assessment methods available.

To explore potential biomarkers of dairy consumption, I developed metabolite scores, which might better reflect and be more specific to dairy consumption than single biomarkers and I assessed the scores for both internal and external validity, which many nutritional metabolomics studies lack[288]. Finally, to explore potential genetic predictors of dairy consumption, I used the largest biobank in the UK with genetic data from approximately 500,000 people, which gave us sufficient power to detect some significant genetic predictors. This work is novel, as there are no previously published genome-wide association studies on dairy consumption.

This PhD also had limitations. The observational nature of the studies included did not allow us to be certain about causality of the associations observed. Replication of such results in RCTs or Mendelian Randomisation studies might provide higher confidence about causality. I did take the first step towards enabling Mendelian Randomisation, which was to identify novel genetic predictors of dairy consumption additional to the lactase persistence variant. These results can be useful as instrumental variables in future

studies. Other limitations were related to the methodological considerations described in section 9.2.2 including the possibility of different types of selection bias (sampling, attrition, non-response and collider bias) and information bias (recall and social desirability bias), as well as residual confounding despite my effort to adjust for many relevant potential confounders. It is difficult to disentangle the effect that presence of bias and confounding could have on the effect estimates we reported. Although I followed some approaches to try and minimise the sources of such errors, bias can be inherent within the initial study design and it is not possible to completely control for it. The use of self-reported methods of dietary assessment in this PhD is more prone to information bias, and also random measurement error and can give less valid and less precise estimates. However, nutritional epidemiology has to rely on available methods, which are currently based on self-report. My use of plasma fatty acids and metabolite scores from metabolomics was an attempt to overcome some of these challenges by using objective assessment.

9.4 Implications for public health

Results from this PhD can be informative for public health research and public health policy. In Chapter 2, I reported trends in total and types of dairy consumption over the past eight years and the contribution of dairy products to nutrient intakes with a higher precision in the estimates than previously reported. Monitoring food consumption patterns in a population can inform policymakers on the compliance of the public to the dietary guidelines, detect areas of concern and further inform public health decision-making. In Chapters 4 and 5, I investigated associations of total and types of dairy consumption with markers of metabolic risk and adiposity, detecting significant associations between certain dairy types and certain markers of adiposity and lipidaemia. Although we cannot be sure about causality of the associations that were observed, we consider our results important for hypothesis generation, which can be further tested in RCTs. Our novel result on the cross-sectional association between low-fat dairy consumption and VAT/SCAT independent of BMI has not previously reported and effects of dairy consumption on body fat distribution would be useful to be characterised in RCTs.

In Chapter 6, I developed metabolite scores predictive of dairy consumption. I was able to validate the score for total dairy consumption and there was some evidence of validation for the scores for milk and butter consumption, but I could not validate scores for yoghurt and cheese. Although this study had limitations, the approach of metabolite scores rather than single biomarkers is important, as metabolite scores can be more specific to foods, such as dairy products, which have a lot of nutrients in common with other foods and thus a more holistic approach reflecting dairy food matrix would be more informative. The use of metabolite scores can have multiple implications from an objective method of dietary assessment in nutrition research to the elucidation of potential pathways linking dairy

consumption to disease, but also the monitoring of the nutritional status in a clinical setting. We are still away from the application of personalised nutrition in the clinical setting based on the metabolomic profile of patients or even more so in a public health setting based on the profile of the people from the general population. However, when personalised nutrition becomes a possibility and is applied in a clinical or public setting, such scores might be very useful. Although currently biomarkers cannot completely replace self-reported methods of dietary assessment, metabolic profiles that predict dairy consumption could complement self-reported methods and account for their limitations. In patients who cannot report their diet e.g. critically ill or with memory problems, biomarkers can prove even more useful for assessment of diet and nutritional status. Complementary to metabolite scores, dairy predictors on the genetic level related to the work we did in Chapter 8, can also be useful for personalised nutrition apart from their usefulness in research.

9.5 Suggestions for future research

This PhD has laid the foundations for advancements in nutrition research, focusing on the link between total and types of dairy products and cardio-metabolic disease, spanning scientific areas of nutritional, molecular and genetic epidemiology. However, further research should build on the findings of this PhD. Some examples are as below.

In Chapter 2, we described dairy consumption trends over the last eight years, but it is of greater interest to monitor consumption trends over decades. Available data on dairy consumption patterns in the UK before the National Diet and Nutrition Survey (NDNS) are related to proxies of individual consumption such as food supply and availability or consumption on the household level. Thus, the NDNS rolling programme will provide valuable information in future years to continue describing and monitoring food consumption patterns in the general UK population, but also specific "at risk" populations or specific ethnic groups.

In Chapters 4 and 5, we identified some associations of total and types of dairy consumption with certain markers of adiposity and lipidaemia. Some of them, e.g. the positive association between butter and LDL-C, are well-characterised from prior RCTs. However, others such as the association between low-fat dairy consumption and VAT/SCAT, have not been previously described. Thus, it would be useful that the hypotheses generated from this work are further tested in RCTs, but also animal or mechanistic studies for the disentanglement of relevant pathways.

In Chapter 6, we adopted an approach developing metabolite scores predictive of dairy consumption and validating some of them and in Chapter 7, we investigated associations between the scores and T2D risk. Limitations of our study included the observational design, the limited set of metabolites we used and the biological sample we used, which might partly explain the failure to validate the scores for yoghurt and cheese. Therefore, it

would be useful if our approach is followed in RCTs, with different sets of metabolites and biological samples e.g. including metabolites related to the gut microbiome, which might be additionally useful and specific for fermented products like yoghurt and cheese.

Finally, the discovery analysis described in Chapter 8 should be further replicated in other studies, but also other populations and ethnic groups and the genetic predictors should be applied in Mendelian Randomisation analyses for the investigation of causal associations between dairy consumption and cardio-metabolic disease.

9.6 Conclusions

The multi-disciplinary approach followed in this PhD, which included aspects of nutritional, molecular and genetic epidemiology, laid the foundations for a greater understanding of the pathways underlying the associations between dairy products and cardio-metabolic disease with a focus on adiposity and lipidaemia. It also generated hypotheses to be further tested in studies of different designs such as the cross-sectional association of low-fat dairy consumption with VAT/SCAT that we observed and some metabolites strongly associated with dairy consumption such as SM-OH C14:1. This PhD also highlighted the usefulness of some approaches such as the use of metabolite scores rather than single metabolites to predict dairy consumption, but also their limitations, which did not allow for validation of the scores and can be further addressed in future research. Such an approach can be useful for future research potentially including more aspects of other disciplines and further advancing this area of research. Finally, this PhD contributed new knowledge on the possible genetic determinants of consumption of dairy types. Though this work highlighted the challenges of doing genome-wide association analyses for dietary factors that are ascertained by self-report, future replication of the work and the use of biomarkers will further enhance this area of science. Such knowledge will also enable the investigation of causality of association with cardio-metabolic disease.

References

[1] Naghavi M, Abajobir AA, Abbafati C, Abbas KM, Abd-Allah F, Abera SF, et al. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the Global Burden of Disease Study 2016. The Lancet. 2017;390(10100):1151–1210.

[2] Federation ID. IDF Diabetes Atlas, 8th edn. Brussels, Belgium: International Diabetes Federation, 2017;.

[3] Gakidou E, Afshin A, Abajobir AA, Abate KH, Abbafati C, Abbas KM, et al. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. The Lancet. 2017;390(10100):1345–1422.

[4] Bian S, Hu J, Zhang K, Wang Y, Yu M, Ma J. Dairy product consumption and risk of hip fracture: a systematic review and meta-analysis. BMC Public Health. 2018;18(1):165.

[5] Mozaffarian D. Dietary and Policy Priorities for Cardiovascular Disease, Diabetes, and Obesity A Comprehensive Review. Circulation. 2016;133(2):187–225.

[6] Chen GC, Wang Y, Tong X, Szeto IM, Smit G, Li ZN, et al. Cheese consumption and risk of cardiovascular disease: a meta-analysis of prospective studies. European journal of nutrition. 2017;56(8):2565–2575.

[7] Guo J, Astrup A, Lovegrove JA, Gijsbers L, Givens DI, Soedamah-Muthu SS. Milk and dairy consumption and risk of cardiovascular diseases and all-cause mortality: dose–response meta-analysis of prospective cohort studies. Eur J Epidemiol. 2017; p. 1–19.

[8] Schwingshackl L, Hoffmann G, Lampousi AM, Knüppel S, Iqbal K, Schwedhelm C, et al.. Food groups and risk of type 2 diabetes mellitus: a systematic review and meta-analysis of prospective studies; 2017.

[9] Wu L, Sun D. Consumption of Yogurt and the Incident Risk of Cardiovascular Disease: A Meta-Analysis of Nine Cohort Studies. Nutrients. 2017;9(3).

[10] Alexander DD, Bylsma LC, Vargas AJ, Cohen SS, Doucette A, Mohamed M, et al. Dairy consumption and CVD: a systematic review and meta-analysis. Br J Nutr. 2016;115(4):737–50.

[11] de Goede J, Soedamah-Muthu SS, Pan A, Gijsbers L, Geleijnse JM. Dairy consumption and risk of stroke: a systematic review and updated dose–response meta-analysis of prospective cohort studies. Journal of the American Heart Association. 2016;5(5):e002787.

[12] Gijsbers L, Ding EL, Malik VS, de Goede J, Geleijnse JM, Soedamah-Muthu SS. Consumption of dairy foods and diabetes incidence: a dose-response meta-analysis of observational studies. Am J Clin Nutr. 2016;103:1111–1124.

[13] Mullie P, Pizot C, Autier P. Daily milk consumption and all-cause mortality, coronary heart disease and stroke: a systematic review and meta-analysis of observational cohort studies. BMC Public Health. 2016;16(1):1236.

[14] Pimpin L, Wu JH, Haskelberg H, Del Gobbo L, Mozaffarian D. Is butter back? A systematic review and meta-analysis of butter consumption and risk of cardiovascular disease, diabetes, and total mortality. PLoS One. 2016;11(6):e0158118.

[15] Qin LQ, Xu JY, Han SF, Zhang ZL, Zhao YY, Szeto IM. Dairy consumption and risk of cardiovascular disease: an updated meta-analysis of prospective cohort studies. Asia Pac J Clin Nutr. 2015;24(1):90–100.

[16] Chen M, Sun Q, Giovannucci E, Mozaffarian D, Manson JE, Willett WC, et al. Dairy consumption and risk of type 2 diabetes: 3 cohorts of US adults and an updated meta-analysis. BMC Med. 2014;12:215.

[17] Hu D, Huang J, Wang Y, Zhang D, Qu Y. Dairy foods and risk of stroke: a meta-analysis of prospective cohort studies. Nutr Metab Cardiovasc Dis. 2014;24(5):460–9.

[18] Aune D, Norat T, Romundstad P, Vatten LJ. Dairy products and the risk of type 2 diabetes: a systematic review and dose-response meta-analysis of cohort studies. Am J Clin Nutr. 2013;98(4):1066–83.

[19] Gao D, Ning N, Wang C, Wang Y, Li Q, Meng Z, et al. Dairy products consumption and risk of type 2 diabetes: systematic review and dose-response meta-analysis. PLoS One. 2013;8(9):e73965.

[20] Soedamah-Muthu SS, Ding EL, Al-Delaimy WK, Hu FB, Engberink MF, Willett WC, et al. Milk and dairy consumption and incidence of cardiovascular diseases and all-cause mortality: dose-response meta-analysis of prospective cohort studies. Am J Clin Nutr. 2011;93(1):158–71.

[21] Tong X, Dong JY, Wu ZW, Li W, Qin LQ. Dairy consumption and risk of type 2 diabetes mellitus: a meta-analysis of cohort studies. Eur J Clin Nutr. 2011;65(9):1027–31.

[22] Elwood PC, Pickering JE, Givens DI, Gallacher JE. The consumption of milk and dairy foods and the incidence of vascular disease and diabetes: an overview of the evidence. Lipids. 2010;45(10):925–39.

[23] England PH. The Eatwell Guide. 2016;.

[24] Agriculture USDoH, Services H, of USD. 2015–2020 Dietary Guidelines for Americans. 8th Edition. 2015;.

[25] Gordon T. The Diet-heart idea Outline of a history. Am J Epidemiol. 1988;127(2):220–225.

[26] de Souza RJ, Mente A, Maroleanu A, Cozma AI, Ha V, Kishibe T, et al. Intake of saturated and trans unsaturated fatty acids and risk of all cause mortality, cardiovascular disease, and type 2 diabetes: systematic review and meta-analysis of observational studies. BMJ. 2015;351:h3978.

[27] Jakobsen MU, O'Reilly EJ, Heitmann BL, Pereira MA, Bälter K, Fraser GE, et al. Major types of dietary fat and risk of coronary heart disease: a pooled analysis of 11 cohort studies. The American journal of clinical nutrition. 2009;89(5):1425–1432.

[28] Mozaffarian D, Micha R, Wallace S. Effects on Coronary Heart Disease of Increasing Polyunsaturated Fat in Place of Saturated Fat: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. Plos Medicine. 2010;7(3).

[29] Hooper L, Martin N, Abdelhamid A, Davey Smith G. Reduction in saturated fat intake for cardiovascular disease. Cochrane Database Syst Rev. 2015;(6):CD011737.

[30] Organisation WH. Guidelines: Saturated fatty acids and trans-fatty acids intakes for adults and children. 2018;.

[31] Nutrition Division F. Codex Alimentarius - Joint FAO/WHO Food Standards Programme. 2011;.

[32] Buttriss J. Nutritional properties of fermented milk products. International Journal of Dairy Technology. 1997;50(1):21–27.

[33] Lindmark Månsson H. Fatty acids in bovine milk fat. Food and nutrition research. 2008;52(1):1821.

[34] Dewettinck K, Rombaut R, Thienpont N, Le TT, Messens K, Van Camp J. Nutritional and technological aspects of milk fat globule membrane material. International dairy journal. 2008;18(5):436–457.

[35] Roe M, Pinchen H, Church S, Finglas P. McCance and Widdowson's the composition of foods seventh summary edition and updated composition of foods integrated dataset. Nutrition Bulletin. 2015;40(1):36–39.

[36] Yantcheva B, Golley S, Topping D, Mohr P. Food avoidance in an Australian adult population sample: the case of dairy products. Public Health Nutr. 2015; p. 1–8.

[37] Khatibzadeh S, Saheb Kashaf M, Micha R, Fahimi S, Shi P, Elmadfa I, et al. A global database of food and nutrient consumption. Bull World Health Organ. 2016;94(12):931–934.

[38] Hobbs DA, Lovegrove JA, Givens ID. In: Watson RR, De Meester F, editors. Chapter 5 - Dairy Products: Their Role in the Diet and Effects on Cardiovascular Disease. AOCS Press; 2016. p. 111–129.

[39] Fagerli RA, Wandel M. Gender differences in opinions and practices with regard to a" healthy diet". Appetite. 1999;32(2):171–190.

[40] Smith AM, Baghurst KI. Public health implications of dietary differences between social status and occupational category groups. Journal of epidemiology and Community Health. 1992;46(4):409–416.

[41] Roos E, Prättälä R, Lahelma E, Kleemola P, Pietinen P. Modern and healthy?: socioeconomic differences in the quality of diet. European journal of clinical nutrition. 1996;50(11):753–760.

[42] Fraser G, Welch A, Luben R, Bingham S, Day N. The effect of age, sex, and education on food consumption of a middle-aged English cohort—EPIC in East Anglia. Preventive medicine. 2000;30(1):26–34.

[43] Hjartåker A, Lagiou A, Slimani N, Lund E, Chirlaque M, Vasilopoulou E, et al. Consumption of dairy products in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort: data from 35955 24-hour dietary recalls in 10 European countries. Public health nutrition. 2002;5(6b):1259–1271.

[44] Kelleher C, Friel S, Nolan G, Forbes B. Effect of social variation on the Irish diet. Proceedings of the Nutrition Society. 2002;61(04):527–536.

[45] Delgado CL. Rising consumption of meat and milk in developing countries has created a new food revolution. The journal of nutrition. 2003;133(11):3907S–3910S.

[46] Sanchez-Villegas A, Martinez J, Prättälä R, Toledo E, Roos G, Martinez-Gonzalez M. A systematic review of socioeconomic differences in food habits in Europe: consumption of cheese and milk. European journal of clinical nutrition. 2003;57(8):917–929.

[47] Deshmukh-Taskar P, Nicklas TA, Yang SJ, Berenson GS. Does food group consumption vary by differences in socioeconomic, demographic, and lifestyle factors in young adults? The Bogalusa Heart Study. Journal of the American Dietetic Association. 2007;107(2):223–234.

[48] Fulgoni V, Nicholls J, Reed A, Buckley R, Kafer K, Huth P, et al. Dairy consumption and related nutrient intake in African-American adults and children in the United States: continuing survey of food intakes by individuals 1994-1996, 1998, and the National Health And Nutrition Examination Survey 1999-2000. Journal of the American Dietetic Association. 2007;107(2):256–264.

[49] Boylan S, Lallukka T, Lahelma E, Pikhart H, Malyutina S, Pajak A, et al. Socioeconomic circumstances and food habits in Eastern, Central and Western European populations. Public health nutrition. 2011;14(04):678–687.

[50] Possa G, de Castro MA, Marchioni DML, Fisberg RM, Fisberg M. Probability and amounts of yogurt intake are differently affected by sociodemographic, economic, and lifestyle factors in adults and the elderly—results from a population-based study. Nutrition Research. 2015;35(8):700–706.

[51] Panahi S, Fernandez M, Marette A, Tremblay A. Yogurt, diet quality and lifestyle factors. European journal of clinical nutrition. 2017;71(5):573.

[52] Cormier H, Thifault E, Garneau V, Tremblay A, Drapeau V, Pérusse L, et al. Association between yogurt consumption, dietary patterns, and cardio-metabolic risk factors. European journal of nutrition. 2016;55(2):577–587.

[53] NatCen Social Research UCLMS MRC Elsie Widdowson Laboratory. National Diet and Nutrition Survey Years 1-8, 2008/09-2015/16; 2018.

[54] Research MHN. Food Standards Agency Standard Recipes Database, 1992-2012. UK Data Service SN: 8159. 2017;.

[55] Hulshof K, van Erp-Baart M, Anttolainen M, Becker W, Church S, Couet C, et al. Intake of fatty acids in western Europe with emphasis on trans fatty acids: the TRANSFAIR Study. European journal of clinical nutrition. 1999;53(2):143.

[56] Tarja Laaninen EPRS. Trans Fats - Overview of recent developments. European Parliament. 2016;.

[57] Health Do. Nutrient analysis of a range of processed foods with particular reference to trans fatty acids. 2013;.

[58] Geng T, Qi L, Huang T. Effects of Dairy Products Consumption on Body Weight and Body Composition Among Adults: An Updated Meta-analysis of 37 Randomized Control Trials. Molecular nutrition and food research. 2018;62(1).

[59] Stonehouse W, Wycherley T, Luscombe-Marsh N, Taylor P, Brinkworth G, Riley M. Dairy Intake Enhances Body Weight and Composition Changes during Energy Restriction in 18–50-Year-Old Adults—A Meta-Analysis of Randomized Controlled Trials. Nutrients. 2016;8(7):394.

[60] Booth AO, Huggins CE, Wattanapenpaiboon N, Nowson CA. Effect of increasing dietary calcium through supplements and dairy food on body weight and body composition: a meta-analysis of randomised controlled trials. Br J Nutr. 2015;114(7):1013–25.

[61] Abargouei AS, Janghorbani M, Salehi-Marzijarani M, Esmaillzadeh A. Effect of dairy consumption on weight and body composition in adults: a systematic review and meta-analysis of randomized controlled clinical trials. Int J Obes (Lond). 2012;36(12):1485–93.

[62] Chen M, Pan A, Malik VS, Hu FB. Effects of dairy intake on body weight and fat: a meta-analysis of randomized controlled trials. Am J Clin Nutr. 2012;96(4):735–47.

[63] Benatar JR, Sidhu K, Stewart RA. Effects of high and low fat dairy food on cardio-metabolic risk factors: a meta-analysis of randomized studies. PLoS One. 2013;8(10):e76480.

[64] Snijder MB, van Dam RM, Stehouwer CD, Hiddink GJ, Heine RJ, Dekker JM. A prospective study of dairy consumption in relation to changes in metabolic risk factors: the Hoorn Study. Obesity (Silver Spring). 2008;16(3):706–9.

[65] Vergnaud AC, Peneau S, Chat-Yung S, Kesse E, Czernichow S, Galan P, et al. Dairy consumption and 6-y changes in body weight and waist circumference in middle-aged French adults. Am J Clin Nutr. 2008;88(5):1248–55.

[66] Rajpathak SN, Rimm EB, Rosner B, Willett WC, Hu FB. Calcium and dairy intakes in relation to long-term weight gain in US men. The American journal of clinical nutrition. 2006;83(3):559–566.

[67] Wang H, Troy LM, Rogers GT, Fox CS, McKeown NM, Meigs JB, et al. Longitudinal association between dairy consumption and changes of body weight and waist circumference: the Framingham Heart Study. Int J Obes (Lond). 2014;38(2):299–305.

[68] Fumeron F, Lamri A, Abi Khalil C, Jaziri R, Porchay-Balderelli I, Lantieri O, et al. Dairy consumption and the incidence of hyperglycemia and the metabolic syndrome: results from a french prospective study, Data from the Epidemiological Study on the Insulin Resistance Syndrome (DESIR). Diabetes Care. 2011;34(4):813–7.

[69] Rautiainen S, Wang L, Lee IM, Manson JE, Buring JE, Sesso HD. Dairy consumption in association with weight change and risk of becoming overweight or obese in middle-aged and older women: a prospective cohort study. The American journal of clinical nutrition. 2016;103(4):979–988.

[70] Mozaffarian D, Hao T, Rimm EB, Willett WC, Hu FB. Changes in diet and lifestyle and long-term weight gain in women and men. New England Journal of Medicine. 2011;364(25):2392–2404.

[71] Samara A, Herbeth B, Ndiaye NC, Fumeron F, Billod S, Siest G, et al. Dairy product consumption, calcium intakes, and metabolic syndrome-related factors over 5 years in the STANISLAS study. Nutrition. 2013;29(3):519–24.

[72] Rosell M, Hakansson NN, Wolk A. Association between dairy food consumption and weight change over 9 y in 19,352 perimenopausal women. Am J Clin Nutr. 2006;84(6):1481–8.

[73] Martinez-Gonzalez MA, Sayon-Orea C, Ruiz-Canela M, de la Fuente C, Gea A, Bes-Rastrollo M. Yogurt consumption, weight change and risk of overweight/obesity: the SUN cohort study. Nutr Metab Cardiovasc Dis. 2014;24(11):1189–96.

[74] Wajchenberg BL. Subcutaneous and visceral adipose tissue: their relation to the metabolic syndrome. Endocr Rev. 2000;21(6):697–738.

[75] Crichton GE, Howe PR, Buckley JD, Coates AM, Murphy KJ. Dairy consumption and cardiometabolic health: outcomes of a 12-month crossover trial. Nutr Metab (Lond). 2012;9:19.

[76] Rideout TC, Marinangeli CP, Martin H, Browne RW, Rempel CB. Consumption of low-fat dairy foods for 6 months improves insulin resistance without adversely affecting lipids or bodyweight in healthy adults: a randomized free-living cross-over study. Nutr J. 2013;12:56.

[77] Zemel MB, Donnelly JE, Smith BK, Sullivan DK, Richards J, Morgan-Hanusa D, et al. Effects of dairy intake on weight maintenance. Nutr Metab (Lond). 2008;5:28.

[78] Josse AR, Atkinson SA, Tarnopolsky MA, Phillips SM. Increased consumption of dairy foods and protein during diet-and exercise-induced weight loss promotes fat mass loss and lean mass gain in overweight and obese premenopausal women. The Journal of nutrition. 2011;141(9):1626–1634.

[79] Van Loan MD, Keim NL, Adams SH, Souza E, Woodhouse LR, Thomas A, et al. Dairy Foods in a Moderate Energy Restricted Diet Do Not Enhance Central Fat, Weight, and Intra-Abdominal Adipose Tissue Losses nor Reduce Adipocyte Size or Inflammatory Markers in Overweight and Obese Adults: A Controlled Feeding Study. J Obes. 2011;2011:989657.

[80] Halkjaer J, Tjonneland A, Overvad K, Sorensen TI. Dietary predictors of 5-year changes in waist circumference. J Am Diet Assoc. 2009;109(8):1356–66.

[81] Zong G, Sun Q, Yu D, Zhu J, Sun L, Ye X, et al. Dairy consumption, type 2 diabetes, and changes in cardiometabolic traits: a prospective cohort study of middle-aged and older Chinese in Beijing and Shanghai. Diabetes Care. 2014;37(1):56–63.

[82] Babio N, Becerra-Tomas N, Martinez-Gonzalez MA, Corella D, Estruch R, Ros E, et al. Consumption of Yogurt, Low-Fat Milk, and Other Low-Fat Dairy Products Is Associated with Lower Risk of Metabolic Syndrome Incidence in an Elderly Mediterranean Population. J Nutr. 2015;145(10):2308–16.

[83] de Goede J, Geleijnse JM, Ding EL, Soedamah-Muthu SS. Effect of cheese consumption on blood lipids: a systematic review and meta-analysis of randomized controlled trials. Nutr Rev. 2015;73(5):259–75.

[84] Ding M, Huang T, Bergholdt HK, Nordestgaard BG, Ellervik C, Qi L. Dairy consumption, systolic blood pressure, and risk of hypertension: Mendelian randomization study. bmj. 2017;356:j1000.

[85] Dong JY, Szeto IM, Makinen K, Gao Q, Wang J, Qin LQ, et al. Effect of probiotic fermented milk on blood pressure: a meta-analysis of randomised controlled trials. Br J Nutr. 2013;110(7):1188–94.

[86] Soedamah-Muthu SS, Verberne LD, Ding EL, Engberink MF, Geleijnse JM. Dairy consumption and incidence of hypertension: a dose-response meta-analysis of prospective cohort studies. Hypertension. 2012;60(5):1131–7.

[87] Zock PL, Katan MB. Butter, margarine and serum lipoproteins. Atherosclerosis. 1997;131(1):7–16.

[88] Brassard D, Tessier-Grenier M, Allaire J, Rajendiran E, She Y, Ramprasath V, et al. Comparison of the impact of SFAs from cheese and butter on cardiometabolic risk factors: a randomized controlled trial–3. The American journal of clinical nutrition. 2017;105(4):800–809.

[89] Engel S, Tholstrup T. Butter increased total and LDL cholesterol compared with olive oil but resulted in higher HDL cholesterol compared with a habitual diet. Am J Clin Nutr. 2015;102(2):309–15.

[90] Dias CB, Phang M, Wood LG, Garg ML. Postprandial lipid responses do not differ following consumption of butter or vegetable oil when consumed with omega-3 polyunsaturated fatty acids. Lipids. 2015;50(4):339–47.

[91] Anderson-Vasquez HE, Perez-Martinez P, Ortega Fernandez P, Wanden-Berghe C. Impact of the consumption of a rich diet in butter and it replacement for a rich diet in extra virgin olive oil on anthropometric, metabolic and lipid profile in postmenopausal women. Nutr Hosp. 2015;31(6):2561–70.

[92] Nestel PJ, Pally S, MacIntosh GL, Greeve MA, Middleton S, Jowett J, et al. Circulating inflammatory and atherogenic biomarkers are not increased following single meals of dairy foods. Eur J Clin Nutr. 2012;66(1):25–31.

[93] Svensson J, Rosenquist A, Ohlsson L. Postprandial lipid responses to an alpha-linolenic acid-rich oil, olive oil and butter in women: a randomized crossover trial. Lipids Health Dis. 2011;10:106.

[94] Hjerpsted J, Leedo E, Tholstrup T. Cheese intake in large amounts lowers LDL-cholesterol concentrations compared with butter intake of equal fat content. Am J Clin Nutr. 2011;94(6):1479–84.

[95] Perez-Martinez P, Ordovas JM, Garcia-Rios A, Delgado-Lista J, Delgado-Casado N, Cruz-Teno C, et al. Consumption of diets with different type of fat influences triacylglycerols-rich lipoproteins particle number and size during the postprandial state. Nutr Metab Cardiovasc Dis. 2011;21(1):39–45.

[96] Ould Mohamedou MM, Zouirech K, El Messal M, El Kebbaj MS, Chraibi A, Adlouni A. Argan Oil Exerts an Antiatherogenic Effect by Improving Lipids and Susceptibility of LDL to Oxidation in Type 2 Diabetes Patients. Int J Endocrinol. 2011;2011:747835.

[97] Masson CJ, Mensink RP. Exchanging saturated fatty acids for (n-6) polyunsaturated fatty acids in a mixed meal may decrease postprandial lipemia and markers of inflammation and endothelial activity in overweight men. J Nutr. 2011;141(5):816–21.

[98] Gagliardi AC, Maranhao RC, de Sousa HP, Schaefer EJ, Santos RD. Effects of margarines and butter consumption on lipid profiles, inflammation markers and lipid transfer to HDL particles in free-living subjects with the metabolic syndrome. Eur J Clin Nutr. 2010;64(10):1141–9.

[99] Palomaki A, Pohjantahti-Maaroos H, Wallenius M, Kankkunen P, Aro H, Husgafvel S, et al. Effects of dietary cold-pressed turnip rapeseed oil and butter on serum lipids, oxidized LDL and arterial elasticity in men with metabolic syndrome. Lipids Health Dis. 2010;9:137.

[100] Lecerf JM, Luc G, Marecaux N, Bal S, Bonte JP, Lacroix B, et al. Small, qualitative changes in fatty acid intake decrease plasma low-density lipoprotein-cholesterol levels in mildly hypercholesterolemic outpatients on their usual high-fat diets. Int J Food Sci Nutr. 2009;60 Suppl 7:151–63.

[101] Jimenez-Gomez Y, Lopez-Miranda J, Blanco-Colio LM, Marin C, Perez-Martinez P, Ruano J, et al. Olive oil and walnut breakfasts reduce the postprandial inflammatory response in mononuclear cells compared with a butter breakfast in healthy men. Atherosclerosis. 2009;204(2):e70–6.

[102] Nestel PJ, Chronopulos A, Cehun M. Dairy fat in cheese raises LDL cholesterol less than that in butter in mildly hypercholesterolaemic subjects. Eur J Clin Nutr. 2005;59(9):1059–63.

[103] Gorguc M, Celik I. Effects of fresh butter consumption on the lipid profile in healthy human male. Journal of Clinical Biochemistry and Nutrition. 2005;36(3):79–82.

[104] Bellido C, Lopez-Miranda J, Blanco-Colio LM, Perez-Martinez P, Muriana FJ, Martin-Ventura JL, et al. Butter and walnuts, but not olive oil, elicit postprandial activation of nuclear transcription factor kappaB in peripheral blood mononuclear cells from healthy men. Am J Clin Nutr. 2004;80(6):1487–91.

[105] Lichtenstein AH, Erkkila AT, Lamarche B, Schwab US, Jalbert SM, Ausman LM. Influence of hydrogenated fat and butter on CVD risk factors: remnant-like particles, glucose and insulin, blood pressure and C-reactive protein. Atherosclerosis. 2003;171(1):97–107.

[106] Thomsen C, Storm H, Holst JJ, Hermansen K. Differential effects of saturated and monounsaturated fats on postprandial lipemia and glucagon-like peptide 1 responses in patients with type 2 diabetes. Am J Clin Nutr. 2003;77(3):605–11.

[107] Mauger JF, Lichtenstein AH, Ausman LM, Jalbert SM, Jauhiainen M, Ehnholm C, et al. Effect of different forms of dietary hydrogenated fats on LDL particle size. Am J Clin Nutr. 2003;78(3):370–5.

[108] Clemente G, Mancini M, Nazzaro F, Lasorella G, Rivieccio A, Palumbo AM, et al. Effects of different dairy products on postprandial lipemia. Nutr Metab Cardiovasc Dis. 2003;13(6):377–83.

[109] Mekki N, Charbonnier M, Borel P, Leonardi J, Juhel C, Portugal H, et al. Butter differs from olive oil and sunflower oil in its effects on postprandial lipemia and triacylglycerol-rich lipoproteins after single mixed meals in healthy young men. J Nutr. 2002;132(12):3642–9.

[110] Han SN, Leka LS, Lichtenstein AH, Ausman LM, Schaefer EJ, Meydani SN. Effect of hydrogenated and saturated, relative to polyunsaturated, fat on immune and inflammatory responses of adults with moderate hypercholesterolemia. Journal of Lipid Research. 2002;43(3):445–452.

[111] Tonstad S, Strom EC, Bergei CS, Ose L, Christophersen B. Serum cholesterol response to replacing butter with a new trans-free margarine in hypercholesterolemic subjects. Nutr Metab Cardiovasc Dis. 2001;11(5):320–6.

[112] Denke MA, Adams-Huet B, Nguyen AT. Individual cholesterol variation in response to a margarine- or butter-based diet: A study in families. JAMA. 2000;284(21):2740–7.

[113] Nielsen NS, Marckmann P, Hoy C. Effect of meal fat quality on oxidation resistance of postprandial VLDL and LDL particles and plasma triacylglycerol level. Br J Nutr. 2000;84(6):855–63.

[114] Judd JT, Baer DJ, Clevidence BA, Muesing RA, Chen SC, Weststrate JA, et al. Effects of margarine compared with those of butter on blood lipid profiles related to cardiovascular disease risk factors in normolipemic adults fed controlled diets. Am J Clin Nutr. 1998;68(4):768–77.

[115] Cox C, Sutherland W, Mann J, de Jong S, Chisholm A, Skeaff M. Effects of dietary coconut oil, butter and safflower oil on plasma lipids, lipoproteins and lathosterol levels. Eur J Clin Nutr. 1998;52(9):650–4.

[116] Noakes M, Clifton PM. Oil blends containing partially hydrogenated or interesterified fats: differential effects on plasma lipids. Am J Clin Nutr. 1998;68(2):242–7.

[117] Chisholm A, Mann J, Sutherland W, Duncan A, Skeaff M, Frampton C. Effect on lipoprotein profile of replacing butter with margarine in a low fat diet: randomised crossover study with hypercholesterolaemic subjects. BMJ. 1996;312(7036):931–4.

[118] Cox C, Mann J, Sutherland W, Chisholm A, Skeaff M. Effects of coconut oil, butter, and safflower oil on lipids and lipoproteins in persons with moderately elevated cholesterol levels. J Lipid Res. 1995;36(8):1787–95.

[119] Sanders K, Johnson L, O'Dea K, Sinclair AJ. The effect of dietary fat level and quality on plasma lipoprotein lipids and plasma fatty acids in normocholesterolemic subjects. Lipids. 1994;29(2):129–38.

[120] Kris-Etherton PM, Derr J, Mitchell DC, Mustad VA, Russell ME, McDonnell ET, et al. The role of fatty acid saturation on plasma lipids, lipoproteins, and apolipoproteins: I. Effects of whole food diets high in cocoa butter, olive oil, soybean oil, dairy butter, and milk chocolate on the plasma lipids of young men. Metabolism. 1993;42(1):121–9.

[121] Seppanen-Laakso T, Vanhanen H, Laakso I, Kohtamaki H, Viikari J. Replacement of butter on bread by rapeseed oil and rapeseed oil-containing margarine: effects on plasma fatty acid composition and serum cholesterol. Br J Nutr. 1992;68(3):639–54.

[122] Khaw KT, Sharp SJ, Finikarides L, Afzal I, Lentjes M, Luben R, et al. Randomised trial of coconut oil, olive oil or butter on blood lipids and other cardiovascular risk factors in healthy men and women. BMJ open. 2018;8(3):e020167.

[123] Jones KW, Eller LK, Parnell JA, Doyle-Baker PK, Edwards AL, Reimer RA. Effect of a dairy- and calcium-rich diet on weight loss and appetite during energy restriction in overweight and obese adults: a randomized trial. Eur J Clin Nutr. 2013;67(4):371–6.

[124] Wennersberg MH, Smedman A, Turpeinen AM, Retterstol K, Tengblad S, Lipre E, et al. Dairy products and metabolic effects in overweight men and women: results from a 6-mo intervention study. Am J Clin Nutr. 2009;90(4):960–8.

[125] Stancliffe RA, Thorpe T, Zemel MB. Dairy attentuates oxidative and inflammatory stress in metabolic syndrome. Am J Clin Nutr. 2011;94(2):422–30.

[126] van Meijl LE, Mensink RP. Effects of low-fat dairy consumption on markers of low-grade systemic inflammation and endothelial function in overweight and obese subjects: an intervention study. Br J Nutr. 2010;104(10):1523–7.

[127] Nestel PJ, Mellett N, Pally S, Wong G, Barlow CK, Croft K, et al. Effects of low-fat or full-fat fermented and non-fermented dairy foods on selected cardiovascular biomarkers in overweight adults. Br J Nutr. 2013;110(12):2242–9.

[128] Shin H, Yoon YS, Lee Y, Kim CI, Oh SW. Dairy product intake is inversely associated with metabolic syndrome in Korean adults: Anseong and Ansan cohort of the Korean Genome and Epidemiology Study. J Korean Med Sci. 2013;28(10):1482–8.

[129] Ivey KL, Lewis JR, Hodgson JM, Zhu K, Dhaliwal SS, Thompson PL, et al. Association between yogurt, milk, and cheese consumption and common carotid artery intima-media thickness and cardiovascular disease risk factors in elderly women. Am J Clin Nutr. 2011;94(1):234–9.

[130] Pei R, Martin DA, DiMarco DM, Bolling BW. Evidence for the Effects of Yogurt on Gut Health and Obesity. Crit Rev Food Sci Nutr. 2015; p. 0.

[131] Drehmer M, Pereira MA, Schmidt MI, Del Carmen BMM, Alvim S, Lotufo PA, et al. Associations of dairy intake with glycemia and insulinemia, independent of obesity, in Brazilian adults: the Brazilian Longitudinal Study of Adult Health (ELSA-Brasil). Am J Clin Nutr. 2015;101(4):775–82.

[132] Struijk EA, Heraclides A, Witte DR, Soedamah-Muthu SS, Geleijnse JM, Toft U, et al. Dairy product intake in relation to glucose regulation indices and risk of type 2 diabetes. Nutr Metab Cardiovasc Dis. 2013;23(9):822–8.

[133] Wallace TM, Levy JC, Matthews DR. Use and abuse of HOMA modeling. Diabetes Care. 2004;27(6):1487–95.

[134] Bordoni A, Danesi F, Dardevet D, Dupont D, Fernandez AS, Gille D, et al. Dairy Products and Inflammation: A Review of the Clinical Evidence. Crit Rev Food Sci Nutr. 2015; p. 0.

[135] Cho NH, Jang HC, Choi SH, Kim HR, Lee HK, Chan JC, et al. Abnormal liver function test predicts type 2 diabetes: a community-based prospective study. Diabetes Care. 2007;30(10):2566–8.

[136] Abbasi A, Bakker SJ, Corpeleijn E, van der AD, Gansevoort RT, Gans RO, et al. Liver function tests and risk prediction of incident type 2 diabetes: evaluation in two independent cohorts. PLoS One. 2012;7(12):e51496.

[137] Targher G, Marchesini G, Byrne CD. Risk of type 2 diabetes in patients with non-alcoholic fatty liver disease: Causal association or epiphenomenon? Diabetes Metab. 2016;42(3):142–56.

[138] Softic S, Cohen DE, Kahn CR. Role of Dietary Fructose and Hepatic De Novo Lipogenesis in Fatty Liver Disease. Dig Dis Sci. 2016;61(5):1282–93.

[139] Jacobs S, Jager S, Jansen E, Peter A, Stefan N, Boeing H, et al. Associations of Erythrocyte Fatty Acids in the De Novo Lipogenesis Pathway with Proxies of Liver Fat Accumulation in the EPIC-Potsdam Study. PLoS One. 2015;10(5):e0127368.

[140] Alam S, Mustafa G, Alam M, Ahmad N. Insulin resistance in development and progression of nonalcoholic fatty liver disease. World J Gastrointest Pathophysiol. 2016;7(2):211–7.

[141] Alberti K, Eckel RH, Grundy SM, Zimmet PZ, Cleeman JI, Donato KA, et al. Harmonizing the metabolic syndrome a joint interim statement of the international diabetes federation task force on epidemiology and prevention; national heart, lung, and blood institute; American heart association; world heart federation; international atherosclerosis society; and international association for the study of obesity. Circulation. 2009;120(16):1640–1645.

[142] Shin JA, Lee JH, Lim SY, Ha HS, Kwon HS, Park YM, et al. Metabolic syndrome as a predictor of type 2 diabetes, and its clinical interpretations and usefulness. J Diabetes Investig. 2013;4(4):334–43.

[143] Galassi A, Reynolds K, He J. Metabolic syndrome and risk of cardiovascular disease: a meta-analysis. The American journal of medicine. 2006;119(10):812–819.

[144] Kim Y, Je Y. Dairy consumption and risk of metabolic syndrome: a meta-analysis. Diabet Med. 2015;.

[145] Trijsburg L, Geelen A, Hollman PC, Hulshof PJ, Feskens EJ, Van't Veer P, et al. BMI was found to be a consistent determinant related to misreporting of energy, protein and potassium intake using self-report and duplicate portion methods. Public Health Nutr. 2016; p. 1–10.

[146] Jessri M, Lou WY, L'Abbe MR. Evaluation of different methods to handle misreporting in obesity research: evidence from the Canadian national nutrition survey. Br J Nutr. 2016;115(1):147–59.

[147] Murakami K, Livingstone MB. Prevalence and characteristics of misreporting of energy intake in US adults: NHANES 2003-2012. Br J Nutr. 2015;114(8):1294–303.

[148] Lutomski JE, van den Broeck J, Harrington J, Shiely F, Perry IJ. Sociodemographic, lifestyle, mental health and dietary factors associated with direction of misreporting of energy intake. Public Health Nutr. 2011;14(3):532–41.

[149] Nielsen BM, Nielsen MM, Toubro S, Pedersen O, Astrup A, Sorensen TI, et al. Past and current body size affect validity of reported energy intake among middle-aged Danish men. J Nutr. 2009;139(12):2337–43.

[150] Rasmussen LB, Matthiessen J, Biltoft-Jensen A, Tetens I. Characteristics of misreporters of dietary intake and physical activity. Public Health Nutr. 2007;10(3):230–7.

[151] Maurer J, Taren DL, Teixeira PJ, Thomson CA, Lohman TG, Going SB, et al. The psychosocial and behavioral characteristics related to energy misreporting. Nutr Rev. 2006;64(2 Pt 1):53–66.

[152] Lissner L, Troiano RP, Midthune D, Heitmann BL, Kipnis V, Subar AF, et al. OPEN about obesity: recovery biomarkers, dietary reporting errors and BMI. Int J Obes (Lond). 2007;31(6):956–61.

[153] Akter S, Kurotani K, Nanri A, Pham NM, Sato M, Hayabuchi H, et al. Dairy consumption is associated with decreased insulin resistance among the Japanese. Nutr Res. 2013;33(4):286–92.

[154] Da Silva MS, Julien P, Couture P, Lemieux S, Vohl MC, Rudkowska I. Associations between dairy intake and metabolic risk parameters in a healthy French-Canadian population. Appl Physiol Nutr Metab. 2014;39(12):1323–31.

[155] Panagiotakos DB, Pitsavos CH, Zampelas AD, Chrysohoou CA, Stefanadis CI. Dairy products consumption is associated with decreased levels of inflammatory markers related to cardiovascular disease in apparently healthy adults: the ATTICA study. J Am Coll Nutr. 2010;29(4):357–64.

[156] Esmaillzadeh A, Azadbakht L. Dairy consumption and circulating levels of inflammatory markers among Iranian women. Public Health Nutr. 2010;13(9):1395–402.

[157] Engberink MF, Geleijnse JM, de Jong N, Smit HA, Kok FJ, Verschuren WM. Dairy intake, blood pressure, and incident hypertension in a general Dutch population. J Nutr. 2009;139(3):582–7.

[158] Bingham SA, Welch AA, McTaggart A, Mulligan AA, Runswick SA, Luben R, et al. Nutritional methods in the European prospective investigation of cancer in Norfolk. Public health nutrition. 2001;4(03):847–858.

[159] Mulligan AA, Luben RN, Bhaniani A, Parry-Smith DJ, O'Connor L, Khawaja AP, et al. A new tool for converting food frequency questionnaire data into nutrient and food group values: FETA research methods and availability. Bmj Open. 2014;4(3).

[160] O'Connor LM, Lentjes MA, Luben RN, Khaw KT, Wareham NJ, Forouhi NG. Dietary dairy product intake and incident type 2 diabetes: a prospective study using dietary data from a 7-day food diary. Diabetologia. 2014;57(5):909–17.

[161] De Lucia Rolfe E, Sleigh A, Finucane FM, Brage S, Stolk RP, Cooper C, et al. Ultrasound measurements of visceral and subcutaneous abdominal thickness to predict abdominal adiposity among older men and women. Obesity (Silver Spring). 2010;18(3):625–31.

[162] Stolk R, Wink O, Zelissen P, Meijer R, Van Gils A, Grobbee D. Validity and reproducibility of ultrasonography for the measurement of intra-abdominal adipose tissue. International Journal of Obesity and Related Metabolic Disorders. 2001;25(9).

[163] Edens MA, van Ooijen PMA, Post WJ, Haagmans MJF, Kristanto W, Sijens PE, et al. Ultrasonography to Quantify Hepatic Fat Content: Validation by H-1 Magnetic Resonance Spectroscopy. Obesity. 2009;17(12):2239–2244.

[164] 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.

[165] Brage S, Ekelund U, Brage N, Hennings MA, Froberg K, Franks PW, et al. Hierarchy of individual calibration levels for heart rate and accelerometry to measure physical activity. Journal of Applied Physiology. 2007;103(2):682–692.

[166] Livingstone M, Prentice AM, Coward W, Ceesay SM, Strain JJ, McKenna PG, et al. Simultaneous measurement of free-living energy expenditure by the doubly labeled water method and heart-rate monitoring. Am J Clin Nutr. 1990;52(1):59–65.

[167] Wijndaele K, Brage S, Besson H, Khaw KT, Sharp SJ, Luben R, et al. Television viewing time independently predicts all-cause and cardiovascular mortality: the EPIC Norfolk study. Int J Epidemiol. 2010; p. dyq105.

[168] White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. Statistics in medicine. 2011;30(4):377–399.

[169] Eddings W, Marchenko Y. Diagnostics for multiple imputation in Stata. Stata Journal. 2012;12(3):353.

[170] Verardi V, Croux C. Robust regression in Stata. 2008;.

[171] Bozdogan H. Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. Psychometrika. 1987;52(3):345–370.

[172] Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. Annals of statistics. 2001; p. 1165–1188.

[173] Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. The American journal of clinical nutrition. 1997;65(4):1220S–1228S.

[174] Kaess B, Pedley A, Massaro J, Murabito J, Hoffmann U, Fox C. The ratio of visceral to subcutaneous fat, a metric of body fat distribution, is a unique correlate of cardiometabolic risk. Diabetologia. 2012;55(10):2622–2630.

[175] Pallister T, Jackson M, Martin T, Glastonbury C, Jennings A, Beaumont M, et al. Untangling the relationship between diet and visceral fat mass through blood metabolomics and gut microbiome profiling. International journal of obesity. 2017;41(7):1106.

[176] Christen T, Sheikine Y, Rocha VZ, Hurwitz S, Goldfine AB, Di Carli M, et al. Increased glucose uptake in visceral versus subcutaneous adipose tissue revealed by PET imaging. JACC Cardiovasc Imaging. 2010;3(8):843–51.

[177] Zemel MB. Role of dietary calcium and dairy products in modulating adiposity. Lipids. 2003;38(2):139–146.

[178] Hall W, Millward D, Long S, Morgan L. Casein and whey exert different effects on plasma amino acid profiles, gastrointestinal hormone secretion and appetite. British Journal of Nutrition. 2003;89(02):239–248.

[179] Veldhorst MA, Nieuwenhuizen AG, Hochstenbach-Waelen A, van Vught AJ, Westerterp KR, Engelen MP, et al. Dose-dependent satiating effect of whey relative to casein or soy. Physiology and behavior. 2009;96(4-5):675–682.

[180] Alfenas RdCG, Bressan J, Paiva ACd. Effects of protein quality on appetite and energy metabolism in normal weight subjects. Arquivos Brasileiros de Endocrinologia and Metabologia. 2010;54(1):45–51.

[181] Thorning TK, Raben A, Tholstrup T, Soedamah-Muthu SS, Givens I, Astrup A. Milk and dairy products: good or bad for human health? An assessment of the totality of scientific evidence. Food and nutrition research. 2016;60(1):32527.

[182] Rich-Edwards JW, Ganmaa D, Pollak MN, Nakamoto EK, Kleinman K, Tserendolgor U, et al. Milk consumption and the prepubertal somatotropic axis. Nutrition journal. 2007;6(1):28.

[183] Van Vught A, Nieuwenhuizen A, Veldhorst M, Brummer RJ, Westerterp-Plantenga M. The effects of dietary protein on the somatotropic axis: a comparison of soy, gelatin,[alpha]-lactalbumin and milk. European journal of clinical nutrition. 2010;64(5):441.

[184] Carroll PV, Christ the members of Growth Hormone Research Society Scientific Committee ER, Bengtsson BA, Carlsson L, Christiansen J, Clemmons D, et al. Growth hormone deficiency in adulthood and the effects of growth hormone replacement: a review. The Journal of Clinical Endocrinology and Metabolism. 1998;83(2):382–395.

[185] Morio B, Fardet A, Legrand P, Lecerf JM. Involvement of dietary saturated fats, from all sources or of dairy origin only, in insulin resistance and type 2 diabetes. Nutr Rev. 2015;.

[186] Maersk M, Belza A, Stødkilde-Jørgensen H, Ringgaard S, Chabanova E, Thomsen H, et al. Sucrose-sweetened beverages increase fat storage in the liver, muscle, and visceral fat depot: a 6-mo randomized intervention study–. The American journal of clinical nutrition. 2011;95(2):283–289.

[187] Smith JD, Hou T, Hu FB, Rimm EB, Spiegelman D, Willett WC, et al. A comparison of different methods for evaluating diet, physical activity, and long-term weight gain in 3 prospective cohort studies. The Journal of nutrition. 2015;145(11):2527–2534.

[188] Riboli E, Kaaks R. The EPIC Project: rationale and study design. European Prospective Investigation into Cancer and Nutrition. International journal of epidemiology. 1997;26(suppl_1):S6.

[189] Day N, Oakes S, Luben R, Khaw KT, Bingham S, Welch A, et al. EPIC-Norfolk: study design and characteristics of the cohort. European Prospective Investigation of Cancer. Br J Cancer. 1999;80 Suppl 1:95–103.

[190] Rhee JJ, Sampson L, Cho E, Hughes MD, Hu FB, Willett WC. Comparison of methods to account for implausible reporting of energy intake in epidemiologic studies. American journal of epidemiology. 2015; p. kwu308.

[191] Wareham NJ, Jakes RW, Rennie KL, Schuit J, Mitchell J, Hennings S, et al. Validity and repeatability of a simple index derived from the short physical activity questionnaire used in the European Prospective Investigation into Cancer and Nutrition (EPIC) study. Public Health Nutr. 2003;6(04):407–413.

[192] Fitzmaurice GM WJ Laird NM. WHO CDC PREVENTION, Applied longitudinal analysis. Hoboken; 2004.

[193] Lydersen S. Statistical review: frequently given comments. Annals of the rheumatic diseases. 2014; p. annrheumdis–2014–206186.

[194] Glymour MM, Weuve J, Berkman LF, Kawachi I, Robins JM. When is baseline adjustment useful in analyses of change? An example with education and cognitive change. Am J Epidemiol. 2005;162(3):267–78.

[195] Rothman KJ. No adjustments are needed for multiple comparisons. Epidemiology. 1990;1(1):43–6.

[196] Seaman S, White I. Inverse probability weighting with missing predictors of treatment assignment or missingness. Communications in Statistics-Theory and Methods. 2014;43(16):3499–3515.

[197] Agency FS. Food portion sizes. 2002;.

[198] Kinross JM, von Roon AC, Holmes E, Darzi A, Nicholson JK. The human gut microbiome: implications for future health care. Current gastroenterology reports. 2008;10(4):396–403.

[199] Adolfsson O, Meydani SN, Russell RM. Yogurt and gut function. The American Journal of Clinical Nutrition. 2004;80(2):245–256.

[200] Mensink RP, Zock PL, Kester AD, Katan MB. Effects of dietary fatty acids and carbohydrates on the ratio of serum total to HDL cholesterol and on serum lipids and apolipoproteins: a meta-analysis of 60 controlled trials. The American journal of clinical nutrition. 2003;77(5):1146–1155.

[201] Micha R, King IB, Lemaitre RN, Rimm EB, Sacks F, Song X, et al. Food sources of individual plasma phospholipid trans fatty acid isomers: the Cardiovascular Health Study. Am J Clin Nutr. 2010;91(4):883–93.

[202] Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. Epidemiology. 2004;15(5):615–25.

[203] Paeratakul S, Popkin BM, Kohlmeier L, Hertz-Picciotto I, Guo X, Edwards LJ. Measurement error in dietary data: implications for the epidemiologic study of the diet-disease relationship. Eur J Clin Nutr. 1998;52(10):722–7.

[204] Willett W. Nutritional epidemiology. Oxford University Press; 2012.

[205] Fraser GE, Yan R. A multivariate method for measurement error correction using pairs of concentration biomarkers. Ann Epidemiol. 2007;17(1):64–73.

[206] Pfrimer K, Vilela M, Resende CM, Scagliusi FB, Marchini JS, Lima NK, et al. Under-reporting of food intake and body fatness in independent older people: a doubly labelled water study. Age Ageing. 2015;44(1):103–8.

[207] Scalbert A, Brennan L, Manach C, Andres-Lacueva C, Dragsted LO, Draper J, et al. The food metabolome: a window over dietary exposure. Am J Clin Nutr. 2014;99(6):1286–308.

[208] Jenkins B, West JA, Koulman A. A review of odd-chain fatty acid metabolism and the role of pentadecanoic acid (C15: 0) and heptadecanoic acid (C17: 0) in health and disease. Molecules. 2015;20(2):2425–2444.

[209] Tserng KY, Kliegman R, Miettinen EL, Kalhan S. A rapid, simple, and sensitive procedure for the determination of free fatty acids in plasma using glass capillary column gas-liquid chromatography. Journal of lipid research. 1981;22(5):852–858.

[210] Persson XMT, Błachnio-Zabielska AU, Jensen MD. Rapid measurement of plasma free fatty acid concentration and isotopic enrichment using LC/MS. Journal of lipid research. 2010;51(9):2761–2765.

[211] Wolk A, Vessby B, Ljung H, Barrefors P. Evaluation of a biological marker of dairy fat intake. Am J Clin Nutr. 1998;68(2):291–5.

[212] Smedman AE, Gustafsson IB, Berglund LG, Vessby BO. Pentadecanoic acid in serum as a marker for intake of milk fat: relations between intake of milk fat and metabolic risk factors. Am J Clin Nutr. 1999;69(1):22–9.

[213] Wolk A, Furuheim M, Vessby B. Fatty acid composition of adipose tissue and serum lipids are valid biological markers of dairy fat intake in men. J Nutr. 2001;131(3):828–33.

[214] Baylin A, Kabagambe EK, Siles X, Campos H. Adipose tissue biomarkers of fatty acid intake. Am J Clin Nutr. 2002;76(4):750–7.

[215] Brevik A, Veierod MB, Drevon CA, Andersen LF. Evaluation of the odd fatty acids 15:0 and 17:0 in serum and adipose tissue as markers of intake of milk and dairy fat. Eur J Clin Nutr. 2005;59(12):1417–22.

[216] Rosell M, Johansson G, Berglund L, Vessby B, Faire Ud, Hellénius ML. The relation between alcohol intake and physical activity and the fatty acids 14: 0, 15: 0 and 17: 0 in serum phospholipids and adipose tissue used as markers for dairy fat intake. British Journal of Nutrition. 2007;93(1):115–121.

[217] Sun Q, Ma J, Campos H, Hu FB. Plasma and erythrocyte biomarkers of dairy fat intake and risk of ischemic heart disease. Am J Clin Nutr. 2007;86(4):929–37.

[218] Krachler B, Norberg M, Eriksson JW, Hallmans G, Johansson I, Vessby B, et al. Fatty acid profile of the erythrocyte membrane preceding development of Type 2 diabetes mellitus. Nutr Metab Cardiovasc Dis. 2008;18(7):503–10.

[219] Thiébaut AC, Rotival M, Gauthier E, Lenoir GM, Boutron-Ruault MC, Joulin V, et al. Correlation between serum phospholipid fatty acids and dietary intakes assessed a few years earlier. Nutrition and cancer. 2009;61(4):500–509.

[220] Saadatian-Elahi M, Slimani N, Chajès V, Jenab M, Goudable J, Biessy C, et al. Plasma phospholipid fatty acid profiles and their association with food intakes: results from a cross-sectional study within the European Prospective Investigation into Cancer and Nutrition. The American journal of clinical nutrition. 2009;89(1):331–346.

[221] Mozaffarian D, Cao H, King IB, Lemaitre RN, Song X, Siscovick DS, et al. Trans-palmitoleic acid, metabolic risk factors, and new-onset diabetes in U.S. adults: a cohort study. Ann Intern Med. 2010;153(12):790–9.

[222] Aslibekyan S, Campos H, Baylin A. Biomarkers of dairy intake and the risk of heart disease. Nutr Metab Cardiovasc Dis. 2012;22(12):1039–45.

[223] Mozaffarian D, de Oliveira Otto MC, Lemaitre RN, Fretts AM, Hotamisligil G, Tsai MY, et al. trans-Palmitoleic acid, other dairy fat biomarkers, and incident diabetes: the Multi-Ethnic Study of Atherosclerosis (MESA). Am J Clin Nutr. 2013;97(4):854–61.

[224] de Oliveira Otto MC, Nettleton JA, Lemaitre RN, Steffen LM, Kromhout D, Rich SS, et al. Biomarkers of dairy fatty acids and risk of cardiovascular disease in the Multi-ethnic Study of Atherosclerosis. J Am Heart Assoc. 2013;2(4):e000092.

[225] Yakoob MY, Shi P, Hu FB, Campos H, Rexrode KM, Orav EJ, et al. Circulating biomarkers of dairy fat and risk of incident stroke in U.S. men and women in 2 large prospective cohorts. Am J Clin Nutr. 2014;100(6):1437–47.

[226] Santaren ID, Watkins SM, Liese AD, Wagenknecht LE, Rewers MJ, Haffner SM, et al. Serum pentadecanoic acid (15:0), a short-term marker of dairy food intake, is inversely associated with incident type 2 diabetes and its underlying disorders. Am J Clin Nutr. 2014;100(6):1532–40.

[227] Guertin KA, Moore SC, Sampson JN, Huang WY, Xiao Q, Stolzenberg-Solomon RZ, et al. Metabolomics in nutritional epidemiology: identifying metabolites associated with diet and quantifying their potential to uncover diet-disease relations in populations. Am J Clin Nutr. 2014;100(1):208–17.

[228] Forouhi NG, Koulman A, Sharp SJ, Imamura F, Kroger J, Schulze MB, et al. Differences in the prospective association between individual plasma phospholipid saturated fatty acids and incident type 2 diabetes: the EPIC-InterAct case-cohort study. Lancet Diabetes Endocrinol. 2014;2(10):810–8.

[229] Nestel PJ, Straznicky N, Mellett NA, Wong G, De Souza DP, Tull DL, et al. Specific plasma lipid classes and phospholipid fatty acids indicative of dairy food consumption associate with insulin sensitivity. Am J Clin Nutr. 2014;99(1):46–53.

[230] Abdullah MM, Cyr A, Lepine MC, Labonte ME, Couture P, Jones PJ, et al. Recommended dairy product intake modulates circulating fatty acid profile in healthy adults: a multi-centre cross-over study. Br J Nutr. 2015;113(3):435–44.

[231] Albani V, Celis-Morales C, Marsaux CF, Forster H, O'Donovan CB, Woolhead C, et al. Exploring the association of diary product intake with the fatty acids C15:0 and C17:0 measured from dried blood spots in a multi-population cohort: findings from the Food4Me study. Mol Nutr Food Res. 2015;.

[232] Yakoob MY, Shi P, Willett WC, Rexrode KM, Campos H, Orav EJ, et al. Circulating Biomarkers of Dairy Fat and Risk of Incident Diabetes Mellitus Among US Men and Women in Two Large Prospective Cohorts. Circulation. 2016; p. CIRCULATIONAHA. 115.018410.

[233] Weitkunat K, Schumann S, Nickel D, Hornemann S, Petzke KJ, Schulze MB, et al. Odd-chain fatty acids as a biomarker for dietary fiber intake: a novel pathway for endogenous production from propionate, 2. The American journal of clinical nutrition. 2017;105(6):1544–1551.

[234] Jenkins BJ, Seyssel K, Chiu S, Pan PH, Lin SY, Stanley E, et al. Odd Chain Fatty Acids; New Insights of the Relationship Between the Gut Microbiota, Dietary Intake, Biosynthesis and Glucose Intolerance. Scientific Reports. 2017;7:44845.

[235] Laursen ASD, Dahm CC, Johnsen SP, Schmidt EB, Overvad K, Jakobsen MU. Adipose tissue fatty acids present in dairy fat and risk of stroke: the Danish Diet, Cancer and Health cohort. European journal of nutrition. 2018; p. 1–11.

[236] Hodson L, Skeaff CM, Fielding BA. Fatty acid composition of adipose tissue and blood in humans and its use as a biomarker of dietary intake. Progress in lipid research. 2008;47(5):348–380.

[237] Arab L, Akbar J. Biomarkers and the measurement of fatty acids. Public health nutrition. 2002;5(6a):865–871.

[238] Lankinen M, Schwab U. Biomarkers of dairy fat. Am J Clin Nutr. 2015;101(5):1101–2.

[239] Ratnayake WM. Concerns about the use of 15:0, 17:0, and trans-16:1n-7 as biomarkers of dairy fat intake in recent observational studies that suggest beneficial effects of dairy food on incidence of diabetes and stroke. Am J Clin Nutr. 2015;101(5):1102–3.

[240] Siler SQ, Neese RA, Hellerstein MK. De novo lipogenesis, lipid kinetics, and whole-body lipid balances in humans after acute alcohol consumption–. The American journal of clinical nutrition. 1999;70(5):928–936.

[241] Wu Z, Palmquist D. Synthesis and biohydrogenation of fatty acids by ruminal microorganisms in vitro. Journal of Dairy Science. 1991;74(9):3035–3046.

[242] Vlaeminck B, Fievez V, Cabrita A, Fonseca A, Dewhurst R. Factors affecting odd-and branched-chain fatty acids in milk: A review. Animal feed science and technology. 2006;131(3-4):389–417.

[243] Horning MG, Martin DB, Karmen A, Vagelos P. Fatty acid synthesis in adipose tissue. J biol Chem. 1961;236:669–672.

[244] Massart-Leën A, Roets E, Peeters G, Verbeke R. Propionate for fatty acid synthesis by the mammary gland of the lactating goat. Journal of Dairy Science. 1983;66(7):1445–1454.

[245] French E, Bertics S, Armentano L. Rumen and milk odd-and branched-chain fatty acid proportions are minimally influenced by ruminal volatile fatty acid infusions. Journal of dairy science. 2012;95(4):2015–2026.

[246] Jaudszus A, Kramer R, Pfeuffer M, Roth A, Jahreis G, Kuhnt K. trans Palmitoleic acid arises endogenously from dietary vaccenic acid. Am J Clin Nutr. 2014;99(3):431–5.

[247] Djousse L. Is plasma pentadecanoic acid a reasonable biomarker of dairy consumption? J Am Heart Assoc. 2013;2(4):e000393.

[248] Gama M, Bizzo H, Antoniassi R. Analytical shortcomings and other considerations related to the identification of biomarkers of dairy fat intake. European journal of clinical nutrition. 2017;71(8):1022.

[249] Kell DB, Oliver SG. Here is the evidence, now what is the hypothesis? The complementary roles of inductive and hypothesis-driven science in the post-genomic era. Bioessays. 2004;26(1):99–105.

[250] Dunn WB, Broadhurst DI, Atherton HJ, Goodacre R, Griffin JL. Systems level studies of mammalian metabolomes: the roles of mass spectrometry and nuclear magnetic resonance spectroscopy. Chemical Society Reviews. 2011;40(1):387–426.

[251] Herrero M, Simó C, García-Cañas V, Ibáñez E, Cifuentes A. Foodomics: MS-based strategies in modern food science and nutrition. Mass spectrometry reviews. 2012;31(1):49–69.

[252] Guasch-Ferré M, Bhupathiraju SN, Hu FB. Use of Metabolomics in Improving Assessment of Dietary Intake. Clinical chemistry. 2018;64(1):82–98.

[253] Bertram HC, Hoppe C, Petersen BO, Duus JO, Molgaard C, Michaelsen KF. An NMR-based metabonomic investigation on effects of milk and meat protein diets given to 8-year-old boys. Br J Nutr. 2007;97(4):758–63.

[254] Pedersen SM, Nielsen NC, Andersen HJ, Olsson J, Simrén M, Öhman L, et al. The serum metabolite response to diet intervention with probiotic acidified milk in irritable bowel syndrome patients is indistinguishable from that of non-probiotic acidified milk by 1H NMR-based metabonomic analysis. Nutrients. 2010;2(11):1141–1155.

[255] Pedersen SMM, Nebel C, Nielsen NC, Andersen HJ, Olsson J, Simren M, et al. A GC-MS-based metabonomic investigation of blood serum from irritable bowel syndrome patients undergoing intervention with acidified milk products. European Food Research and Technology. 2011;233(6):1013–1021.

[256] Hjerpsted JB, Ritz C, Schou SS, Tholstrup T, Dragsted LO. Effect of cheese and butter intake on metabolites in urine using an untargeted metabolomics approach. Metabolomics. 2014;10(6):1176–1185.

[257] Zheng H, Yde CC, Clausen MR, Kristensen M, Lorenzen J, Astrup A, et al. Metabolomics investigation to shed light on cheese as a possible piece in the French paradox puzzle. J Agric Food Chem. 2015;63(10):2830–9.

[258] Zheng H, Yde CC, Dalsgaard TK, Arnberg K, Mølgaard C, Michaelsen KF, et al. Nuclear magnetic resonance-based metabolomics reveals that dairy protein fractions affect urinary urea excretion differently in overweight adolescents. European Food Research and Technology. 2015;240(3):489–497.

[259] Meikle PJ, Barlow CK, Mellett NA, Mundra PA, Bonham MP, Larsen A, et al. Postprandial Plasma Phospholipids in Men Are Influenced by the Source of Dietary Fat–3. The Journal of nutrition. 2015;145(9):2012–2018.

[260] Zheng H, Lorenzen JK, Astrup A, Larsen LH, Yde CC, Clausen MR, et al. Metabolic effects of a 24-week energy-restricted intervention combined with low or high dairy intake in overweight women: an NMR-based metabolomics investigation. Nutrients. 2016;8(3):108.

[261] Pallister T, Haller T, Thorand B, Altmaier E, Cassidy A, Martin T, et al. Metabolites of milk intake: a metabolomic approach in UK twins with findings replicated in two European cohorts. European journal of nutrition. 2017;56(7):2379–2391.

[262] Munger LH, Trimigno A, Picone G, Freiburghaus C, Pimentel G, Burton KJ, et al. Identification of urinary food intake biomarkers for milk, cheese, and soy-based drink by untargeted GC-MS and NMR in healthy humans. Journal of proteome research. 2017;16(9):3321–3335.

[263] Lees HJ, Swann JR, Wilson ID, Nicholson JK, Holmes E. Hippurate: the natural history of a mammalian–microbial cometabolite. Journal of proteome research. 2013;12(4):1527–1546.

[264] Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. Annals of internal medicine. 2015;162(1):W1–W73.

[265] Lind MV, Savolainen OI, Ross AB. The use of mass spectrometry for analysing metabolite biomarkers in epidemiology: methodological and statistical considerations for application to large numbers of biological samples. European journal of epidemiology. 2016;31(8):717–733.

[266] Wishart DS. Metabolomics: applications to food science and nutrition research. Trends in Food Science and Technology. 2008;19(9):482–493.

[267] Pham-Tuan H, Kaskavelis L, Daykin CA, Janssen HG. Method development in high-performance liquid chromatography for high-throughput profiling and metabonomic studies of biofluid samples. Journal of Chromatography B. 2003;789(2):283–301.

[268] Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, Daykin CA, et al. Proposed minimum reporting standards for chemical analysis. Metabolomics. 2007;3(3):211–221.

[269] Dunn WB, Broadhurst D, Begley P, Zelena E, Francis-McIntyre S, Anderson N, et al. Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. Nat Protoc. 2011;6(7):1060–83.

[270] Trygg J, Gullberg J, Johansson AI, Jonsson P, Antti H, Marklund SL, et al. Extraction and GC/MS analysis of the human blood plasma metabolome. Analytical Chemistry. 2005;77(24):8086–8094.

[271] Wang LY, Summerhill K, Rodriguez-Canas C, Mather I, Patel P, Eiden M, et al. Development and validation of a robust automated analysis of plasma phospholipid fatty acids for metabolic phenotyping of large epidemiological studies. Genome medicine. 2013;5(4):39.

[272] IUPAC I. IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN). Nomenclature and symbolism for amino acids and peptides. Recommendations 1983. Biochem J. 1984;219(2):345–373.

[273] van den Berg RA, Hoefsloot HC, Westerhuis JA, Smilde AK, van der Werf MJ. Centering, scaling, and transformations: improving the biological information content of metabolomics data. BMC genomics. 2006;7(1):142.

[274] Saccenti E, Hoefsloot HC, Smilde AK, Westerhuis JA, Hendriks MM. Reflections on univariate and multivariate analysis of metabolomics data. Metabolomics. 2014;10(3):0.

[275] Hendriks MM, van Eeuwijk FA, Jellema RH, Westerhuis JA, Reijmers TH, Hoefsloot HC, et al. Data-processing strategies for metabolomics studies. TrAC Trends in Analytical Chemistry. 2011;30(10):1685–1698.

[276] Alsberg BK, Goodacre R, Rowland JJ, Kell D. Classification of pyrolysis mass spectra by fuzzy multivariate rule induction-comparison with regression, K-nearest neighbour, neural and decision-tree methods. Analytica Chimica Acta. 1997;348(1):389–407.

[277] Bryan K, Brennan L, Cunningham P. MetaFIND: A feature analysis tool for metabolomics data. BMC bioinformatics. 2008;9(1):470.

[278] Fan Y, Murphy TB, Byrne JC, Brennan L, Fitzpatrick JM, Watson RWG. Applying random forests to identify biomarker panels in serum 2D-DIGE data for the detection and staging of prostate cancer. Journal of proteome research. 2011;10(3):1361–1373.

[279] Touw WG, Bayjanov JR, Overmars L, Backus L, Boekhorst J, Wels M, et al. Data mining in the Life Sciences with Random Forest: a walk in the park or lost in the jungle? Briefings in Bioinformatics. 2012; p. bbs034.

[280] Acharjee A, Finkers R, Visser RG, Maliepaard C. Comparison of regularized regression methods for omics data. Metabolomics. 2013;3(126):2153–0769.1000126.

[281] Chen T, Cao Y, Zhang Y, Liu J, Bao Y, Wang C, et al. Random forest in clinical metabolomics for phenotypic discrimination and biomarker selection. Evidence-Based Complementary and Alternative Medicine. 2013;2013.

[282] Hoerl AE, Kennard RW. Ridge regression: Biased estimation for nonorthogonal problems. Technometrics. 1970;12(1):55–67.

[283] Tibshirani R. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society Series B (Methodological). 1996; p. 267–288.

[284] Kyung M, Gill J, Ghosh M, Casella G. Penalized regression, standard errors, and Bayesian lassos. Bayesian Analysis. 2010;5(2):369–411.

[285] Zou H, Hastie T. Regression shrinkage and selection via the elastic net, with applications to microarrays. JR Stat Soc Ser B. 2003;67:301–20.

[286] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine learning in Python. Journal of machine learning research. 2011;12(Oct):2825–2830.

[287] Kluyver T, Ragan-Kelley B, Pérez F, Granger BE, Bussonnier M, Frederic J, et al. Jupyter Notebooks-a publishing format for reproducible computational workflows. In: ELPUB;. p. 87–90.

[288] Brennan L. The nutritional metabolomics crossroads: how to ensure success for dietary biomarkers; 2017.

[289] Kröger J, Zietemann V, Enzenbach C, Weikert C, Jansen EH, Döring F, et al. Erythrocyte membrane phospholipid fatty acids, desaturase activity, and dietary fatty acids in relation to risk of type 2 diabetes in the European Prospective Investigation into Cancer and Nutrition–Potsdam Study. Am J Clin Nutr. 2010; p. ajcn. 005447.

[290] Patel PS, Sharp SJ, Jansen E, Luben RN, Khaw KT, Wareham NJ, et al. Fatty acids measured in plasma and erythrocyte-membrane phospholipids and derived by food-frequency questionnaire and the risk of new-onset type 2 diabetes: a pilot study in the European Prospective Investigation into Cancer and Nutrition (EPIC)–Norfolk cohort. Am J Clin Nutr. 2010; p. ajcn. 29182.

[291] Hodge AM, English DR, O'dea K, Sinclair AJ, Makrides M, Gibson RA, et al. Plasma phospholipid and dietary fatty acids as predictors of type 2 diabetes: interpreting the role of linoleic acid–. The American journal of clinical nutrition. 2007;86(1):189–197.

[292] Imamura F, Fretts A, Marklund M, Ardisson Korat AV, Yang WS, Lankinen M, et al. Fatty acid biomarkers of dairy fat consumption and incidence of type 2 diabetes: a pooled analysis of prospective cohort studies. PLoS Med. 2018; p. (in press)–(in press).

[293] Lee HS, Xu T, Lee Y, Kim NH, Kim YJ, Kim JM, et al. Identification of putative biomarkers for type 2 diabetes using metabolomics in the Korea Association REsource (KARE) cohort. Metabolomics. 2016;12(12):178.

[294] Wang-Sattler R, Yu Z, Herder C, Messias AC, Floegel A, He Y, et al. Novel biomarkers for pre-diabetes identified by metabolomics. Mol Syst Biol. 2012;8:615.

[295] Shi L, Brunius C, Lehtonen M, Auriola S, Bergdahl IA, Rolandsson O, et al. Plasma metabolites associated with type 2 diabetes in a Swedish population: a case–control study nested in a prospective cohort. Diabetologia. 2018; p. 1–13.

[296] Floegel A, Stefan N, Yu Z, Muhlenbruch K, Drogan D, Joost HG, et al. Identification of serum metabolites associated with risk of type 2 diabetes using a targeted metabolomic approach. Diabetes. 2013;62(2):639–48.

[297] Fall T, Salihovic S, Brandmaier S, Nowak C, Ganna A, Gustafsson S, et al. Non-targeted metabolomics combined with genetic analyses identifies bile acid synthesis and phospholipid metabolism as being associated with incident type 2 diabetes. Diabetologia. 2016;59(10):2114–24.

[298] van den Berg SA, Guigas B, Bijland S, Ouwens M, Voshol PJ, Frants RR, et al. High levels of dietary stearate promote adiposity and deteriorate hepatic insulin sensitivity. Nutr Metab (Lond). 2010;7:24.

[299] Reynoso R, Salgado LM, Calderon V. High levels of palmitic acid lead to insulin resistance due to changes in the level of phosphorylation of the insulin receptor and insulin receptor substrate-1. Mol Cell Biochem. 2003;246(1-2):155–62.

[300] Anderson EK, Hill AA, Hasty AH. Stearic acid accumulation in macrophages induces toll-like receptor 4/2-independent inflammation leading to endoplasmic reticulum stress-mediated apoptosis. Arterioscler Thromb Vasc Biol. 2012;32(7):1687–95.

[301] Weigert C, Brodbeck K, Staiger H, Kausch C, Machicao F, Haring HU, et al. Palmitate, but not unsaturated fatty acids, induces the expression of interleukin-6 in human myotubes through proteasome-dependent activation of nuclear factor-kappaB. J Biol Chem. 2004;279(23):23942–52.

[302] Postic C, Girard J. Contribution of de novo fatty acid synthesis to hepatic steatosis and insulin resistance: lessons from genetically engineered mice. J Clin Invest. 2008;118(3):829–38.

[303] Corgneau M, Scher J, Ritie-Pertusa L, Le Dt, Petit J, Nikolova Y, et al. Recent advances on lactose intolerance: Tolerance thresholds and currently available answers. Critical reviews in food science and nutrition. 2017;57(15):3344–3356.

[304] Yang Q, Lin SL, Au Yeung SL, Kwok MK, Xu L, Leung GM, et al. Genetically predicted milk consumption and bone health, ischemic heart disease and type 2 diabetes: a Mendelian randomization study. Eur J Clin Nutr. 2017;.

[305] Bersaglieri T, Sabeti PC, Patterson N, Vanderploeg T, Schaffner SF, Drake JA, et al. Genetic signatures of strong recent positive selection at the lactase gene. Am J Hum Genet. 2004;74(6):1111–20.

[306] Gerbault P, Liebert A, Itan Y, Powell A, Currat M, Burger J, et al. Evolution of lactase persistence: an example of human niche construction. Philos Trans R Soc Lond B Biol Sci. 2011;366(1566):863–77.

[307] Ingram CJ, Mulcare CA, Itan Y, Thomas MG, Swallow DM. Lactose digestion and the evolutionary genetics of lactase persistence. Hum Genet. 2009;124(6):579–91.

[308] Hartwig FP, Horta BL, Smith GD, de Mola CL, Victora CG. Association of lactase persistence genotype with milk consumption, obesity and blood pressure: a Mendelian randomization study in the 1982 Pelotas (Brazil) Birth Cohort, with a systematic review and meta-analysis. International journal of epidemiology. 2016; p. dyw074.

[309] Savaiano DA. Lactose digestion from yogurt: mechanism and relevance. Am J Clin Nutr. 2014;99(5 Suppl):1251S–5S.

[310] Corella D, Arregui M, Coltell O, Portoles O, Guillem-Saiz P, Carrasco P, et al. Association of the LCT-13910C>T polymorphism with obesity and its modulation by dairy products in a Mediterranean population. Obesity (Silver Spring). 2011;19(8):1707–14.

[311] Pimentel G, Burton KJ, Rosikiewicz M, Freiburghaus C, von Ah U, Münger LH, et al. Blood lactose after dairy product intake in healthy men. British Journal of Nutrition. 2017;118(12):1070–1077.

[312] Burgess S, Thompson SG. Mendelian randomization: methods for using genetic variants in causal estimation. CRC Press; 2015.

[313] Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. Statistics in medicine. 2008;27(8):1133–1163.

[314] Bergholdt HK, Nordestgaard BG, Ellervik C. Milk intake is not associated with low risk of diabetes or overweight-obesity: a Mendelian randomization study in 97,811 Danish individuals. Am J Clin Nutr. 2015;102(2):487–96.

[315] Bergholdt HK, Nordestgaard BG, Varbo A, Ellervik C. Milk intake is not associated with ischaemic heart disease in observational or Mendelian randomization analyses in 98,529 Danish adults. Int J Epidemiol. 2015;44(2):587–603.

[316] Biobank U. Genotyping and quality control of UK Biobank, a large-scale, extensively phenotypes prospective resource. 2015;.

[317] Biobank U. Genotype imputation and genetic association studies of UK Biobank. 2015;.

[318] Liu B, Young H, Crowe FL, Benson VS, Spencer EA, Key TJ, et al. Development and evaluation of the Oxford WebQ, a low-cost, web-based method for assessment of previous 24 h dietary intakes in large-scale prospective studies. Public health nutrition. 2011;14(11):1998–2005.

[319] Loh PR, Kichaev G, Gazal S, Schoech AP, Price AL. Mixed-model association for biobank-scale datasets. Nature genetics. 2018; p. 1.

[320] Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Res. 2012;40(Database issue):D930–4.

[321] Arnold M, Raffler J, Pfeufer A, Suhre K, Kastenmüller G. SNiPA: an interactive, genetic variant-centered annotation browser. Bioinformatics. 2015;31(8):1334–1336.

[322] Sotos-Prieto M, Guillen M, Guillem-Saiz P, Portoles O, Corella D. The rs1466113 polymorphism in the somatostatin receptor 2 gene is associated with obesity and food intake in a Mediterranean population. Ann Nutr Metab. 2010;57(2):124–31.

[323] McCaffery JM, Papandonatos GD, Peter I, Huggins GS, Raynor HA, Delahanty LM, et al. Obesity susceptibility loci and dietary intake in the Look AHEAD Trial. Am J Clin Nutr. 2012;95(6):1477–86.

[324] Pirastu N, Robino A, Lanzara C, Athanasakis E, Esposito L, Tepper BJ, et al. Genetics of food preferences: a first view from silk road populations. J Food Sci. 2012;77(12):S413–8.

[325] Pirastu N, Kooyman M, Traglia M, Robino A, Willems SM, Pistis G, et al. A Genome-Wide Association Study in isolated populations reveals new genes associated to common food likings. Rev Endocr Metab Disord. 2016;17(2):209–19.

[326] de Oliveira Otto MC, Lemaitre RN, Sun Q, King IB, Wu JHY, Manichaikul A, et al. Genome-wide association meta-analysis of circulating odd-numbered chain saturated fatty acids: Results from the CHARGE Consortium. PLoS One. 2018;13(5):e0196951.

[327] Garcia-Etxebarria K, Jauregi-Miguel A, Romero-Garmendia I, Plaza-Izurieta L, Legarda M, Irastorza I, et al. Ancestry-based stratified analysis of Immunochip data identifies novel associations with celiac disease. Eur J Hum Genet. 2016;24(12):1831–1834.

[328] Murray JA. The widening spectrum of celiac disease. The American Journal of Clinical Nutrition. 1999;69(3):354–365.

[329] Beulens JW, Grobbee DE, Sluijs I, Spijkerman AM, Van Der Schouw YT. Dietary phylloquinone and menaquinones intakes and risk of type 2 diabetes. Diabetes care. 2010;33(8):1699–1705.

[330] Gast GCM, de Roos NM, Sluijs I, Bots ML, Beulens JW, Geleijnse JM, et al. A high menaquinone intake reduces the incidence of coronary heart disease. Nutrition, Metabolism and Cardiovascular Diseases. 2009;19(7):504–510.

[331] Miller PE, Alexander DD, Perez V. Effects of whey protein and resistance exercise on body composition: a meta-analysis of randomized controlled trials. Journal of the American College of Nutrition. 2014;33(2):163–175.

[332] Takahira M, Noda K, Fukushima M, Zhang B, Mitsutake R, Uehara Y, et al. Randomized, double-blind, controlled, comparative trial of formula food containing soy protein vs. milk protein in visceral fat obesity. -FLAVO study. Circ J. 2011;75(9):2235–43.

[333] Boirie Y, Dangin M, Gachon P, Vasson MP, Maubois JL, Beaufrère B. Slow and fast dietary proteins differently modulate postprandial protein accretion. Proceedings of the National Academy of Sciences. 1997;94(26):14930–14935.

[334] Giezenaar C, Trahair LG, Luscombe-Marsh ND, Hausken T, Standfield S, Jones KL, et al. Effects of randomized whey-protein loads on energy intake, appetite, gastric emptying, and plasma gut-hormone concentrations in older men and women. The American journal of clinical nutrition. 2017;106(3):865–877.

[335] Lorenzen J, Frederiksen R, Hoppe C, Hvid R, Astrup A. The effect of milk proteins on appetite regulation and diet-induced thermogenesis. European journal of clinical nutrition. 2012;66(5):622–627.

[336] Law M, Lee YT, Vien S, Luhovyy BL, Anderson GH. The effect of dairy products consumed with high glycemic carbohydrate on subjective appetite, food intake, and postprandial glycemia in older adults. Applied Physiology, Nutrition, and Metabolism. 2017;42(11):1210–1216.

[337] Trowman R, Dumville JC, Hahn S, Torgerson DJ. A systematic review of the effects of calcium supplementation on body weight. British Journal of Nutrition. 2006;95(06):1033–1038.

[338] Christensen R, Lorenzen JK, Svith CR, Bartels E, Melanson E, Saris W, et al. Effect of calcium from dairy and dietary supplements on faecal fat excretion: a meta-analysis of randomized controlled trials. Obesity reviews. 2009;10(4):475–486.

[339] Zemel MB, Shi H, Greer B, Dirienzo D, Zemel PC. Regulation of adiposity by dietary calcium. The FASEB Journal. 2000;14(9):1132–1138.

[340] Zemel M, Richards J, Mathis S, Milstead A, Gebhardt L, Silva E. Dairy augmentation of total and central fat loss in obese subjects. International journal of obesity. 2005;29(4):391–397.

[341] Bovee-Oudenhoven I, Termont D, Weerkamp AH, Faassen-Peters M, Van der Meer R. Dietary calcium inhibits the intestinal colonization and translocation of Salmonella in rats. Gastroenterology. 1997;113(2):550–557.

[342] Schepens MA, ten Bruggencate SJ, Schonewille AJ, Brummer RJM, van der Meer R, Bovee-Oudenhoven IM. The protective effect of supplemental calcium on colonic permeability depends on a calcium phosphate-induced increase in luminal buffering capacity. British Journal of Nutrition. 2012;107(7):950–956.

[343] Whigham LD, Watras AC, Schoeller DA. Efficacy of conjugated linoleic acid for reducing fat mass: a meta-analysis in humans. Am J Clin Nutr. 2007;85(5):1203–1211.

[344] Brown JM, Boysen MS, Jensen SS, Morrison RF, Storkson J, Lea-Currie R, et al. Isomer-specific regulation of metabolism and PPARγ signaling by CLA in human preadipocytes. Journal of lipid research. 2003;44(7):1287–1300.

[345] Brown JM, Boysen MS, Chung S, Fabiyi O, Morrison RF, Mandrup S, et al. Conjugated Linoleic Acid Induces Human Adipocyte Delipidation Autocrine/Paracrine regulation of MEK/ERK signaling by adipocytokines. Journal of Biological Chemistry. 2004;279(25):26735–26747.

[346] Knapen M, Jardon K, Vermeer C. Vitamin K-induced effects on body fat and weight: results from a 3-year vitamin K2 intervention study. European journal of clinical nutrition. 2018;72(1):136.

[347] Rosqvist F, Smedman A, Lindmark-Månsson H, Paulsson M, Petrus P, Straniero S, et al. Potential role of milk fat globule membrane in modulating plasma lipoproteins, gene expression, and cholesterol metabolism in humans: a randomized study. The American journal of clinical nutrition. 2015;102(1):20–30.

[348] Sjogren P, Rosell M, Skoglund-Andersson C, Zdravkovic S, Vessby B, de Faire U, et al. Milk-derived fatty acids are associated with a more favorable LDL particle size distribution in healthy men. The Journal of nutrition. 2004;134(7):1729–1735.

[349] Motard-Belanger A, Charest A, Grenier G, Paquin P, Chouinard Y, Lemieux S, et al. Study of the effect of trans fatty acids from ruminants on blood lipids and other risk factors for cardiovascular disease. Am J Clin Nutr. 2008;87(3):593–9.

[350] Conway V, Couture P, Richard C, Gauthier S, Pouliot Y, Lamarche B. Impact of buttermilk consumption on plasma lipids and surrogate markers of cholesterol homeostasis in men and women. Nutrition, Metabolism and Cardiovascular Diseases. 2013;23(12):1255–1262.

[351] Guo Z, Liu X, Zhang Q, Shen Z, Tian F, Zhang H, et al. Influence of consumption of probiotics on the plasma lipid profile: a meta-analysis of randomised controlled trials. Nutrition, Metabolism and Cardiovascular Diseases. 2011;21(11):844–850.

[352] Pittas AG, Lau J, Hu FB, Dawson-Hughes B. The role of vitamin D and calcium in type 2 diabetes. A systematic review and meta-analysis. The Journal of Clinical Endocrinology and Metabolism. 2007;92(6):2017–2029.

[353] Autier P, Mullie P, Macacu A, Dragomir M, Boniol M, Coppens K, et al. Effect of vitamin D supplementation on non-skeletal disorders: a systematic review of meta-analyses and randomised trials. The Lancet Diabetes and Endocrinology. 2017;5(12):986–1004.

[354] Wang X, Juan QF, He YW, Zhuang L, Fang YY, Wang YH. Multiple effects of probiotics on different types of diabetes: a systematic review and meta-analysis of randomized, placebo-controlled trials. Journal of Pediatric Endocrinology and Metabolism. 2017;30(6):611–622.

[355] Chung S, Brown JM, Provo JN, Hopkins R, McIntosh MK. Conjugated linoleic acid promotes human adipocyte insulin resistance through NFκB-dependent cytokine production. Journal of Biological Chemistry. 2005;280(46):38445–38456.

[356] Wang X, Gupta J, Kerslake M, Rayat G, Proctor SD, Chan CB. Trans-11 vaccenic acid improves insulin secretion in models of type 2 diabetes in vivo and in vitro. Molecular nutrition and food research. 2016;60(4):846–857.

[357] Kennedy A, Overman A, LaPoint K, Hopkins R, West T, Chuang CC, et al. Conjugated linoleic acid-mediated inflammation and insulin resistance in human adipocytes are attenuated by resveratrol. Journal of lipid research. 2009;50(2):225–232.

[358] Fekete AA, Givens DI, Lovegrove JA. Casein-derived lactotripeptides reduce systolic and diastolic blood pressure in a meta-analysis of randomised clinical trials. Nutrients. 2015;7(1):659–681.

[359] Nakamura Y, Yamamoto N, Sakai K, Okubo A, Yamazaki S, Takano T. Purification and characterization of angiotensin I-converting enzyme inhibitors from sour milk. Journal of dairy science. 1995;78(4):777–783.

[360] Van Mierlo L, Arends L, Streppel M, Zeegers M, Kok F, Grobbee D, et al. Blood pressure response to calcium supplementation: a meta-analysis of randomized controlled trials. Journal of human hypertension. 2006;20(8):571–580.

[361] Aburto NJ, Hanson S, Gutierrez H, Hooper L, Elliott P, Cappuccio FP. Effect of increased potassium intake on cardiovascular risk factors and disease: systematic review and meta-analyses. 2013;.

[362] Jee SH, Miller ER, Guallar E, Singh VK, Appel LJ, Klag MJ. The effect of magnesium supplementation on blood pressure: a meta-analysis of randomized clinical trials. American Journal of Hypertension. 2002;15(8):691–696.

[363] Rothman KJ, Greenland S, Lash TL. Modern epidemiology. 2008;.

[364] Hebert JR, Clemow L, Pbert L, Ockene IS, Ockene JK. Social desirability bias in dietary self-report may compromise the validity of dietary intake measures. International journal of epidemiology. 1995;24(2):389–398.

Appendix A

Supplemental results

This Appendix includes supplemental tables and figures related to results from

- Chapter 2 (section A.1, pages 285-288)
- Chapters 4 (section A.2, pages 289-304)
- Chapters 5 (section A.3, pages 299-316)
- Chapters 6 (section A.4, pages 317-357)
- Chapters 8 (section A.5, pages 358-359)

A.1 Chapter 2

The following tables include descriptive statistics (mean, standard error, median and IQR) of total and types of dairy consumption among adults <65 years and elderly participants $\geq$65 years across the old years and the years of the rolling programme of the National Diet and Nutrition Survey.

Table A.1 Means and SE of total and types of dairy products within consumers and percentage of consumers among adults <65 years and ≥65 years in the old survey years and the first 8 years of the rolling programme of the National Diet and Nutrition Survey (NDNS) (2008-2016)

	Survey year	**Old survey †**			**2008/2009**			**2009/2010**			**2010/2011**			**2011/2012**			**2012/2013**			**2013/2014**			**2014/2015**			**2015/2016**		
<65 years	**N**	1,723			488			420			444			512			535			546			556			597		
Dairy products (g/d)		**% ‡**	**Mean**	**SE**	**% ‡**	**Mean**	**SE**	**% ‡**	**Mean**	**SE**	**% ‡**	**Mean**	**SE**	**% ‡**	**Mean**	**SE**	**% ‡**	**Mean**	**SE**	**% ‡**	**Mean**	**SE**	**% ‡**	**Mean**	**SE**	**% ‡**	**Mean**	**SE**
Milk	Full-fat	80	68.8	3.8	77	55.3	7.2	79.5	46	5.3	72.5	47.1	4.6	75	35.4	4.8	78.5	38	5.8	78	45.1	7	75.2	41.4	8	76.2	46.7	5.8
	Low-fat	99.2	184.8	5	99.8	150	7.4	98.8	153.2	7.3	98	141.7	8.5	100	152.5	7.6	98.1	159.5	7.4	98.4	167	8.8	98.4	151.8	7.6	97.2	159	9.6
	Total	99.7	239.3	5.1	100	192.3	7.8	99.8	188.2	8.7	99.3	174.4	8.7	100	179.2	7.8	99.3	186.7	7.3	99.3	199.3	9.4	99.3	181.8	7.5	98.7	193.8	9.8
Yoghurt	Full-fat	27.5	24	2.3	29.9	32.1	3.9	29	29.8	3	32	25.9	2.5	30.9	28	3.1	29.7	30.1	3.6	30	32	4.6	30.2	26.3	2.8	31.3	29.3	2.8
	Low-fat	30.5	36	1.8	29.3	33.3	3.2	27.1	38.5	3.7	32.9	35.5	4.8	31.1	37.1	2.8	28.2	40.1	3.8	32.6	32.7	4.1	33.5	33	3	31.3	43.5	4.6
	Total	43.5	40.4	2	46.1	42.7	3.3	44.5	42.4	3	49.3	40.7	4.2	48.4	41.3	2.5	46.4	43.1	3.5	48.5	43.2	4	49.5	37.8	2.7	48.1	45.4	3.7
Cheese	High-fat	78.4	21.8	0.6	76.8	25.3	1.5	79.5	24.8	1.7	78.6	24.4	1.4	78.7	23.4	1.1	80.7	24.8	1.3	80.8	24.4	1.4	82.6	23.6	1.2	80.9	21.9	1.3
	Low-fat	13.6	20.6	1.7	11.1	16.1	3.2	8.6	19.4	3.6	11.7	15.5	4	10.7	13.5	2.6	10.8	15.5	2	11.4	17.6	4.5	9.5	17	3.9	13.7	13.3	1.7
	Total	79.9	24.8	0.7	77.9	27.5	1.5	80.7	26.4	1.8	80	26.3	1.7	79.7	24.8	1.2	81.9	26.2	1.3	82.2	26.6	1.7	83.5	25.4	1.4	81.6	24.1	1.4
Cream		81.5	6.4	0.3	78.1	5.5	0.5	75.5	7.3	0.9	72.5	6.8	0.7	73.4	7	0.6	72.1	6.6	0.7	72.2	6.4	0.6	73	6.6	0.7	74	5.7	0.4
Butter		82.6	7.2	0.3	81.6	6.4	0.5	77.4	8.2	0.6	80.9	7.2	0.5	80.3	6.9	0.5	82.1	8	0.5	81.5	6.8	0.5	84.5	7.6	0.6	83.1	8.1	0.5
Fermented dairy products §		86.5	43.2	1.3	85.9	47.1	2.3	86.9	46.4	2.6	86.9	47	2.9	88.3	45	2	89	46.7	2.7	88.3	48.4	2.7	89.6	43.8	2.5	88.8	47	2.5
Total dairy products	High-fat ‖	98.5	85.1	3.2	98.2	73.2	5.9	98.6	68.8	4.9	97.5	65.6	3.3	97.5	57.6	4.2	97.9	63.1	4.7	97.4	65.5	5.3	98.2	63	6.4	97.8	66.7	4.9
	Low-fat #	99.3	205.3	5.2	99.8	171.1	8.2	98.8	174	7.6	98.6	162.5	9.9	100	173.6	8.3	98.9	178.4	7.7	98.7	189.4	9.4	98.9	170.7	7.8	98.2	181.8	9.9
	Total *	99.9	287.8	5.3	100	242.6	9	99.8	240.5	9.2	100	224.2	10.4	100	229.6	9	99.8	239.7	7.9	99.8	249.9	10.4	99.8	231.1	8	99.8	245	10.3
≥65 years	**N**	1,733			124			116			87			131			160			178			184			151		
		% ‡	**Mean**	**SE**	**% ‡**	**Mean**	**SE**	**% ‡**	**Mean**	**SE**	**% ‡**	**Mean**	**SE**	**% ‡**	**Mean**	**SE**	**% ‡**	**Mean**	**SE**	**% ‡**	**Mean**	**SE**	**% ‡**	**Mean**	**SE**	**% ‡**	**Mean**	**SE**
Milk	Full-fat	87.9	165.6	6	76.6	114.2	19.4	75	78.1	14.3	67.8	49.8	13.6	75.6	40.5	9.7	79.4	42.1	8.5	78.1	37.5	6.8	71.7	57.9	10.8	73.5	33.1	7.8
	Low-fat	98.9	173.9	5.3	98.4	164.6	15.3	100	208.4	18.7	97.7	179.5	12.9	100	219.7	16.4	98.8	201	11.8	98.3	184.5	14.9	100	188	13.9	99.3	204.3	15.6
	Total	99.9	315.9	5.3	99.2	252.5	16.9	100	265.8	16.8	100	208.4	14	100	250.2	16.8	98.8	235	12.5	99.4	212.7	14.6	100	228.6	13.1	100	227.9	14.8
Yoghurt	Full-fat	13.8	22.5	1.9	25.8	30.3	4.9	31.9	36.5	8.8	41.4	46.4	6.7	27.5	30.5	7.7	21.9	30	5.9	26.4	33.1	5.6	26.6	30.6	4.3	29.8	45.6	10.2
	Low-fat	13.5	44.8	2.9	23.4	44.5	8.5	31	32.8	5.5	25.3	44.4	9.3	32.1	39.7	4.1	31.9	44.1	4.7	31.5	41.9	5.3	29.3	49.8	9.7	34.4	40.4	6.3
	Total	19.5	46.9	2.5	39.5	46.6	8	47.4	46.3	6	50.6	60.7	8.2	46.6	44.7	5.2	39.4	53.4	5.4	44.9	48.7	5.6	44	51	7.3	46.4	64	9.2
Cheese	High-fat	69.6	14.6	0.4	68.5	22.7	3	71.6	19.7	2	77	21.7	2.8	72.5	19.7	1.7	72.5	19.4	1.4	73	17.7	1.3	77.7	20.2	1.7	75.5	19.3	2
	Low-fat	7	15.4	2	12.9	21.2	4.9	6.9	14	4.1	5.7	22.8	5.2	8.4	4.9	1.5	9.4	6.1	1.2	10.7	18.1	5.9	7.6	14.4	4.8	6.6	12.1	3.4
	Total	71.3	16	0.5	71.8	25.1	3.2	74.1	20.4	2	77	23.3	2.8	73.3	20.2	1.7	73.8	19.8	1.3	75.3	19.8	1.7	78.8	21.7	1.9	76.8	19.9	2.1
Cream		83.8	4.5	0.2	75	6.9	1.2	81.9	8.9	1.4	78.2	5.9	1	77.9	6.9	1.1	79.4	8.1	1.1	77.5	9.4	2.1	71.7	7.1	1.2	75.5	7.4	1.1
Butter		84.8	10.6	0.3	84.7	8.1	1	86.2	11.2	1.1	85.1	10.5	1.3	80.2	10.3	1.1	82.5	8.4	0.8	84.3	7.9	1.1	82.6	11.1	1.1	86.1	7.9	0.8
Fermented dairy products §		74.6	28.7	1.2	81.5	46.2	4.8	81.9	44.8	5.7	85.1	57.6	5.9	81.7	44.3	4	80.6	44.4	3.9	82.6	45	3.1	84.2	47.6	4.5	82.8	52.6	5.9
Total dairy products	High-fat ‖	99.5	167.6	5.5	96.8	119.1	16.5	98.3	91	11.1	97.7	65.2	10.4	98.5	60.1	8	98.8	61.7	7.6	98.9	56.2	6.3	98.9	71.7	8.5	98.7	53.1	6.2
	Low-fat #	99	185.3	5.4	99.2	186.3	14.5	100	231	19.8	97.7	212.5	12.3	100	241.7	17.3	98.8	223.5	12.8	98.3	209.3	15.3	100	212.1	14.2	100	232.9	16.1
	Total *	100	350.1	5.4	99.2	302.8	15.7	100	320.6	18	100	271	14.6	100	300.5	18.8	99.4	282.4	13.6	100	262.6	14.4	100	282.4	13.6	100	285.9	15.1

†Old survey years are 2000/2001 for adults <65 years and 1994/1995 for adults ≥ 65 years
‡Percentage of consumers
§Sum of yoghurt and cheese
‖Sum of full-fat milk, high-fat cheese, total cream and butter
#Sum of low-fat milk, total yoghurt, and low-fat cheese
*Sum of total milk, total yoghurt, total cheese, total cream and butter

Table A.2 Median and inter-quartile range (IQR) of total and types of dairy products within consumers and percentage of non-consumers among adults <65 years in the year 2000/2001 and the first 8 years of the rolling programme of the National Diet and Nutrition Survey (NDNS) (2008-2016)

	Survey year	**Old survey †**			**2008/2009**			**2009/2010**			**2010/2011**			**2011/2012**			**2012/2013**			**2013/2014**			**2014/2015**			**2015/2016**		
<65 years	**N**	1,723			488			420			444			512			535			546			556			597		
Dairy products (g/d)		**% ‡**	**Median**	**IQR**	**% ‡**	**Median**	**IQR**	**% ‡**	**Median**	**IQR**	**% ‡**	**Median**	**IQR**	**% ‡**	**Median**	**IQR**	**% ‡**	**Median**	**IQR**	**% ‡**	**Median**	**IQR**	**% ‡**	**Median**	**IQR**	**% ‡**	**Median**	**IQR**
Milk	Full-fat	20	3.6	0.2-43.8	23	2.2	0.1-23.9	20.5	2.9	0.2-33.7	27.5	2.1	0-20.5	25	1.5	0-8.4	21.5	1.8	0.1-10.5	22	1.9	0-15.8	24.8	1.5	0-8.7	23.8	1.8	0.1-19.5
	Low-fat	0.8	150.5	38.6-271.3	0.2	110.1	42.3-203.6	1.2	116.7	50.7-198.5	2	114.2	29.9-198	0	117.4	47.5-222.2	1.9	130	41.9-226.7	1.6	120.3	30.9-253.1	1.6	112	34.9-226.8	2.8	103.1	28.7-230.6
	Total	0.3	204.7	112.5-320.7	0	154.7	79.6-256.1	0.2	150.1	78.8-244.6	0.7	152.5	74.1-249.6	0	151.6	75-255.9	0.7	164.3	91.5-259.4	0.7	166.1	78.3-273.9	0.7	149.9	72.9-263.8	1.3	144.7	62.6-279.6
Yoghurt	Full-fat	72.5	0	0-1	70.1	0	0-2.1	71	0	0-3.8	68	0	0-4.2	69.1	0	0-4.3	70.3	0	0-5.4	70	0	0-3.6	69.8	0	0-1.6	68.7	0	0-7.7
	Low-fat	69.5	0	0-0.8	70.7	0	0-0.8	72.9	0	0-0.4	67.1	0	0-3.4	68.9	0	0-1.4	71.8	0	0-1	67.4	0	0-1.9	66.5	0	0-1.2	68.7	0	0-1.3
	Total	56.5	0	0-20.1	53.9	0	0-22.8	55.5	0	0-28.3	50.7	0	0-25.7	51.6	0	0-27.5	53.6	0	0-26.2	51.5	0	0-26.6	50.5	0	0-24.3	51.9	0	0-30.2
Cheese	High-fat	21.6	12	1.9-26.2	23.2	14.3	0.7-30	20.5	13.3	1.5-29.3	21.4	13.3	1.7-28.3	21.3	11.9	2-30	19.3	13.3	2.4-32.3	19.2	11	2-30	17.4	11.9	2.3-30	19.1	11.3	2.8-26.3
	Low-fat	86.4	0	0-0	88.9	0	0-0	91.4	0	0-0	88.3	0	0-0	89.3	0	0-0	89.2	0	0-0	88.6	0	0-0	90.5	0	0-0	86.3	0	0-0
	Total	20.1	13.9	2.4-29.7	22.1	15.4	2-33.1	19.3	14.3	2.4-30	20	14.3	2.1-30	20.3	12.9	2.3-32.2	18.1	15.4	3.3-35	17.8	13.4	2.6-31.9	16.5	14.3	2.7-32.5	18.4	12.3	2.8-28.7
Cream		18.5	0.9	0-5.9	21.9	0.8	0-4.4	24.5	0.8	0-5.7	27.5	0.8	0-4.3	26.6	1.1	0-5.3	27.9	1	0-5.6	27.8	1	0-5.8	27	0.7	0-5.2	26	1.2	0-5.6
Butter		17.4	2.9	0.6-8.4	18.4	2.3	0.5-6.6	22.6	2.9	0.2-8.5	19.1	2.7	0.4-8.2	19.7	2.8	0.4-7.9	17.9	4	0.7-9.8	18.5	3.3	0.9-8.1	15.5	3.2	0.5-9	16.9	3.4	0.6-8.8
Fermented dairy products §		13.5	23.8	6.3-50.8	14.1	30.1	8.2-57.1	13.1	28.2	7.1-59.7	13.1	28.2	7.9-61.2	11.7	27.5	6.9-57.1	11	29	9.3-61.4	11.7	29.1	7.1-61.1	10.4	26.7	6.4-54.5	11.2	26.8	9.5-61.9
Total dairy products	High-fat ‖	1.5	40.3	17.1-89.2	1.8	36.1	15.6-76.2	1.4	41.6	16-87.9	2.5	32.9	15.8-76.3	2.5	31.5	13.6-60	2.1	38.4	17.6-69.9	2.6	35.3	15.6-73.6	1.8	34.6	13.3-67.6	2.2	32.7	15.6-74.3
	Low-fat #	0.7	165.5	52-300.3	0.2	133.1	58.3-235.8	1.2	132.1	62.6-233	1.4	127.9	48.8-220.1	0	143.6	62.8-259.9	1.1	152.4	67.1-252.9	1.3	153.8	46-280.7	1.1	129.6	56-247.9	1.8	134.8	51.8-277.8
	Total *	0.1	257.5	152.9-380.1	0	203.2	115.6-332.4	0.2	202.5	128-309.7	0	193.3	114.1-302	0	198	107.7-311.6	0.2	216.2	135.1-317.9	0.2	232.1	114.6-339	0.2	201.7	107.5-315.5	0.2	198.1	113.9-339.6
≥65 years	**N**																											
		% ‡	**Median**	**IQR**	**% ‡**	**Median**	**IQR**	**% ‡**	**Median**	**IQR**	**% ‡**	**Median**	**IQR**	**% ‡**	**Median**	**IQR**	**% ‡**	**Median**	**IQR**	**% ‡**	**Median**	**IQR**	**% ‡**	**Median**	**IQR**	**% ‡**	**Median**	**IQR**
Milk	Full-fat	12.1	41.6	1.3-245.2	23.4	2.2	0.1-99.2	25	5.2	0-61.1	32.2	0.7	0-13.2	24.4	0.9	0-9.3	20.6	1.4	0-11.6	21.9	1.3	0-18.4	28.3	1.6	0-17.7	26.5	1.8	0-22.1
	Low-fat	1.1	97.7	18-287.7	1.6	143.2	40-248.6	0	167.4	70.2-306.6	2.3	163.6	96.8-245	0	186.4	104.1-350.9	1.3	182.6	97.5-264.5	1.7	151.8	68.9-273.6	0	179	75.9-285.5	0.7	172.6	94.2-265.7
	Total	0.1	289.2	179.9-413.4	0.8	213.1	135.8-336.4	0	232.6	155.4-361.7	0	185.4	112.2-291.1	0	208.3	127.6-371.4	1.3	204	137.6-284.6	0.6	188.2	92.1-297.8	0	221.5	117.1-304.6	0	203	128.9-280.8
Yoghurt	Full-fat	86.2	0	0-0	74.2	0	0-2.1	68.1	0	0-6	58.6	0	0-34.4	72.5	0	0-2.5	78.1	0	0-0	73.6	0	0-3.5	73.4	0	0-0	70.2	0	0-7.5
	Low-fat	86.5	0	0-0	76.6	0	0-0.1	69	0	0-0.3	74.7	0	0-0.4	67.9	0	0-19.9	68.1	0	0-23.5	68.5	0	0-2.4	70.7	0	0-5.1	65.6	0	0-21
	Total	80.5	0	0-0	60.5	0	0-27.5	52.6	0	0-37.3	49.4	0.4	0-54.4	53.4	0	0-34.3	60.6	0	0-34.9	55.1	0	0-38.9	56	0	0-31.7	53.6	0	0-47.6
Cheese	High-fat	30.4	6.4	0-15.1	31.5	6.4	0-20.5	28.4	12.5	1.4-22.7	23	9.1	2.1-20.5	27.5	8.3	0-22.5	27.5	9.3	0-21.7	27	9	0-20.9	22.3	10.8	0.7-23.2	24.5	9.4	1-20.9
	Low-fat	93	0	0-0	87.1	0	0-0	93.1	0	0-0	94.3	0	0-0	91.6	0	0-0	90.6	0	0-0	89.3	0	0-0	92.4	0	0-0	93.4	0	0-0
	Total	28.7	7	0-16.8	28.2	8.6	0-23.6	25.9	12.9	1.9-23.7	23	11	2.1-25.7	26.7	9.2	0-24	26.2	10	0-22	24.7	10	0-22.2	21.2	12.9	1.4-25.4	23.2	10.5	1.5-21.4
Cream		16.2	0.4	0-4.3	25	1.5	0-5.9	18.1	1.5	0.1-9.5	21.8	1.3	0-6	22.1	0.9	0-6.6	20.6	1.6	0-6.7	22.5	1.5	0-7.5	28.3	0.6	0-5.8	24.5	2	0-8
Butter		15.2	3.9	0.8-14.1	15.3	3	1.1-10.8	13.8	7.9	2.7-15.2	14.9	6.3	1.7-15.4	19.8	3.5	0.9-12.5	17.5	3.4	0.7-10.6	15.7	3	0.7-9	17.4	4.4	0.3-14.9	13.9	4.6	1.8-9.8
Fermented dairy products §		25.4	9.9	0.2-26.4	18.5	22.6	3.6-62.2	18.1	25.4	5.4-60.9	14.9	35	8.9-72.6	18.3	27.4	1.9-55	19.4	22.5	7-51.3	17.4	25	5.4-56.3	15.8	28.1	5.5-59.1	17.2	21.9	4-68.5
Total dairy products	High-fat ‖	0.5	74.3	25.3-263.9	3.2	42.9	19.3-147.2	1.7	45.9	21.4-106.9	2.3	31	15.4-69.6	1.5	30.8	16.5-56.1	1.3	27.5	15.8-65.4	1.1	30.9	14-63.6	1.1	37.8	14.7-71.2	1.3	36.7	17.4-68.3
	Low-fat #	1	113.8	23.7-305.8	0.8	161.5	55.8-265.2	0	211.7	77.5-326.9	2.3	201.2	115.1-286.2	0	209.3	111.4-382.5	1.3	200.6	102.9-302.2	1.7	181.1	72.8-290.7	0	201.2	83.3-307.7	0	203	109.4-294.4
	Total *	0	327	215.5-446.5	0.8	269.7	180.9-395.6	0	296.1	199-422.9	0	263.5	156.2-361.2	0	265.3	172.6-441.5	0.6	258.8	174.6-357.5	0	235.2	130.8-338.5	0	283.6	167.7-370	0	251.1	171.9-357.1

†Old survey years are 2000/2001 for adults <65 years and 1994/1995 for adults ≥65 years
‡Percentage of non-consumers
§Sum of yoghurt and cheese
‖Sum of full-fat milk, high-fat cheese, total cream and butter
#Sum of low-fat milk, total yoghurt, and low-fat cheese
*Sum of total milk, total yoghurt, total cheese, total cream and butter

Table A.3 Means and SE of total and types of dairy products (kcal/2000 kcal) within consumers and percentage of consumers among adults <65 years and ≥65 years in the old survey years and the first 8 years of the rolling programme of the National Diet and Nutrition Survey (NDNS) (2008-2016)

	Survey year	Old survey †			2008/2009			2009/2010			2010/2011			2011/2012			2012/2013			2013/2014			2014/2015			2015/2016		
<65 years	**N**	1,723			488			420			444			512			535			546			556			597		
Dairy products (kcal/2000 kcal)		**% ‡**	**Mean**	**SE**	**% ‡**	**Mean**	**SE**	**% ‡**	**Mean**	**SE**	**% ‡**	**Mean**	**SE**	**% ‡**	**Mean**	**SE**	**% ‡**	**Mean**	**SE**	**% ‡**	**Mean**	**SE**	**% ‡**	**Mean**	**SE**	**% ‡**	**Mean**	**SE**
Milk	Full-fat	80	69.3	3.7	77	52.4	6.3	79.5	49.7	5.3	72.5	52.9	5.9	75	38.6	5.5	78.5	39.9	5.8	78	51.4	7.4	75.2	43.9	7.6	76.2	49.5	5.7
	Low-fat	99.2	196.3	5	99.8	162.4	7.2	98.8	172.3	8.5	98	157.8	8.2	100	174.4	8.2	98.1	176.3	7.3	98.4	184.9	9.8	98.4	167.9	8.7	97.2	176.6	9.5
	Total	99.7	251	4.9	100	202.4	6.9	99.8	210	9.7	99.3	194.5	8.8	100	203.5	8.3	99.3	204.8	7	99.3	221.9	9.8	99.3	199.7	8.2	98.7	213.3	8.9
Yoghurt	Full-fat	27.5	24.3	2.2	29.9	32.1	3.4	29	31.1	3.4	32	28.6	3.3	30.9	31.4	3.3	29.7	32	4	30	34	5	30.2	26.4	2.6	31.3	31.6	3
	Low-fat	30.5	40.7	2.2	29.3	37.8	4.7	27.1	45.8	4.8	32.9	37.5	4.3	31.1	42.6	3.3	28.2	50.5	5.6	32.6	34.5	4.7	33.5	36.8	3.9	31.3	52.6	6.7
	Total	43.5	43.8	2.2	46.1	45.5	3.7	44.5	47.5	3.5	49.3	43.9	4	48.4	46.9	2.9	46.4	50.3	4.3	48.5	45.7	4.4	49.5	40.4	3	48.1	52.1	4.6
Cheese	High-fat	78.4	22.4	0.6	76.8	25.4	1.3	79.5	26.5	2	78.6	26.5	1.4	78.7	24.7	1	80.7	26.2	1.2	80.8	25.5	1.2	82.6	24.9	1.3	80.9	23.2	1.2
	Low-fat	13.6	22	1.6	11.1	20.5	4.4	8.6	21.4	3.8	11.7	19.1	6.2	10.7	14.1	2.9	10.8	19.8	2.8	11.4	23	6.4	9.5	20.8	5.3	13.7	14.9	1.8
	Total	79.9	25.7	0.7	77.9	28.3	1.5	80.7	28.3	2	80	29	1.9	79.7	26.2	1.1	81.9	28	1.3	82.2	28.6	1.6	83.5	27	1.7	81.6	25.7	1.3
Cream		81.5	6.4	0.3	78.1	5.9	0.6	75.5	7	0.7	72.5	7	0.7	73.4	7.3	0.6	72.1	6.9	0.7	72.2	7.1	0.7	73	6.5	0.7	74	6.1	0.4
Butter		82.6	7.3	0.2	81.6	6.5	0.5	77.4	8.4	0.6	80.9	7.6	0.5	80.3	7.4	0.5	82.1	8.4	0.5	81.5	7.5	0.5	84.5	7.7	0.5	83.1	8.7	0.6
Fermented dairy products §		86.5	45.8	1.4	85.9	49.3	2.4	86.9	50.8	2.8	86.9	51.2	2.9	88.3	49.3	2.1	89	52.2	3	88.3	51.6	2.8	89.6	46.7	2.7	88.8	52.1	2.9
Total dairy products	High-fat ‖	98.5	86.1	3.1	98.2	71.4	5.1	98.6	73.1	4.7	97.5	72.1	4.2	97.5	61.8	4.7	97.9	66.4	4.6	97.4	72.4	5.4	98.2	66	6	97.8	70.9	4.8
	Low-fat #	99.3	218.4	5.2	99.8	185.3	8	98.8	195.5	9.1	98.6	180.4	9.2	100	198.3	9	98.9	198.7	7.7	98.7	209.1	10.1	98.9	188.4	8.9	98.2	202.7	9.8
	Total *	99.9	301.9	5.1	100	254.9	7.7	99.8	266.1	10.3	100	248.2	10	100	258.4	9.2	99.8	263.2	7	99.8	276.1	10	99.8	251.6	8.4	99.8	269.7	9.1
≥65 years	**N**	1,733			124			116			87			131			160			178			184			151		
		% ‡	**Mean**	**SE**	**% ‡**	**Mean**	**SE**	**% ‡**	**Mean**	**SE**	**% ‡**	**Mean**	**SE**	**% ‡**	**Mean**	**SE**	**% ‡**	**Mean**	**SE**	**% ‡**	**Mean**	**SE**	**% ‡**	**Mean**	**SE**	**% ‡**	**Mean**	**SE**
Milk	Full-fat	87.9	203.4	7.4	76.6	120.1	19.4	75	93.7	16.3	67.8	58.7	15.3	75.6	49.9	13.3	79.4	50.3	10.2	78.1	48.1	9	71.7	72.8	15	73.5	43.5	11.5
	Low-fat	98.9	219	6.7	98.4	210	20.4	100	255	24.6	97.7	233.3	17.8	100	248.6	16.2	98.8	246.8	14.1	98.3	228.2	14.9	100	236	17.5	99.3	256.3	17.4
	Total	99.9	393.4	6.4	99.2	302.1	19.8	100	323.9	23	100	266.9	17.1	100	286.2	16.4	98.8	287.4	13	99.4	264.5	14.5	100	287	17.1	100	287.4	16
Yoghurt	Full-fat	13.8	27.9	2.2	25.8	39	7.7	31.9	39.8	8.8	41.4	61	7.4	27.5	33.2	7.4	21.9	34.7	6.2	26.4	41.1	6.8	26.6	33.7	4.1	29.8	60.4	14.8
	Low-fat	13.5	62.2	4.6	23.4	64	14.4	31	41.1	7.3	25.3	62	12.7	32.1	49.3	5.8	31.9	50.7	5.3	31.5	53.7	7.1	29.3	64.1	11.7	34.4	48	7.3
	Total	19.5	62.8	3.8	39.5	63.8	13.5	47.4	53.9	6.4	50.6	81.7	9.1	46.6	52.7	6.1	39.4	61.6	5.7	44.9	61.6	7.2	44	62.7	8.8	46.4	80.9	12.6
Cheese	High-fat	69.6	17.4	0.5	68.5	26	3.3	71.6	21.1	1.8	77	24.1	2.6	72.5	22.3	1.8	72.5	22.8	1.7	73	22.2	1.6	77.7	24.1	2	75.5	23.9	2.1
	Low-fat	7	21.9	3.1	12.9	27.1	6	6.9	19	6.6	5.7	31.7	6.9	8.4	5.2	1.7	9.4	8.1	1.7	10.7	22	7.3	7.6	17.8	6.4	6.6	13.6	3.3
	Total	71.3	19.5	0.6	71.8	29.2	3.5	74.1	22.2	1.9	77	26.4	2.7	73.3	22.8	1.8	73.8	23.4	1.7	75.3	24.7	2.3	78.8	26	2	76.8	24.4	2.2
Cream		83.8	5.3	0.3	75	9	1.9	81.9	9.3	1.2	78.2	7.2	1.1	77.9	7.1	1.1	79.4	9.4	1.3	77.5	11	2.6	71.7	7.8	1.2	75.5	9.2	1.5
Butter		84.8	12.9	0.4	84.7	9.8	1.1	86.2	13.2	1.3	85.1	12.8	1.4	80.2	10.7	1	82.5	10	0.9	84.3	9.4	1	82.6	13.2	1.4	86.1	10	1
Fermented dairy products §		74.6	36.6	1.7	81.5	58.9	7.5	81.9	50.6	6	85.1	73	7	81.7	51.4	4.8	80.6	51.6	4.2	82.6	56.7	4.1	84.2	57.8	5.3	82.8	65.8	7.7
Total dairy products	High-fat ‖	99.5	205.1	6.8	96.8	129.2	16.4	98.3	105.8	12.4	97.7	76.3	11.5	98.5	69.9	10.6	98.8	73.1	9.1	98.9	70.2	8.1	98.9	87.6	11.4	98.7	67.9	8.9
	Low-fat #	99	234.5	6.9	99.2	239.5	20.1	100	281.5	24.9	97.7	277.8	18.6	100	274.5	17.2	98.8	272.8	14.2	98.3	259.5	15.3	100	265.6	18.1	100	292.4	18.1
	Total *	100	436	6.4	99.2	365.9	17.9	100	385.6	22.7	100	345.6	18	100	342.8	18.3	99.4	342.5	13.1	100	326.2	14.1	100	351.6	17.5	100	360.1	15.9

†Old survey years are 2000/2001 for adults <65 years and 1994/1995 for adults ≥65 years
‡Percentage of consumers
§Sum of yoghurt and cheese
‖Sum of full-fat milk, high-fat cheese, total cream and butter
#Sum of low-fat milk, total yoghurt, and low-fat cheese
*Sum of total milk, total yoghurt, total cheese, total cream and butter

A.2 Chapter 4

Table A.4 Associations of total and types of dairy consumption with clinical markers of body weight and composition from multiple linear regression models †

	BMI (kg/m^2)			**Waist (cm)**			**Waist / Hip circumference**		
Mean (SD)	26.9 (4.8)			91.0 (13.5)			0.9 (0.1)		
Participants (N)	12,064			12,058			12,043		
Dairy consumption (servings †/d)	b	95% CI		b	95% CI		b	95% CI	
Milk									
Demographic, energy ‡	0.1	0.01	0.2	0.09	-0.19	0.37	0	0	0
+ SES, lifestyle, diet	0.264	0.16	0.36	0.48	0.19	0.77	0	0	0
+ BMI				-0.153	-0.29	0	0	0	0
Yoghurt									
Demographic, energy ‡	0.06	-0.1	0.22	-0.21	-0.7	0.27	-0.01	-0.01	0
+ SES, lifestyle, diet	0.14	-0.03	0.31	0.39	-0.1	0.88	0	0	0
+ BMI				-0.22	-0.46	0.02	0	0	0
Cheese									
Demographic, energy ‡	-0.19	-0.42	0.04	-0.34	-0.93	0.26	0	-0.01	0
+ SES, lifestyle, diet	0.06	-0.15	0.27	0.15	-0.41	0.71	0	0	0.01
+ BMI				0.14	-0.14	0.43	0	0	0
Fermented dairy products									
Demographic, energy ‡	-0.04	-0.16	0.09	-0.26	-0.62	0.09	0	-0.01	0
+ SES, lifestyle, diet	0.11	-0.03	0.24	0.29	-0.08	0.66	0	0	0
+ BMI				-0.07	-0.26	0.12	0	0	0
Full-fat milk									
Demographic, energy ‡	-0.41	-0.58	-0.24	-1.15	-1.61	-0.69	0	-0.01	0
+ SES, lifestyle, diet	-0.13	-0.31	0.05	-0.61	-1.08	-0.13	0	-0.01	0
+ BMI				-0.303	-0.56	-0.04	0.003	-0.01	0
Low-fat milk									
Demographic, energy ‡	0.16	0.06	0.26	0.25	-0.04	0.54	0	0	0
+ SES, lifestyle, diet	0.284	0.18	0.38	0.55	0.25	0.84	0	0	0
+ BMI				-0.14	-0.28	0.01	0	0	0
Full-fat yoghurt									
Demographic, energy ‡	-1.38	-1.78	-0.97	-3.6	-4.81	-2.4	-0.02	-0.03	-0.01
+ SES, lifestyle, diet	-0.453	-0.85	-0.06	-0.85	-1.99	0.3	0	-0.01	0
+ BMI				-0.01	-0.53	0.52	0	0	0.01
Low-fat yoghurt									
Demographic, energy ‡	0.25	0.08	0.43	0.24	-0.28	0.75	0	-0.01	0
+ SES, lifestyle, diet	0.223	0.04	0.41	0.56	0.03	1.08	0	0	0
+ BMI				-0.25	-0.51	0.01	0	0	0
High-fat cheese									
Demographic, energy ‡	-0.63	-0.93	-0.33	-0.83	-1.59	-0.07	0	-0.01	0
+ SES, lifestyle, diet	-0.11	-0.39	0.17	0.14	-0.56	0.84	0	0	0.01
+ BMI				0.393	0.03	0.75	0	0	0
Low-fat cheese									
Demographic, energy ‡	0.44	0.1	0.79	0.38	-0.51	1.27	0	-0.01	0.01
+ SES, lifestyle, diet	0.28	-0.03	0.59	0.18	-0.69	1.05	0	0	0.01
+ BMI				-0.25	-0.75	0.25	0	-0.01	0
Butter									
Demographic, energy ‡	-0.06	-0.17	0.05	-0.17	-0.49	0.14	0	0	0
+ SES, lifestyle, diet	0.153	0.03	0.27	0.22	-0.12	0.57	0	0	0
+ BMI				-0.1	-0.26	0.06	0	0	0
Ice-cream									
Demographic, energy ‡	0.83	0.38	1.28	2.45	1.13	3.76	0.01	0	0.02
+ SES, lifestyle, diet	0.733	0.31	1.15	2.2	1.06	3.34	0.01	0	0.02
+ BMI				-0.01	-0.6	0.57	0	0	0.01

Table A.4 (continued)

	BMI (kg/m^2)			**Waist (cm)**			**Waist / Hip circumference**		
Mean (SD)	26.9 (4.8)			91.0 (13.5)			0.9 (0.1)		
Participants (N)	12,064			12,058			12,043		
Dairy consumption (servings †/d)	b	95% CI		b	95% CI		b	95% CI	
Low-fat fermented dairy products									
Demographic, energy ‡	0.12	-0.02	0.26	-0.1	-0.5	0.31	0	-0.01	0
+ SES, lifestyle, diet	0.163	0.01	0.31	0.33	-0.09	0.75	0	0	0
+ BMI				-0.2	-0.41	0.01	0	0	0
High-fat dairy products									
Demographic, energy ‡	-0.18	-0.26	-0.1	-0.46	-0.69	-0.23	0	0	0
+ SES, lifestyle, diet	0.02	-0.08	0.12	-0.09	-0.36	0.19	0	0	0
+ BMI				-0.08	-0.21	0.05	0	0	0
Low-fat dairy products									
Demographic, energy ‡	0.13	0.05	0.21	0.09	-0.14	0.33	0	0	0
+ SES, lifestyle, diet	0.224	0.13	0.31	0.44	0.19	0.69	0	0	0
+ BMI				-0.153	-0.27	-0.03	0	0	0
Total dairy products									
Demographic, energy ‡	0.02	-0.04	0.08	-0.07	-0.26	0.12	0	0	0
+ SES, lifestyle, diet	0.214	0.13	0.28	0.39	0.16	0.61	0	0	0
+ BMI				-0.123	-0.23	-0.01	0	0	0

†Associations are per serving/day (Milk: 1 average glass (200g); Yoghurt: 125g carton; Cheese: medium serving (40g); Single cream: 1 tablespoon (15g); Double cream: 1 tablespoon (30g); Butter: 1 teaspoon (10g); Ice-cream: 1 average scoop/tub (60g) as defined by Food Standards Agency 2002)[197]

‡Model 1: age (years), sex, test-site (Cambridge, Ely, Wisbech), ethnicity (white, non-white), total energy intake (kcal/d), mutual adjustment for dairy products; Model 2: Model 1 + educational level (low, medium, high), age when full-time education finished (years), socio-economic status based on occupation (low: technical/semi-routine and routine occupations; medium: lower managerial / intermediate occupations; high: professional/higher managerial occupations), income (<£20,000, £20,000-40,000, >£40,000), marital status (single, married, widowed/separated), smoking status (never, former, current smoker), pack-years of smoking, energy expenditure due to physical activity (kj/kg/d), lipid-lowering medication (Yes, No), anti-hypertensive medication (Yes, No), hormone-replacement therapy (Yes, No, Men). Model 3: Model 2 + intakes (g/d) of fruit, vegetables (not including potatoes), potatoes, legumes, processed cereals, whole-grain cereals, poultry and eggs, red meat, processed meat, fish, sauces, margarine, nuts, sweet snacks, sugar-sweetened beverages, artificially sweetened beverages, fruit juice, regular coffee, decaffeinated coffee, tea, alcoholic beverages, plasma vitamin C levels (μmol/l), dietary supplement use (Yes, No). Model 4: Model 3 + BMI (kg/m^2)

Table A.5 Associations of total and types of dairy consumption with markers of body composition assessed with Dual Energy X-Ray Absorptiometry (DEXA) from multiple linear regression models †

	Body fat mass (kg)			**Peripheral fat mass (kg)**			**VAT/SCAT**			**VAT (kg)**			**SCAT (kg)**			**Body lean mass (kg)**			**Appendicular lean mass (kg)**		
Mean (SD)	26.5 (9.4)			11.6 (4.1)			0.8 (0.7)			1.0 (0.8)			1.4 (0.7)			48.7 (10.0)			22.2 (5.5)		
Participants (N)	11,523			11,523			11,253			11253			11,253			11,523			11,523		
Dairy consumption (servings †/d)	b	95% CI		b	95% CI		% change ‡	95% CI		% change ‡	95% CI		% change ‡	95% CI		b	95% CI		b	95% CI	
Milk																					
Demographic, energy §	0.08	-0.13	0.29	0.09	0.01	0.17	-1.9	-3.32	-0.46	-0.01	-0.02	0.01	0.84	-0.26	1.95	0.44	0.28	0.59	0.2	0.11	0.29
+ SES, lifestyle, diet	0.37	0.15	0.59	0.16	0.08	0.25	-1.15	-2.75	0.48	0	-0.01	0.02	2.28	1.09	3.49	0.56	0.38	0.73	0.26	0.17	0.36
+ BMI	-0.08	-0.18	0.02	0	-0.06	0.05	-2.20*	-3.72	-0.66	-0.02**	-0.03	-0.01	0.58	-0.16	1.32	0.33**	0.19	0.46	0.12*	0.04	0.19
Yoghurt																					
Demographic, energy §	-0.19	-0.55	0.17	0	-0.13	0.13	-7.1	-9.68	-4.45	-0.04	-0.07	-0.01	1.4	-0.71	3.56	0.45	0.17	0.73	0.21	0.07	0.35
+ SES, lifestyle, diet	0.17	-0.19	0.54	0.1	-0.03	0.24	-2.36	-5.13	0.5	0	-0.03	0.02	2.89	0.6	5.24	0.26	-0.05	0.58	0.12	-0.04	0.28
+ BMI	0	-0.17	0.17	0	-0.09	0.1	-4.01*	-6.63	-1.32	-0.02*	-0.04	0	1.41*	0.16	2.69	0.07	-0.17	0.32	0.06	-0.07	0.19
Cheese																					
Demographic, energy §	-0.13	-0.6	0.34	0	-0.18	0.18	-3.75	-6.7	-0.7	-0.03	-0.06	0	0.42	-1.99	2.9	0.26	-0.1	0.61	-0.01	-0.19	0.18
+ SES, lifestyle, diet	0.18	-0.25	0.61	0.09	-0.09	0.26	-1.21	-4.35	2.04	0	-0.03	0.03	2.19	-0.25	4.69	0.17	-0.18	0.53	-0.04	-0.23	0.15
+ BMI	0.06	-0.12	0.23	0.03	-0.08	0.13	-1.51	-4.52	1.58	0	-0.02	0.02	0.79	-0.66	2.27	0.08	-0.2	0.36	-0.07	-0.22	0.08
Fermented dairy products																					
Demographic, energy §	-0.16	-0.44	0.11	0	-0.1	0.1	-5.66	-7.55	-3.72	-0.04	-0.05	-0.02	0.99	-0.55	2.55	0.37	0.17	0.58	0.12	0.01	0.23
+ SES, lifestyle, diet	0.18	-0.1	0.45	0.1	-0.01	0.2	-1.87	-4.02	0.33	0	-0.02	0.02	2.59	0.93	4.28	0.23	0	0.45	0.05	-0.07	0.18
+ BMI	0.03	-0.1	0.15	0.01	-0.06	0.08	-2.96*	-5	-0.88	-0.01*	-0.03	0	1.16*	0.22	2.11	0.08	-0.1	0.26	0	-0.09	0.1
Full-fat milk																					
Demographic, energy §	-0.94	-1.29	-0.59	-0.27	-0.41	-0.14	-2.83	-5.48	-0.11	-0.07	-0.1	-0.04	-4.6	-6.42	-2.74	-0.13	-0.45	0.19	-0.1	-0.27	0.07
+ SES, lifestyle, diet	-0.48	-0.83	-0.13	-0.13	-0.27	0.01	-1.54	-4.42	1.43	-0.04	-0.07	-0.02	-2.25	-4.21	-0.26	0.09	-0.25	0.43	0.01	-0.17	0.2
+ BMI	-0.23*	-0.41	-0.06	-0.06	-0.15	0.03	-0.4	-3.19	2.47	-0.03*	-0.05	-0.01	-1.33*	-2.59	-0.04	0.15	-0.1	0.4	0.04	-0.09	0.16
Low-fat milk																					
Demographic, energy §	0.21	-0.01	0.43	0.14	0.05	0.22	-1.79	-3.24	-0.33	0	-0.02	0.01	1.49	0.37	2.63	0.49	0.33	0.65	0.23	0.14	0.32
+ SES, lifestyle, diet	0.43	0.21	0.65	0.18	0.1	0.27	-1.13	-2.74	0.51	0.01	-0.01	0.02	2.58	1.38	3.8	0.58	0.4	0.75	0.27	0.18	0.37
+ BMI	-0.07	-0.17	0.03	0	-0.05	0.05	-2.31*	-3.84	-0.76	-0.02*	-0.03	-0.01	0.71	-0.03	1.46	0.34**	0.2	0.47	0.12*	0.05	0.2
Full-fat yoghurt																					
Demographic, energy §	-2.72	-3.6	-1.85	-1.01	-1.34	-0.68	-11.8	-19.04	-3.92	-0.18	-0.24	-0.11	-11.25	-15.77	-6.49	-0.5	-1.15	0.16	-0.31	-0.67	0.05
+ SES, lifestyle, diet	-0.78	-1.59	0.03	-0.41	-0.73	-0.08	-0.42	-8.74	8.65	-0.02	-0.09	0.04	-3.41	-8.09	1.5	-0.27	-0.94	0.39	-0.17	-0.54	0.19
+ BMI	-0.14	-0.57	0.28	-0.06	-0.3	0.18	1.2	-6.34	9.35	0	-0.04	0.05	1.46	-1.4	4.41	0.39	-0.18	0.96	0.22	-0.09	0.52
Low-fat yoghurt																					
Demographic, energy §	0.18	-0.22	0.57	0.15	0	0.3	-6.5	-9.22	-3.7	-0.02	-0.05	0	3.51	1.07	6.01	0.6	0.3	0.89	0.28	0.14	0.43
+ SES, lifestyle, diet	0.32	-0.08	0.71	0.17	0.02	0.32	-2.57	-5.42	0.38	0	-0.03	0.03	3.93	1.36	6.57	0.35	0.01	0.68	0.16	0	0.33
+ BMI	0.02	-0.16	0.2	0.01	-0.09	0.11	-4.59*	-7.28	-1.82	-0.03*	-0.04	-0.01	1.41*	0.08	2.76	0.03	-0.24	0.29	0.04	-0.1	0.18
High-fat cheese																					
Demographic, energy §	-0.75	-1.37	-0.13	-0.26	-0.49	-0.02	-4.11	-7.78	-0.3	-0.05	-0.1	-0.01	-2.6	-5.6	0.5	0.08	-0.39	0.55	-0.1	-0.35	0.14
+ SES, lifestyle, diet	-0.05	-0.62	0.52	-0.08	-0.31	0.15	-1.81	-5.78	2.32	0.01	-0.03	0.05	0.97	-2.31	4.37	0.11	-0.36	0.58	-0.08	-0.34	0.17
+ BMI	0.24*	0.02	0.47	0.03	-0.1	0.17	-1.11	-4.87	2.8	0.02	-0.01	0.05	1.85	-0.11	3.85	0.2	-0.18	0.58	0	-0.2	0.21

Table A.5 (continued)

	Body fat mass (kg)			**Peripheral fat mass (kg)**			**VAT/SCAT**			**VAT (kg)**			**SCAT (kg)**			**Body lean mass (kg)**			**Appendicular lean mass (kg)**		
Mean (SD)	26.5 (9.4)			11.6 (4.1)			0.8 (0.7)			1.0 (0.8)			1.4 (0.7)			48.7 (10.0)			22.2 (5.5)		
Participants (N)	11,523			11,523			11,253			11253			11,253			11,523			11,523		
Dairy consumption (servings †/d)	b	95% CI		b	95% CI		% change ‡	95% CI		% change ‡	95% CI		% change ‡	95% CI		b	95% CI		b	95% CI	
Low-fat cheese																					
Demographic, energy §	0.76	0.04	1.48	0.43	0.12	0.74	-3.14	-8.13	2.13	0	-0.04	0.03	5.2	1.35	9.2	0.5	-0.01	1.02	0.13	-0.15	0.4
+ SES, lifestyle, diet	0.47	-0.16	1.11	0.31	0.04	0.59	-0.32	-5.15	4.74	-0.01	-0.05	0.02	3.77	0.07	7.6	0.25	-0.26	0.77	0.01	-0.26	0.28
+ BMI	-0.18	-0.42	0.07	0.02	-0.13	0.18	-2.11	-6.83	2.85	-0.03*	-0.06	0	-0.56	-2.72	1.66	-0.08	-0.49	0.34	-0.16	-0.38	0.05
Butter																					
Demographic, energy §	0	-0.23	0.23	0.04	-0.05	0.12	0.75	-0.89	2.43	0	-0.02	0.02	-0.36	-1.63	0.91	0.04	-0.15	0.22	0.03	-0.07	0.13
+ SES, lifestyle, diet	0.3	0.05	0.55	0.13	0.03	0.22	1.7	-0.15	3.58	0.02	0	0.04	0.7	-0.69	2.1	0.03	-0.18	0.24	0.02	-0.1	0.13
+ BMI	-0.05	-0.16	0.07	0	-0.06	0.06	0.45	-1.35	2.29	0.01	0	0.02	-0.63	-1.46	0.2	-0.1	-0.25	0.05	-0.04	-0.12	0.05
Ice-cream																					
Demographic, energy §	1.78	0.78	2.78	0.59	0.23	0.96	-0.53	-5.97	5.23	0.11	0.04	0.17	10.09	4.07	16.46	0.28	-0.34	0.9	0.22	-0.11	0.54
+ SES, lifestyle, diet	1.69	0.72	2.65	0.64	0.28	1	-1.4	-6.8	4.31	0.07	0.01	0.13	8.71	3.38	14.31	0.42	-0.24	1.07	0.18	-0.16	0.51
+ BMI	0.23	-0.16	0.62	0.07	-0.14	0.27	-4.43	-9.95	1.43	0	-0.05	0.05	2.51	-0.22	5.32	-0.44	-0.94	0.07	-0.18	-0.46	0.09
Low-fat fermented dairy products																					
Demographic, energy §	0.01	-0.31	0.32	0.08	-0.04	0.2	-6.16	-8.35	-3.91	-0.03	-0.05	-0.01	2.13	0.33	3.97	0.45	0.22	0.68	0.18	0.06	0.3
+ SES, lifestyle, diet	0.24	-0.07	0.55	0.14	0.02	0.26	-1.9	-4.26	0.53	-0.01	-0.03	0.01	3.04	1.14	4.98	0.26	0	0.51	0.09	-0.04	0.22
+ BMI	-0.03	-0.17	0.1	0.01	-0.07	0.09	-3.49*	-5.73	-1.2	-0.02*	-0.04	-0.01	0.99	-0.05	2.04	0.04	-0.16	0.24	0.01	-0.1	0.11
High-fat dairy products																					
Demographic, energy §	-0.29	-0.46	-0.12	-0.06	-0.13	0.01	-0.6	-1.85	0.66	-0.02	-0.03	-0.01	-1.7	-2.64	-0.75	0.04	-0.1	0.18	0	-0.08	0.08
+ SES, lifestyle, diet	0	-0.21	0.2	0.02	-0.06	0.1	1.07	-0.5	2.66	0	-0.01	0.02	-0.68	-1.83	0.49	0.05	-0.13	0.22	0.01	-0.09	0.1
+ BMI	-0.06	-0.16	0.04	-0.01	-0.06	0.04	0.76	-0.76	2.31	0	-0.01	0.02	-0.90*	-1.61	-0.18	0.03	-0.1	0.15	0.01	-0.06	0.07
Low-fat dairy products																					
Demographic, energy §	0.11	-0.07	0.29	0.1	0.04	0.17	-3.17	-4.37	-1.96	-0.01	-0.03	0	1.49	0.54	2.44	0.47	0.33	0.6	0.21	0.13	0.28
+ SES, lifestyle, diet	0.33	0.14	0.52	0.15	0.08	0.23	-1.4	-2.81	0.03	0	-0.01	0.02	2.51	1.46	3.58	0.46	0.31	0.61	0.2	0.12	0.29
+ BMI	-0.05	-0.14	0.03	0	-0.05	0.05	-2.58**	-3.91	-1.23	-0.02**	-0.03	-0.01	0.79*	0.16	1.43	0.25**	0.13	0.36	0.09*	0.02	0.15
Total dairy products																					
Demographic, energy §	0	-0.14	0.14	0.05	0	0.11	-1.76	-2.73	-0.78	-0.01	-0.02	0	0.37	-0.39	1.14	0.29	0.18	0.4	0.13	0.07	0.18
+ SES, lifestyle, diet	0.34	0.17	0.51	0.15	0.09	0.22	0.02	-1.27	1.33	0.01	0	0.02	1.8	0.81	2.8	0.33	0.19	0.47	0.15	0.07	0.22
+ BMI	-0.05	-0.13	0.03	0	-0.04	0.04	-1.24	-2.48	0.01	-0.01	-0.02	0	0.13	-0.47	0.74	0.12*	0.02	0.23	0.04	-0.01	0.1

†Associations are per serving/day (Milk: 1 average glass (200g); Yoghurt: 125g carton; Cheese: medium serving (40g); Single cream: 1 tablespoon (15g); Double cream: 1 tablespoon (30g); Butter: 1 teaspoon (10g); Ice-cream: 1 average scoop/tub (60g) as defined by Food Standards Agency 2002)

‡% change used for log-transformed outcomes after back-transformation (exponentiation) and defined as [(eb-1)*100]%

SModel 1: age (years), sex, test-site (Cambridge, Ely, Wisbech), ethnicity (white, non-white), total energy intake (kcal/d), mutual adjustment for dairy products; Model 2: Model 1 + educational level (low, medium, high), age when full-time education finished (years), socio-economic status based on occupation (low: technical/semi-routine and routine occupations; medium: lower managerial / intermediate occupations; high: professional/higher managerial occupations), income (<£20,000, £20,000-40,000, >£40,000), marital status (single, married, widowed/separated), smoking status (never, former, current smoker), pack-years of smoking, energy expenditure due to physical activity (kj/kg/d), lipid-lowering medication (Yes, No), anti-hypertensive medication (Yes, No), hormone-replacement therapy (Yes, No, Men).; Model 3: Model 2 + intakes (g/d) of fruit, vegetables (not including potatoes), potatoes, legumes, processed cereals, whole-grain cereals, poultry and eggs, red meat, processed meat, fish, sauces, margarine, nuts, sweet snacks, sugar-sweetened beverages, artificially sweetened beverages, fruit juice, regular coffee, decaffeinated coffee, tea, alcoholic beverages, plasma vitamin C levels (μmol/l), dietary supplement use (Yes, No). ;Model 4: Model 3 + BMI (kg/m^2)

*p<0.05 for the maximally adjusted model

**p<0.00025 (critical p after correction for multiple testing) for the maximally adjusted model

Abbreviations: SCAT: Subcutaneous Adipose Tissue; VAT: Visceral Adipose Tissue

Table A.6 Associations of total and types of dairy consumption with lipid markers from multiple linear regression models †

	Total / HDL-C			**Total cholesterol (mmol/l)**			**HDL-C (mmol/l)**			**LDL-C (mmol/l)**			**Triglycerides (mmol/l)**			**ApoA1 (g/l)**			**ApoB (g/l)**			**NEFA (umol/l)**		
Mean (SD)	3.7 (1.2)			5.4 (1.0)			1.5 (0.4)			3.4 (0.9)			1.2 (0.9)			1.4 (0.2)			1.0 (0.2)			341.3 (163.5)		
Participants (N)	11,988			11,989			11,988			11,897			11,988			9,647			9,650			9,638		
Dairy consumption (servings †/d)	b	95% CI		b	95% CI		b	95% CI		b	95% CI		% change ‡	95% CI		b	95% CI		b	95% CI		% change ‡	95% CI	
Milk																								
Demographic §	0.03	0.01	0.06	-0.02	-0.04	0.01	-0.02	-0.03	-0.01	-0.01	-0.03	0.01	0.84	-0.51	2.21	-0.01	-0.02	-0.01	0	-0.01	0	-1.49	-2.81	-0.15
+ SES, lifestyle	0.05	0.02	0.08	-0.02	-0.05	0	-0.03	-0.04	-0.02	-0.01	-0.04	0.01	2.31	0.82	3.81	-0.01	-0.02	-0.01	0	-0.01	0	-0.61	-2.12	0.93
+ BMI	0.035	0	0.06	-0.035	-0.06	0	-0.024	-0.03	-0.01	-0.02	-0.04	0.01	1.615	0.2	3.05	-0.015	-0.02	0	-0.01	-0.01	0	-0.83	-2.33	0.7
Yoghurt																								
Demographic §	-0.02	-0.07	0.02	-0.07	-0.11	-0.02	0	-0.01	0.02	-0.05	-0.09	-0.02	-3.63	-5.74	-1.48	-0.01	-0.02	0	-0.02	-0.03	-0.01	-0.17	-2.81	2.55
+ SES, lifestyle	0.03	-0.01	0.08	-0.02	-0.07	0.03	-0.02	-0.03	0	-0.01	-0.05	0.03	0.25	-1.91	2.46	-0.02	-0.03	-0.01	-0.01	-0.02	0	-0.65	-3.4	2.17
+ BMI	0.02	-0.03	0.06	-0.02	-0.07	0.03	-0.015	-0.03	0	-0.01	-0.05	0.03	-0.67	-2.75	1.44	-0.015	-0.03	0	-0.015	-0.02	0	-0.84	-3.58	1.99
Cheese																								
Demographic §	-0.04	-0.09	0.01	0	-0.05	0.05	0.02	0	0.04	-0.02	-0.06	0.03	-0.38	-3.19	2.51	0	-0.01	0.01	0	-0.02	0.01	2.86	-0.02	5.82
+ SES, lifestyle	0	-0.05	0.05	0	-0.05	0.06	0.01	-0.01	0.03	0	-0.05	0.05	2.22	-0.7	5.23	0	-0.02	0.01	0	-0.01	0.02	2.59	-0.52	5.8
+ BMI	0.02	-0.02	0.07	0	-0.05	0.06	0.01	-0.01	0.02	0	-0.05	0.05	2.13	-0.66	5	0	-0.02	0.01	0	-0.01	0.01	2.67	-0.46	5.9
Fermented dairy products																								
Demographic §	-0.03	-0.06	0	-0.04	-0.07	-0.01	0.01	0	0.02	-0.04	-0.07	-0.01	-2.34	-3.96	-0.7	-0.01	-0.01	0	-0.01	-0.02	0	1.15	-0.73	3.06
+ SES, lifestyle	0.02	-0.01	0.05	-0.01	-0.04	0.03	-0.01	-0.02	0	-0.01	-0.04	0.02	1.02	-0.73	2.8	-0.01	-0.02	0	-0.01	-0.01	0	0.72	-1.35	2.84
+ BMI	0.02	-0.01	0.05	-0.01	-0.05	0.02	-0.01	-0.02	0	-0.01	-0.04	0.02	0.41	-1.26	2.11	-0.015	-0.02	0	-0.01	-0.01	0	0.64	-1.43	2.76
Full-fat milk																								
Demographic §	0.02	-0.03	0.07	0.02	-0.03	0.06	-0.02	-0.03	0	0.02	-0.03	0.06	-1.28	-3.66	1.16	0	-0.01	0.01	0	-0.01	0.01	-1.72	-3.99	0.6
+ SES, lifestyle	0.06	0.01	0.1	-0.01	-0.06	0.04	-0.02	-0.04	-0.01	0	-0.05	0.04	1.13	-1.57	3.89	0	-0.01	0.01	0	-0.01	0.01	-1.1	-3.64	1.51
+ BMI	0.065	0.01	0.11	-0.01	-0.06	0.04	-0.035	-0.04	-0.01	0	-0.04	0.05	1.95	-0.58	4.55	0	-0.01	0.01	0	-0.01	0.01	-0.79	-3.32	1.8
Low-fat milk																								
Demographic §	0.03	0.01	0.06	-0.02	-0.05	0	-0.02	-0.03	-0.01	-0.01	-0.03	0.01	1.09	-0.29	2.48	-0.01	-0.02	-0.01	0	-0.01	0	-1.46	-2.81	-0.09
+ SES, lifestyle	0.05	0.02	0.08	-0.02	-0.05	0	-0.03	-0.04	-0.02	-0.01	-0.04	0.01	2.37	0.88	3.88	-0.01	-0.02	-0.01	0	-0.01	0	-0.58	-2.1	0.97
+ BMI	0.035	0	0.05	-0.035	-0.06	0	-0.024	-0.03	-0.01	-0.02	-0.04	0.01	1.605	0.18	3.04	-0.015	-0.02	-0.01	-0.01	-0.01	0	-0.83	-2.34	0.71
Full-fat yoghurt																								
Demographic §	-0.22	-0.32	-0.12	0.02	-0.1	0.14	0.11	0.07	0.15	-0.02	-0.13	0.08	-17.1	-21.9	-11.9	0.03	0	0.06	-0.01	-0.05	0.02	0.47	-6.75	8.26
+ SES, lifestyle	-0.02	-0.12	0.08	0.07	-0.04	0.19	0.03	-0.01	0.07	0.05	-0.06	0.15	-6.5	-12	-0.7	0	-0.02	0.03	0	-0.03	0.03	3.03	-4.95	11.67
+ BMI	0.02	-0.09	0.12	0.08	-0.04	0.2	0.01	-0.03	0.05	0.05	-0.06	0.16	-4.68	-10.37	1.37	-0.01	-0.04	0.02	0	-0.03	0.04	3.41	-4.64	12.12
Low-fat yoghurt																								
Demographic §	0	-0.04	0.05	-0.08	-0.13	-0.03	-0.01	-0.03	0	-0.06	-0.1	-0.02	-1.9	-4.05	0.3	-0.02	-0.03	-0.01	-0.02	-0.03	-0.01	-0.25	-3.02	2.61
+ SES, lifestyle	0.04	-0.01	0.09	-0.03	-0.08	0.02	-0.02	-0.04	-0.01	-0.02	-0.06	0.02	1.06	-1.18	3.35	-0.02	-0.03	-0.01	-0.01	-0.02	0	-1.02	-3.82	1.86
+ BMI	0.02	-0.03	0.06	-0.03	-0.08	0.02	-0.025	-0.03	0	-0.02	-0.06	0.02	-0.22	-2.35	1.96	-0.025	-0.03	0	-0.015	-0.02	0	-1.25	-4.05	1.62
High-fat cheese																								
Demographic §	-0.1	-0.17	-0.04	0.06	-0.01	0.13	0.05	0.02	0.07	0	-0.06	0.06	-1.92	-5.72	2.04	0.02	0	0.03	0	-0.02	0.01	2.3	-1.27	5.99
+ SES, lifestyle	-0.01	-0.08	0.05	0.06	-0.01	0.13	0.03	0	0.05	0.02	-0.04	0.09	1.67	-2.41	5.91	0	-0.01	0.02	0	-0.02	0.02	2.3	-1.73	6.5
+ BMI	0.02	-0.04	0.08	0.06	-0.01	0.13	0.025	0	0.05	0.03	-0.03	0.1	2.71	-1.13	6.7	0	-0.01	0.02	0	-0.02	0.02	2.53	-1.54	6.77

Table A.6 (continued)

	Total / HDL-C			Total cholesterol (mmol/l)			HDL-C (mmol/l)			LDL-C (mmol/l)			Triglycerides (mmol/l)			ApoA1 (g/l)			ApoB (g/l)			NEFA (umol/l)		
Mean (SD)	3.7 (1.2)			5.4 (1.0)			1.5 (0.4)			3.4 (0.9)			1.2 (0.9)			1.4 (0.2)			1.0 (0.2)			341.3 (163.5)		
Participants (N)	11,988			11,989			11,988			11,897			11,988			9,647			9,650			9,638		
Dairy consumption (servings †/d)	b	95% CI		b	95% CI		b	95% CI		b	95% CI		% change ‡	95% CI		b	95% CI		b	95% CI		% change ‡	95% CI	
Low-fat cheese																								
Demographic §	0.05	-0.02	0.12	-0.09	-0.16	-0.01	-0.03	-0.06	0	-0.04	-0.11	0.02	1.6	-2.38	5.74	-0.02	-0.04	0	0	-0.02	0.02	3.73	-0.9	8.57
+ SES, lifestyle	0.02	-0.05	0.09	-0.07	-0.14	0.01	-0.02	-0.06	0.01	-0.03	-0.1	0.04	2.84	-1.08	6.91	-0.02	-0.04	0.01	0.01	-0.01	0.02	2.97	-1.61	7.77
+ BMI	0.02	-0.04	0.08	-0.07	-0.15	0.01	-0.02	-0.05	0.01	-0.03	-0.1	0.03	1.47	-2.38	5.47	-0.01	-0.03	0.01	0	-0.02	0.02	2.85	-1.79	7.7
Butter																								
Demographic §	0.01	-0.02	0.04	0.07	0.04	0.09	0.01	0	0.02	0.07	0.04	0.09	-1.01	-2.49	0.49	0	0	0.01	0.01	0.01	0.02	-1.23	-2.83	0.39
+ SES, lifestyle	0.03	0	0.06	0.06	0.03	0.1	0.01	-0.01	0.02	0.07	0.04	0.1	0.5	-1.16	2.19	0	-0.01	0	0.01	0	0.02	-0.55	-2.32	1.26
+ BMI	0.02	-0.01	0.05	0.065	0.02	0.09	0.01	0	0.02	0.064	0.03	0.09	-0.25	-1.82	1.34	0	-0.01	0.01	0.015	0	0.02	-0.8	-2.54	0.98
Ice-cream																								
Demographic §	0.16	0.05	0.26	0.04	-0.06	0.14	-0.04	-0.07	0	0.08	-0.01	0.18	3.83	-1.43	9.37	-0.02	-0.04	0.01	0.02	0	0.05	-7.74	-12.8	-2.43
+ SES, lifestyle	0.13	0.02	0.24	0.02	-0.08	0.13	-0.04	-0.07	0	0.05	-0.04	0.14	6.07	0.83	11.58	-0.02	-0.05	0.01	0.02	0	0.05	-8.72	-13.7	-3.4
+ BMI	0.06	-0.05	0.17	0	-0.1	0.1	-0.03	-0.06	0.01	0.02	-0.07	0.11	2.31	-2.86	7.75	-0.01	-0.04	0.02	0.02	-0.01	0.04	-9.285	-14.2	-4.08
Low-fat fermented dairy products																								
Demographic §	-0.01	-0.04	0.02	-0.07	-0.1	-0.03	0	-0.02	0.01	-0.05	-0.08	-0.02	-2.43	-4.16	-0.66	-0.01	-0.02	-0.01	-0.01	-0.02	-0.01	0.8	-1.37	3.02
+ SES, lifestyle	0.03	-0.01	0.06	-0.03	-0.07	0.01	-0.02	-0.03	0	-0.02	-0.05	0.02	0.88	-0.94	2.74	-0.02	-0.02	-0.01	-0.01	-0.02	0	0.29	-2.02	2.65
+ BMI	0.02	-0.02	0.05	-0.03	-0.07	0.01	-0.015	-0.03	0	-0.02	-0.05	0.02	-0.07	-1.82	1.7	-0.01a	-0.02	0	-0.01	-0.02	0	0.12	-2.2	2.5
High-fat dairy products																								
Demographic §	0	-0.02	0.02	0.06	0.03	0.08	0.01	0	0.02	0.05	0.03	0.07	-1.15	-2.27	-0.01	0	0	0.01	0.01	0	0.01	-0.96	-2.14	0.24
+ SES, lifestyle	0.03	0.01	0.06	0.06	0.03	0.09	0	-0.01	0.01	0.06	0.03	0.08	0.4	-1.01	1.82	0	-0.01	0.01	0.01	0	0.02	-0.3	-1.8	1.22
+ BMI	0.035	0.01	0.06	0.064	0.03	0.09	0.01	0	0.01	0.054	0.03	0.08	0.35	-1	1.72	0	0	0.01	0.01*	0	0.02	-0.38	-1.87	1.13
Low-fat dairy products																								
Demographic §	0.02	0	0.04	-0.04	-0.06	-0.02	-0.02	-0.02	-0.01	-0.02	-0.04	0	-0.21	-1.31	0.9	-0.01	-0.02	-0.01	-0.01	-0.01	0	-0.79	-1.95	0.37
+ SES, lifestyle	0.04	0.02	0.06	-0.02	-0.05	0	-0.02	-0.03	-0.01	-0.01	-0.03	0.01	1.8	0.55	3.06	-0.01	-0.02	-0.01	0	-0.01	0	-0.32	-1.66	1.03
+ BMI	0.035	0	0.05	-0.035	-0.05	0	-0.024	-0.03	-0.01	-0.02	-0.04	0.01	1.08	-0.11	2.28	-0.01**	-0.02	-0.01	-0.01*	-0.01	0	-0.53	-1.85	0.82
Total dairy products																								
Demographic §	0.01	0	0.03	0.01	0	0.03	0	-0.01	0	0.02	0	0.03	-0.57	-1.45	0.31	0	-0.01	0	0	0	0.01	-1.05	-1.97	-0.13
+ SES, lifestyle	0.04	0.02	0.06	0.02	0	0.04	-0.01	-0.02	0	0.03	0.01	0.05	1.47	0.35	2.61	-0.01	-0.01	0	0	0	0.01	-0.62	-1.8	0.58
+ BMI	0.035	0.01	0.05	0.02	0	0.04	-0.015	-0.01	0	0.025	0	0.04	0.66	-0.42	1.76	-0.01*	-0.01	0	0	0	0.01	-0.86	-2.04	0.33

†Associations are per serving/day (Milk: 1 average glass (200g); Yoghurt: 125g carton; Cheese: medium serving (40g); Single cream: 1 tablespoon (15g); Double cream: 1 tablespoon (30g); Butter: 1 teaspoon (10g); Ice-cream: 1 average scoop/tub (60g) as defined by Food Standards Agency 2002)[197]

‡% change used for log-transformed outcomes after back-transformation (exponentiation) and defined as [(eb-1)*100]%

SModel 1: age (years), sex, test-site (Cambridge, Ely, Wisbech), ethnicity (white, non-white), total energy intake (kcal/d), mutual adjustment for dairy products.; Model 2: Model 1 + educational level (low, medium, high), age when full-time education finished (years), socio-economic status based on occupation (low: technical/semi-routine and routine occupations; medium: lower managerial / intermediate occupations; high: professional/higher managerial occupations), income (<£20,000, £20,000-40,000, >£40,000), marital status (single, married, widowed/separated), smoking status (never, former, current smoker), pack-years of smoking, energy expenditure due to physical activity (kj/kg/d), lipid-lowering medication (Yes, No), anti-hypertensive medication (Yes, No), hormone-replacement therapy (Yes, No, Men).; Model 3: Model 2 + intakes (g/d) of fruit, vegetables (not including potatoes), potatoes, legumes, processed cereals, whole-grain cereals, poultry and eggs, red meat, processed meat, fish, sauces, margarine, nuts, sweet snacks, sugar-sweetened beverages, artificially sweetened beverages, fruit juice, regular coffee, decaffeinated coffee, tea, alcoholic beverages, plasma vitamin C levels (μmol/l), dietary supplement use (Yes, No).; Model 4: Model 3 + BMI (kg/m^2)

*p<0.05 for the maximally adjusted model

**p<0.00025 (critical p-value after correction for multiple testing) for the maximally adjusted model

Abbreviations: ApoA1: Apolipoprotein A1; ApoB: Apolipoprotein B; NEFA: Non-Esterified Fatty Acids

Table A.7 Associations of total and types of dairy consumption with cardio-metabolic markers after adjustment for dairy nutrients †

	Low-fat dairy products			**Milk**			**High-fat dairy products**			**Total milk**			**High-fat dairy products**		
	VAT/SCAT			**Total lean mass (kg)**			**Total cholesterol (mmol/l)**			**HDL-C (mmol/l)**			**LDL-C (mmol/l)**		
N	11,253			11,523			11,989			11,988			11,897		
Nutrient	% change	95% CI		b	95% CI		b	95% CI		b	95% CI		b	95% CI	
Protein	-0.02	-0.05	0	0.39	0.17	0.61	0.04	-0.01	0.08	-0.02	-0.03	-0.01	0.04	-0.01	0.08
Total sugars	-0.02	-0.04	0	0.39	0.17	0.61	0.05	0.01	0.1	-0.02	-0.03	-0.01	0.04	0	0.08
Lactose	-0.03	-0.05	-0.01	0.39	0.17	0.61	0.06	0	0.11	-0.02	-0.03	-0.01	0.04	-0.01	0.09
Total fat	-0.03	-0.04	-0.02	0.38	0.23	0.53	0.02	-0.03	0.07	-0.02	-0.03	-0.01	0.03	-0.02	0.07
Saturated fat	-0.03	-0.04	-0.02	0.38	0.23	0.53	0.02	-0.03	0.07	-0.02	-0.03	-0.01	0.02	-0.02	0.07
Monounsaturated fat	-0.03	-0.04	-0.02	0.38	0.23	0.53	0.02	-0.03	0.07	-0.02	-0.03	-0.01	0.03	-0.02	0.07
Calcium	-0.03	-0.05	0	0.39	0.17	0.61	0.04	-0.01	0.09	-0.02	-0.03	-0.01	0.04	-0.01	0.08
Potassium	-0.03	-0.05	0	0.38	0.16	0.6	0.05	-0.01	0.11	-0.02	-0.03	-0.01	0.03	-0.02	0.09
Magnesium	-0.02	-0.05	0	0.38	0.16	0.6	0.03	-0.02	0.09	-0.02	-0.03	-0.01	0.04	-0.01	0.09
Phosphorous	-0.03	-0.05	0	0.39	0.17	0.61	0.04	-0.01	0.08	-0.02	-0.03	-0.01	0.04	-0.01	0.08
Zinc	-0.03	-0.05	0	0.39	0.17	0.61	0.04	-0.01	0.08	-0.02	-0.03	-0.01	0.04	-0.01	0.08
Selenium	-0.03	-0.05	0	0.39	0.17	0.61	0.04	-0.01	0.08	-0.02	-0.03	-0.01	0.04	-0.01	0.08
Vitamin A	-0.03	-0.04	-0.02	0.38	0.23	0.53	0.01	-0.04	0.06	-0.02	-0.03	-0.01	0.02	-0.03	0.06
Vitamin B12	-0.02	-0.04	0	0.39	0.17	0.61	0.03	-0.02	0.09	-0.02	-0.03	-0.01	0.04	-0.01	0.08

†Significant associations were selected after False Discovery Rate correction ; Serving sizes as defined by the Food Standards Agency 2002[197]: Milk: 1 average glass (200g); Yoghurt: 125g carton; Cheese: medium serving (40g); Single cream: 1 tablespoon (15g); Double cream: 1 tablespoon (30g); Butter: 1 teaspoon (10g); Ice-cream: 1 average scoop/tub (60g) ; Dairy nutrients intakes were calculated from the food frequency questionnaire using in-house software

Abbreviations: SCAT: Subcutaneous Adipose Tissue; VAT: Visceral Adipose Tissue

Table A.8 Associations of total and types of dairy consumption with glycaemic markers from multiple linear regression models

	HbA1c (mmol/mol)			**HOMA2-IR**			**HOMA2-%B**			**Fasting glucose (mmol/l)**			**2-hour glucose (mmol/l)**			**Fasting insulin (pmol/l)**		
Mean (SD)	37.0 (5.1)			0.9 (0.7)			90.9 (36.7)			4.8 (0.7)			5.3 (1.7)			48.1 (39.2)		
Participants (N)	11,960			9,637			9,637			11,969			11,795			9,662		
Dairy consumption (servings †/d)	b	95% CI		% change ‡	95% CI		% change ‡	95% CI		% change ‡	95% CI		% change ‡	95% CI		% change ‡	95% CI	
Milk																		
Demographic, energy §	0.11	0.03	0.19	-0.63	-2.28	1.04	-0.95	-2.02	0.14	0.11	-0.11	0.34	0.01	-0.64	0.67	-0.67	-2.32	0.99
+ SES, lifestyle, diet	0.07	-0.02	0.16	0.5	-1.29	2.33	-0.66	-1.84	0.54	0.27	0.02	0.52	0.63	-0.09	1.36	0.36	-1.42	2.17
+ BMI	0.03	-0.06	0.12	-1.3	-2.79	0.22	-1.264	-2.33	-0.18	0.16	-0.08	0.4	0.2	-0.49	0.9	-1.36	-2.85	0.15
Yoghurt																		
Demographic, energy §	-0.04	-0.18	0.11	-3.2	-6.03	-0.29	-1.68	-3.51	0.18	-0.33	-0.72	0.06	-1.09	-2.3	0.12	-3.1	-5.89	-0.22
+ SES, lifestyle, diet	0.07	-0.08	0.23	-0.33	-3.22	2.65	-0.01	-2.01	2.03	-0.19	-0.6	0.22	-1.09	-2.26	0.1	-0.26	-3.13	2.69
+ BMI	0.03	-0.13	0.18	-1.4	-3.71	0.96	-0.74	-2.44	1	-0.28	-0.67	0.12	-1.424	-2.52	-0.3	-1.26	-3.57	1.11
Cheese																		
Demographic, energy §	0.22	0.04	0.39	-2.48	-5.85	1	-2.65	-4.64	-0.62	-0.13	-0.58	0.33	-1.68	-3.26	-0.07	-2.7	-6.01	0.73
+ SES, lifestyle, diet	0.32	0.14	0.5	-2.42	-5.76	1.04	-2	-4.07	0.12	-0.1	-0.58	0.38	-0.78	-2.39	0.86	-2.64	-5.94	0.77
+ BMI	0.31*	0.13	0.49	-1.09	-4.01	1.93	-1.53	-3.47	0.45	-0.07	-0.55	0.41	-0.88	-2.38	0.65	-1.21	-4.15	1.82
Fermented dairy products																		
Demographic, energy §	0.07	-0.04	0.18	-2.9	-4.96	-0.79	-2.1	-3.39	-0.79	-0.25	-0.52	0.03	-1.32	-2.2	-0.44	-2.93	-4.97	-0.84
+ SES, lifestyle, diet	0.18	0.06	0.3	-1.2	-3.36	1.01	-0.88	-2.3	0.56	-0.16	-0.46	0.15	-0.97	-1.88	-0.05	-1.25	-3.39	0.93
+ BMI	0.144	0.03	0.26	-1.27	-3.1	0.59	-1.08	-2.34	0.2	-0.19	-0.49	0.1	-1.214	-2.06	-0.35	-1.24	-3.07	0.63
Full-fat milk																		
Demographic, energy §	0.27	0.11	0.42	-4.78	-7.91	-1.53	-3.38	-5.37	-1.35	-0.01	-0.5	0.47	-0.81	-1.99	0.38	-4.96	-8.07	-1.75
+ SES, lifestyle, diet	0.18	0.01	0.35	-2.72	-6.01	0.68	-2.72	-4.78	-0.62	0.27	-0.22	0.76	-0.16	-1.41	1.1	-3.03	-6.3	0.35
+ BMI	0.194	0.03	0.35	-1.13	-3.93	1.75	-1.39	-3.24	0.5	0.42	-0.04	0.89	-0.31	-1.58	0.97	-1.38	-4.17	1.5
Low-fat milk																		
Demographic, energy §	0.1	0.01	0.18	-0.22	-1.89	1.48	-0.7	-1.79	0.39	0.13	-0.1	0.35	0.12	-0.55	0.79	-0.25	-1.91	1.44
+ SES, lifestyle, diet	0.06	-0.03	0.16	0.65	-1.15	2.49	-0.56	-1.75	0.64	0.27	0.02	0.52	0.68	-0.04	1.42	0.51	-1.27	2.33
+ BMI	0.02	-0.07	0.11	-1.3	-2.8	0.21	-1.264	-2.33	-0.17	0.15	-0.09	0.39	0.24	-0.46	0.94	-1.36	-2.85	0.15
Full-fat yoghurt																		
Demographic, energy §	0.15	-0.22	0.52	-12.42	-18.61	-5.75	-5.14	-9.39	-0.68	-0.82	-1.96	0.33	-5.36	-8.3	-2.33	-12.02	-18.2	-5.37
+ SES, lifestyle, diet	0.48	0.1	0.85	-2.38	-9.41	5.2	-0.15	-4.68	4.59	-0.27	-1.47	0.94	-2.55	-5.62	0.63	-2.2	-9.16	5.3
+ BMI	0.494	0.11	0.87	1.49	-4.62	7.99	1.61	-2.76	6.17	-0.1	-1.23	1.05	-1.49	-4.65	1.77	1.86	-4.26	8.36
Low-fat yoghurt																		
Demographic, energy §	-0.06	-0.21	0.09	-1.87	-4.94	1.29	-1.2	-3.17	0.81	-0.27	-0.67	0.14	-0.51	-1.78	0.77	-1.81	-4.85	1.31
+ SES, lifestyle, diet	0.02	-0.14	0.19	-0.07	-3.12	3.08	0.01	-2.09	2.16	-0.18	-0.61	0.24	-0.91	-2.13	0.32	-0.02	-3.05	3.11
+ BMI	-0.03	-0.19	0.13	-1.74	-4.13	0.71	-1.03	-2.79	0.77	-0.3	-0.71	0.11	-1.414	-2.54	-0.26	-1.62	-4.01	0.83
High-fat cheese																		
Demographic, energy §	0.22	0	0.45	-3.71	-8.27	1.07	-2.69	-5.43	0.12	-0.52	-1.09	0.05	-1.85	-3.76	0.11	-3.82	-8.34	0.92
+ SES, lifestyle, diet	0.36	0.13	0.59	-2.72	-7.04	1.79	-1.13	-4.05	1.87	-0.56	-1.17	0.05	-0.91	-2.88	1.1	-2.8	-7.09	1.68
+ BMI	0.374	0.14	0.6	1.02	-2.73	4.91	0.1	-2.58	2.87	-0.47	-1.07	0.14	-0.56	-2.37	1.27	1.01	-2.76	4.92

Table A.8 (continued)

	HbA1c (mmol/mol)			**HOMA2-IR**			**HOMA2-%B**			**Fasting glucose (mmol/l)**			**2-hour glucose (mmol/l)**			**Fasting insulin (pmol/l)**		
Mean (SD)	37.0 (5.1)			0.9 (0.7)			90.9 (36.7)			4.8 (0.7)			5.3 (1.7)			48.1 (39.2)		
Participants (N)	11,960			9,637			9,637			11,969			11,795			9,662		
Dairy consumption (servings †/d)	b	95% CI		% change ‡	95% CI		% change ‡	95% CI		% change ‡	95% CI		% change ‡	95% CI		% change ‡	95% CI	
Low-fat cheese																		
Demographic, energy §	0.2	-0.08	0.48	-0.86	-5.79	4.32	-2.59	-5.48	0.39	0.49	-0.22	1.21	-1.4	-3.95	1.22	-1.23	-6.08	3.87
+ SES, lifestyle, diet	0.27	-0.02	0.56	-2.04	-7	3.18	-2.99	-5.97	0.09	0.52	-0.2	1.24	-0.58	-3.21	2.12	-2.44	-7.31	2.69
+ BMI	0.22	-0.06	0.51	-3.88	-8.42	0.88	-3.454	-6.2	-0.62	0.48	-0.25	1.22	-1.37	-3.87	1.21	-4.14	-8.65	0.59
Butter																		
Demographic, energy §	0.1	0	0.2	1.27	-0.65	3.22	1.1	-0.09	2.31	-0.09	-0.33	0.15	-1.22	-1.98	-0.46	1.32	-0.59	3.26
+ SES, lifestyle, diet	0.09	-0.02	0.21	2.1	-0.05	4.31	1.68	0.39	2.99	-0.23	-0.5	0.05	-0.67	-1.55	0.23	2.18	0.05	4.35
+ BMI	0.09	-0.02	0.2	0.79	-0.87	2.49	0.64	-0.52	1.82	-0.24	-0.51	0.03	-0.904	-1.72	-0.08	0.84	-0.83	2.54
Ice-cream																		
Demographic, energy §	0.11	-0.24	0.46	8.02	1.51	14.94	3.37	-0.83	7.74	0.36	-0.57	1.3	3.26	0.38	6.23	7.65	1.19	14.51
+ SES, lifestyle, diet	-0.04	-0.4	0.32	1.77	-3.89	7.76	0.36	-3.54	4.42	0.27	-0.67	1.21	1.09	-1.79	4.05	1.42	-4.21	7.38
+ BMI	-0.11	-0.47	0.26	-4.42	-9.15	0.56	-2.73	-6.24	0.9	-0.22	-1.12	0.7	-0.74	-3.44	2.03	-4.6	-9.3	0.35
Low-fat fermented dairy products																		
Demographic, energy §	0.02	-0.1	0.15	-2.68	-4.98	-0.32	-1.94	-3.4	-0.45	-0.16	-0.47	0.15	-1.18	-2.18	-0.17	-2.69	-4.97	-0.36
+ SES, lifestyle, diet	0.12	-0.01	0.26	-0.82	-3.2	1.63	-0.8	-2.38	0.8	-0.05	-0.39	0.29	-0.98	-1.99	0.03	-0.86	-3.22	1.55
+ BMI	0.08	-0.05	0.21	-1.83	-3.79	0.18	-1.36	-2.72	0.01	-0.12	-0.45	0.21	-1.374	-2.33	-0.41	-1.78	-3.74	0.23
High-fat dairy products																		
Demographic, energy §	0.13	0.06	0.2	-0.47	-1.92	1	-0.27	-1.21	0.68	-0.06	-0.25	0.13	-1.19	-1.75	-0.62	-0.49	-1.93	0.97
+ SES, lifestyle, diet	0.12	0.03	0.21	0.71	-1.02	2.47	0.54	-0.57	1.66	-0.12	-0.36	0.12	-0.75	-1.46	-0.04	0.67	-1.04	2.42
+ BMI	0.134	0.03	0.22	0.91	-0.49	2.33	0.38	-0.62	1.4	-0.07	-0.3	0.17	-0.804	-1.47	-0.12	0.88	-0.52	2.31
Low-fat dairy products																		
Demographic, energy §	0.08	0.01	0.14	-1.15	-2.5	0.23	-1.18	-2.04	-0.31	0.02	-0.16	0.21	-0.31	-0.87	0.25	-1.18	-2.53	0.19
+ SES, lifestyle, diet	0.09	0.01	0.17	0.03	-1.48	1.57	-0.7	-1.68	0.3	0.15	-0.07	0.36	0.11	-0.51	0.73	-0.08	-1.57	1.44
+ BMI	0.05	-0.03	0.13	-1.374	-2.64	-0.09	-1.214	-2.11	-0.31	0.05	-0.15	0.26	-0.29	-0.89	0.31	-1.404	-2.66	-0.12
Total dairy products																		
Demographic, energy §	0.1	0.04	0.15	-0.3	-1.4	0.82	-0.44	-1.14	0.27	-0.01	-0.16	0.14	-0.75	-1.19	-0.3	-0.32	-1.42	0.78
+ SES, lifestyle, diet	0.1	0.02	0.18	0.99	-0.38	2.38	0.31	-0.58	1.21	0.02	-0.17	0.21	-0.31	-0.88	0.27	0.92	-0.43	2.29
+ BMI	0.084	0	0.15	-0.37	-1.49	0.76	-0.46	-1.26	0.35	-0.05	-0.23	0.14	-0.664	-1.21	-0.11	-0.38	-1.5	0.75

†Associations are per serving/day (Milk: 1 average glass (200g); Yoghurt: 125g carton; Cheese: medium serving (40g); Single cream: 1 tablespoon (15g); Double cream: 1 tablespoon (30g); Butter: 1 teaspoon (10g); Ice-cream: 1 average scoop/tub (60g) as defined by Food Standards Agency 2002)citeFSA20021

‡% change used for log-transformed outcomes after back-transformation (exponentiation) and defined as [(eb-1)*100]%

SModel 1: age (years), sex, test-site (Cambridge, Ely, Wisbech), ethnicity (white, non-white), total energy intake (kcal/d), mutual adjustment for dairy products; Model 2: Model 1 + educational level (low, medium, high), age when full-time education finished (years), socio-economic status based on occupation (low: technical/semi-routine and routine occupations; medium: lower managerial / intermediate occupations; high: professional/higher managerial occupations), income (<£20,000, £20,000-40,000, >£40,000), marital status (single, married, widowed/separated), smoking status (never, former, current smoker), pack-years of smoking, energy expenditure due to physical activity (kj/kg/d), lipid-lowering medication (Yes, No), anti-hypertensive medication (Yes, No), hormone-replacement therapy (Yes, No, Men).; Model 3: Model 2 + intakes (g/d) of fruit, vegetables (not including potatoes), potatoes, legumes, processed cereals, whole-grain cereals, poultry and eggs, red meat, processed meat, fish, sauces, margarine, nuts, sweet snacks, sugar-sweetened beverages, artificially sweetened beverages, fruit juice, regular coffee, decaffeinated coffee, tea, alcoholic beverages, plasma vitamin C levels (μmol/l), dietary supplement use (Yes, No); Model 4: Model 3 + BMI (kg/m^2)

A.3 Chapter 5

Table A.9 Associations of total and types of dairy consumption with markers of hepatic function from multiple linear regression models

	Hepatic fat score			**ALT (IU/l)**			**GGT (IU/l)**		
Mean (SD)	4.3 (1.3)			28.9 (16.6)			34.8 (35.3)		
Participants (N)	10,108			11,982			11,986		
Dairy consumption (servings †/d)	b ‡	95% CI		% change §	95% CI		% change §	95% CI	
Milk									
Demographic, energy ‖	0	-0.02	0.01	1.36	0.38	2.35	0.89	-0.08	1.87
+ SES, lifestyle, diet	0	-0.01	0.01	2.09	0.98	3.22	1.69	0.62	2.78
+ BMI	-0.01	-0.02	0.01	1.27	0.21	2.35	1.17	0.15	2.21
Yoghurt									
Demographic, energy ‖	0	-0.02	0.02	1.04	-0.52	2.63	-2.91	-4.53	-1.26
+ SES, lifestyle, diet	0.01	-0.02	0.03	0.46	-1.22	2.17	-0.58	-2.29	1.17
+ BMI	0	-0.02	0.02	0.1	-1.6	1.83	-1.27	-2.97	0.45
Cheese									
Demographic, energy ‖	-0.01	-0.03	0.02	-1.55	-3.65	0.6	-0.78	-2.72	1.19
+ SES, lifestyle, diet	0	-0.03	0.02	-1.11	-3.25	1.08	0.57	-1.44	2.62
+ BMI	0	-0.02	0.03	-1.05	-3.17	1.11	0.93	-1.07	2.96
Fermented dairy products									
Demographic, energy ‖	0	-0.02	0.01	0.01	-1.16	1.2	-2.02	-3.19	-0.84
+ SES, lifestyle, diet	0	-0.01	0.02	-0.16	-1.47	1.17	-0.1	-1.38	1.19
+ BMI	0	-0.02	0.02	-0.36	-1.67	0.96	-0.36	-1.61	0.91
Full-fat milk									
Demographic, energy ‖	-0.02	-0.04	0	-0.52	-2.35	1.35	-1.03	-2.79	0.77
+ SES, lifestyle, diet	-0.02	-0.04	0.01	1.26	-0.73	3.29	-0.18	-2.14	1.82
+ BMI	-0.01	-0.04	0.01	1.51	-0.39	3.45	0.21	-1.59	2.04
Low-fat milk									
Demographic, energy ‖	0	-0.01	0.01	1.57	0.57	2.57	1.1	0.12	2.1
+ SES, lifestyle, diet	0	-0.01	0.01	2.14	1.02	3.27	1.79	0.72	2.88
+ BMI	-0.01	-0.02	0.01	1.265	0.19	2.34	1.23	0.2	2.27
Full-fat yoghurt									
Demographic, energy ‖	-0.07	-0.13	-0.02	0.46	-4.29	5.44	-10.91	-14.43	-7.24
+ SES, lifestyle, diet	-0.03	-0.09	0.02	2.01	-2.92	7.19	-4.39	-8.29	-0.32
+ BMI	-0.02	-0.07	0.04	2.91	-2.13	8.21	-3.16	-6.87	0.7
Low-fat yoghurt									
Demographic, energy ‖	0.01	-0.01	0.03	1.11	-0.5	2.75	-1.71	-3.47	0.07
+ SES, lifestyle, diet	0.01	-0.01	0.03	0.31	-1.4	2.05	-0.06	-1.88	1.79
+ BMI	0	-0.02	0.02	-0.19	-1.93	1.57	-1.01	-2.8	0.81
High-fat cheese									
Demographic, energy ‖	-0.02	-0.05	0.02	-3.22	-5.8	-0.57	-0.9	-3.16	1.42
+ SES, lifestyle, diet	-0.01	-0.05	0.02	-2.15	-4.91	0.7	0.77	-1.7	3.32
+ BMI	0	-0.04	0.03	-1.48	-4.21	1.32	1.44	-1.02	3.96
Low-fat cheese									
Demographic, energy ‖	0.01	-0.03	0.05	1.06	-2.23	4.45	-0.58	-3.92	2.87
+ SES, lifestyle, diet	0.01	-0.03	0.05	0.27	-2.92	3.57	0.28	-2.99	3.66
+ BMI	0.01	-0.03	0.05	-0.47	-3.69	2.85	0.17	-3.13	3.58
Butter									
Demographic, energy ‖	0	-0.01	0.01	0.54	-0.63	1.72	-0.69	-1.77	0.41
+ SES, lifestyle, diet	0	-0.01	0.02	1.46	0.16	2.78	-0.61	-1.78	0.58
+ BMI	0	-0.02	0.01	1.05	-0.16	2.29	-0.84	-1.96	0.29
Ice-cream									
Demographic, energy ‖	0.02	-0.03	0.07	1.64	-2.41	5.87	-4.1	-7.97	-0.08
+ SES, lifestyle, diet	0.01	-0.04	0.06	0.85	-3.24	5.1	-3.18	-6.98	0.77
+ BMI	-0.02	-0.07	0.03	-2.37	-5.98	1.38	-5.62	-9.28	-1.8

Table A.9

	Hepatic fat score			**ALT (IU/l)**			**GGT (IU/l)**		
Mean (SD)	4.3 (1.3)			28.9 (16.6)			34.8 (35.3)		
Participants (N)	10,108			11,982			11,986		
Dairy consumption (servings †/d)	b ‡	95% CI		% change §	95% CI		% change §	95% CI	
Low-fat fermented dairy products									
Demographic, energy ‖	0	-0.02	0.02	0.91	-0.4	2.25	-2.35	-3.69	-0.99
+ SES, lifestyle, diet	0.01	-0.01	0.02	0.32	-1.1	1.76	-0.34	-1.77	1.12
+ BMI	0	-0.02	0.02	-0.08	-1.52	1.38	-0.86	-2.29	0.58
High-fat dairy products									
Demographic, energy ‖	-0.01	-0.01	0	0.06	-0.79	0.92	-0.64	-1.45	0.18
+ SES, lifestyle, diet	0	-0.02	0.01	1.26	0.21	2.33	-0.42	-1.4	0.57
+ BMI	-0.01	-0.02	0.01	1.24	0.23	2.26	-0.31	-1.24	0.63
Low-fat dairy products									
Demographic, energy ‖	0	-0.01	0.01	1.26	0.46	2.07	-0.09	-0.88	0.71
+ SES, lifestyle, diet	0	-0.01	0.01	1.42	0.49	2.37	1.04	0.15	1.94
+ BMI	-0.01	-0.02	0.01	0.79	-0.11	1.71	0.56	-0.3	1.42
Total dairy products									
Demographic, energy ‖	0	-0.01	0.01	0.83	0.19	1.47	-0.43	-1.06	0.21
+ SES, lifestyle, diet	0	-0.01	0.01	1.63	0.79	2.48	0.29	-0.5	1.09
+ BMI	-0.01	-0.02	0	1.03	0.22	1.85	-0.1	-0.86	0.66

†Associations are per serving/day (Milk: 1 average glass (200g); Yoghurt: 125g carton; Cheese: medium serving (40g); Single cream: 1 tablespoon (15g); Double cream: 1 tablespoon (30g); Butter: 1 teaspoon (10g); Ice-cream: 1 average scoop/tub (60g) as defined by Food Standards Agency 2002)[197]

‡Beta coefficient and 95% CI derived from poisson regression

§% change used for log-transformed outcomes after back-transformation (exponentiation) and defined as [(eb-1)*100]%

parallelModel 1: age (years), sex, test-site (Cambridge, Ely, Wisbech), ethnicity (white, non-white), total energy intake (kcal/d), mutual adjustment for dairy products; Model 2: Model 1 + educational level (low, medium, high), age when full-time education finished (years), socio-economic status based on occupation (low: technical/semi-routine and routine occupations; medium: lower managerial / intermediate occupations; high: professional/higher managerial occupations), income (<£20,000, £20,000-40,000, >£40,000), marital status (single, married, widowed/separated), smoking status (never, former, current smoker), pack-years of smoking, energy expenditure due to physical activity (kj/kg/d), lipid-lowering medication (Yes, No), anti-hypertensive medication (Yes, No), hormone-replacement therapy (Yes, No, Men).; Model 3: Model 2 + intakes (g/d) of fruit, vegetables (not including potatoes), potatoes, legumes, processed cereals, whole-grain cereals, poultry and eggs, red meat, processed meat, fish, sauces, margarine, nuts, sweet snacks, sugar-sweetened beverages, artificially sweetened beverages, fruit juice, regular coffee, decaffeinated coffee, tea, alcoholic beverages, plasma vitamin C levels (μmol/l), dietary supplement use (Yes, No).; Model 4: Model 3 + BMI (kg/m^2)

Abbreviations: ALT: Alanine Transaminase; GGT: Gamma-Glutamyl Transferase

Table A.10 Associations of total and types of dairy consumption with blood pressure, CRP and adiponectin from multiple linear regression models

	Systolic blood pressure (mmHg)			**Diastolic blood pressure (mmHg)**			**CRP (mg/l)**			**Adiponectin (ug/ml)**		
Mean (SD)	123.5 (15.8)			74.5 (10.2)			2.9 (5.5)			7.1 (3.8)		
Participants (N)	12,063			12,062			9,438			9,642		
Dairy consumption (servings †/d)	b	95% CI		b	95% CI		% change ‡	95% CI		% change ‡	95% CI	
Milk												
Demographic, energy §	-0.28	-0.6	0.05	-0.28	-0.53	-0.04	3.69	0.2	7.3	-1.23	-2.44	0
+ SES, lifestyle, diet	0.11	-0.26	0.48	-0.15	-0.41	0.12	6.16	2.26	10.21	-1.52	-2.87	-0.16
+ BMI	0	-0.35	0.36	-0.23	-0.49	0.02	3.09	-0.22	6.51	-0.96	-2.28	0.39
Yoghurt												
Demographic, energy §	-0.25	-0.85	0.36	-0.33	-0.78	0.12	-6.58	-12.91	0.2	1.56	-0.78	3.96
+ SES, lifestyle, diet	0.01	-0.58	0.6	0.14	-0.36	0.64	5.06	-2.08	12.73	0.54	-1.79	2.92
+ BMI	-0.2	-0.78	0.38	-0.18	-0.63	0.27	1.84	-3.88	7.9	1.91	-0.39	4.25
Cheese												
Demographic, energy §	-0.78	-1.48	-0.07	-0.57	-1.05	-0.09	2.13	-5.38	10.24	1.88	-0.81	4.63
+ SES, lifestyle, diet	-0.21	-0.94	0.52	-0.22	-0.72	0.28	8.46	0.33	17.25	2.13	-0.61	4.95
+ BMI	-0.28	-0.99	0.43	-0.23	-0.72	0.26	5.12	-1.62	12.31	2.44	-0.18	5.13
Fermented dairy products												
Demographic, energy §	-0.48	-0.9	-0.05	-0.43	-0.75	-0.12	-2.86	-7.47	1.97	1.7	0.04	3.38
+ SES, lifestyle, diet	-0.08	-0.53	0.37	-0.02	-0.37	0.34	6.5	1.09	12.19	1.21	-0.56	3.01
+ BMI	-0.23	-0.67	0.2	-0.2	-0.54	0.13	3.2	-1.22	7.83	2.134	0.38	3.92
Full-fat milk												
Demographic, energy §	-1.15	-1.76	-0.53	-0.66	-1.18	-0.14	2.28	-3.97	8.93	-1.1	-3.3	1.16
+ SES, lifestyle, diet	-0.53	-1.21	0.14	-0.43	-0.96	0.1	5.48	-1.44	12.89	-1.41	-3.83	1.08
+ BMI	-0.29	-0.91	0.32	-0.22	-0.71	0.28	5.69	-0.14	11.86	-2.02	-4.32	0.33
Low-fat milk												
Demographic, energy §	-0.19	-0.52	0.14	-0.25	-0.5	0	3.84	0.31	7.5	-1.24	-2.48	0
+ SES, lifestyle, diet	0.14	-0.23	0.51	-0.13	-0.4	0.14	6.19	2.28	10.24	-1.53	-2.88	-0.16
+ BMI	0.02	-0.34	0.38	-0.23	-0.49	0.02	2.95	-0.37	6.37	-0.9	-2.23	0.46
Full-fat yoghurt												
Demographic, energy §	-2.21	-3.73	-0.69	-1.29	-2.43	-0.15	-35.88	-46.68	-22.88	8.1	1.44	15.2
+ SES, lifestyle, diet	-0.72	-2.34	0.9	-0.03	-1.18	1.13	-11.07	-26.01	6.89	3.82	-2.85	10.94
+ BMI	-0.43	-1.91	1.05	0.09	-1.02	1.21	-0.02	-14.92	17.5	1.69	-4.45	8.22
Low-fat yoghurt												
Demographic, energy §	0.04	-0.61	0.7	-0.19	-0.67	0.3	-1.92	-8.86	5.55	0.74	-1.69	3.24
+ SES, lifestyle, diet	0.1	-0.52	0.72	0.17	-0.36	0.69	7.18	-0.49	15.43	0.16	-2.25	2.63
+ BMI	-0.17	-0.78	0.44	-0.22	-0.69	0.25	2.04	-3.9	8.35	1.93	-0.45	4.37
High-fat cheese												
Demographic, energy §	-1.31	-2.26	-0.36	-0.93	-1.54	-0.31	0.65	-8.61	10.85	4.18	0.7	7.78
+ SES, lifestyle, diet	-0.39	-1.4	0.63	-0.44	-1.09	0.21	9.46	-1.45	21.58	4.6	1.04	8.28
+ BMI	-0.26	-1.24	0.73	-0.22	-0.88	0.44	10.13	0.42	20.79	3.77	0.43	7.23
Low-fat cheese												
Demographic, energy §	-0.05	-1.09	0.99	-0.06	-0.79	0.66	4.35	-7.76	18.04	-1.46	-5.51	2.76
+ SES, lifestyle, diet	0	-1.02	1.02	0.06	-0.67	0.8	7.23	-4.49	20.39	-1.11	-5.29	3.24
+ BMI	-0.31	-1.29	0.67	-0.24	-0.96	0.47	-0.39	-9.23	9.3	0.68	-3.35	4.87
Butter												
Demographic, energy §	-0.49	-0.89	-0.09	-0.27	-0.54	0.01	0.29	-3.56	4.29	-1.07	-2.48	0.36
+ SES, lifestyle, diet	-0.11	-0.58	0.36	-0.01	-0.32	0.3	2.07	-2.29	6.62	-1.2	-2.8	0.43
+ BMI	-0.19	-0.64	0.26	-0.09	-0.39	0.22	-1.09	-4.82	2.78	-0.06	-1.69	1.6
Ice-cream												
Demographic, energy §	-0.4	-1.73	0.92	0.3	-0.57	1.17	13.52	-1.94	31.42	-2.61	-7.91	2.99
+ SES, lifestyle, diet	0.02	-1.29	1.34	0.56	-0.34	1.46	10.59	-4.41	27.95	-1.74	-7.02	3.83
+ BMI	-0.62	-1.97	0.73	0.03	-0.87	0.93	0.78	-13.39	17.26	1.46	-4.17	7.42

Table A.10

	Systolic blood pressure (mmHg)			**Diastolic blood pressure (mmHg)**			CRP (mg/l)			**Adiponectin (ug/ml)**		
Mean (SD)	123.5 (15.8)			74.5 (10.2)			2.9 (5.5)			7.1 (3.8)		
Participants (N)	12,063			12,062			9,438			9,642		
Dairy consumption (servings †/d)	b	95% CI		b	95% CI		% change ‡	95% CI		% change ‡	95% CI	
Low-fat fermented dairy products												
Demographic, energy §	-0.23	-0.72	0.27	-0.29	-0.65	0.07	-3.89	-9.19	1.73	0.92	-0.97	2.85
+ SES, lifestyle, diet	0	-0.5	0.5	0.1	-0.3	0.5	5.72	-0.3	12.09	0.29	-1.66	2.28
+ BMI	-0.23	-0.71	0.26	-0.2	-0.57	0.17	1.57	-3.17	6.55	1.69	-0.26	3.67
High-fat dairy products												
Demographic, energy §	-0.67	-0.97	-0.37	-0.4	-0.62	-0.18	-0.07	-3.03	2.98	-0.47	-1.54	0.6
+ SES, lifestyle, diet	-0.23	-0.62	0.15	-0.16	-0.42	0.11	1.76	-1.95	5.62	-0.4	-1.71	0.93
+ BMI	-0.19	-0.56	0.19	-0.11	-0.37	0.14	0.48	-2.7	3.77	0.05	-1.26	1.38
Low-fat dairy products												
Demographic, energy §	-0.23	-0.51	0.05	-0.28	-0.48	-0.07	1.37	-1.53	4.36	-0.52	-1.54	0.51
+ SES, lifestyle, diet	0.07	-0.25	0.39	-0.08	-0.31	0.15	6.09	2.74	9.55	-0.84	-2	0.32
+ BMI	-0.07	-0.37	0.24	-0.224	-0.44	0	2.86	0.03	5.77	-0.06	-1.21	1.1
Total dairy products												
Demographic, energy §	-0.42	-0.64	-0.19	-0.31	-0.47	-0.15	0.75	-1.61	3.17	-0.56	-1.39	0.27
+ SES, lifestyle, diet	-0.01	-0.31	0.29	-0.06	-0.26	0.15	4.47	1.37	7.68	-0.88	-1.94	0.19
+ BMI	-0.14	-0.43	0.16	-0.18	-0.37	0.02	1.08	-1.52	3.74	0.11	-0.95	1.19

†Associations are per serving/day (Milk: 1 average glass (200g); Yoghurt: 125g carton; Cheese: medium serving (40g); Single cream: 1 tablespoon (15g); Double cream: 1 tablespoon (30g); Butter: 1 teaspoon (10g); Ice-cream: 1 average scoop/tub (60g) as defined by Food Standards Agency 2002)[197]

‡% change used for log-transformed outcomes after back-transformation (exponentiation) and defined as [(eb-1)*100]%

SModel 1: age (years), sex, test-site (Cambridge, Ely, Wisbech), ethnicity (white, non-white), total energy intake (kcal/d), mutual adjustment for dairy products; Model 2: Model 1 + educational level (low, medium, high), age when full-time education finished (years), socio-economic status based on occupation (low: technical/semi-routine and routine occupations; medium: lower managerial / intermediate occupations; high: professional/higher managerial occupations), income (<£20,000, £20,000-£40,000, >40,000), marital status (single, married, widowed/separated), smoking status (never, former, current smoker), pack-years of smoking, energy expenditure due to physical activity (kj/kg/d), lipid-lowering medication (Yes, No), anti-hypertensive medication (Yes, No), hormone-replacement therapy (Yes, No, Men).; Model 3: Model 2 + intakes (g/d) of fruit, vegetables (not including potatoes), potatoes, legumes, processed cereals, whole-gran cereals, poultry and eggs, red meat, processed meat, fish, sauces, margarine, nuts, sweet snacks, sugar-sweetened beverages, artificially sweetened beverages, fruit juice, regular coffee, decaffeinated coffee, tea, alcoholic beverages, plasma vitamin C levels (μmol/l), dietary supplement use (Yes, No).; Model 4: Model 3 + BMI (kg/m^2)

Abbreviations: CRP: C-Reactive Protein

Table A.11 Associations of total and types of dairy consumption with metabolic syndrome z-score from multiple linear regression models

	Metabolic syndrome z-score		
Mean (SD)	-0.002 (0.62)		
Participants (N)	11,931		
Dairy consumption (servings †/d)	b	95% CI	
Milk			
Demographic, energy ‡	0.02	0.01	0.04
+ SES, lifestyle, diet	0.04	0.03	0.06
Yoghurt			
Demographic, energy ‡	-0.01	-0.04	0.01
+ SES, lifestyle, diet	0.03	0	0.05
Cheese			
Demographic, energy ‡	-0.02	-0.05	0.02
+ SES, lifestyle, diet	0.02	-0.01	0.05
Fermented dairy products			
Demographic, energy ‡	-0.01	-0.03	0
+ SES, lifestyle, diet	0.03	0.01	0.04
Full-fat milk			
Demographic, energy ‡	-0.02	-0.05	0.01
+ SES, lifestyle, diet	0.01	-0.02	0.04
Low-fat milk			
Demographic, energy ‡	0.03	0.01	0.04
+ SES, lifestyle, diet	0.04	0.03	0.06
Full-fat yoghurt			
Demographic, energy ‡	-0.2	-0.27	-0.13
+ SES, lifestyle, diet	-0.04	-0.1	0.02
Low-fat yoghurt			
Demographic, energy ‡	0.01	-0.02	0.04
+ SES, lifestyle, diet	0.04	0.01	0.06
High-fat cheese			
Demographic, energy ‡	-0.06	-0.11	-0.02
+ SES, lifestyle, diet	0	-0.04	0.04
Low-fat cheese			
Demographic, energy ‡	0.05	0.01	0.1
+ SES, lifestyle, diet	0.05	0.01	0.09
Butter			
Demographic, energy ‡	-0.01	-0.03	0
+ SES, lifestyle, diet	0.01	-0.01	0.03
Ice-cream			
Demographic, energy ‡	0.08	0.02	0.13
+ SES, lifestyle, diet	0.07	0.02	0.13

Table A.11

	Metabolic syndrome z-score		
Mean (SD)	-0.002 (0.62)		
Participants (N)	11,931		
Dairy consumption (servings †/d)	b	95% CI	
Low-fat fermented dairy products			
Demographic, energy ‡	0	-0.02	0.02
+ SES, lifestyle, diet	0.03	0.01	0.05
High-fat dairy products			
Demographic, energy ‡	-0.02	-0.03	-0.01
+ SES, lifestyle, diet	0	-0.01	0.02
Low-fat dairy products			
Demographic, energy ‡	0.02	0	0.03
+ SES, lifestyle, diet	0.04	0.02	0.05
Total dairy products			
Demographic, energy ‡	0	-0.01	0.01
+ SES, lifestyle, diet	0.03	0.01	0.04

†Associations are per serving/day (Milk: 1 average glass (200g); Yoghurt: 125g carton; Cheese: medium serving (40g); Single cream: 1 tablespoon (15g); Double cream: 1 tablespoon (30g); Butter: 1 teaspoon (10g); Ice-cream: 1 average scoop/tub (60g) as defined by Food Standards Agency 2002)[197]

‡Model 1: age (years), sex, test-site (Cambridge, Ely, Wisbech), ethnicity (white, non-white), total energy intake (kcal/d), mutual adjustment for dairy products; Model 2: Model 1 + educational level (low, medium, high), age when full-time education finished (years), socio-economic status based on occupation (low: technical/semi-routine and routine occupations; medium: lower managerial / intermediate occupations; high: professional/higher managerial occupations), income (<£20,000, £20,000-40,000, >£40,000), marital status (single, married, widowed/separated), smoking status (never, former, current smoker), pack-years of smoking, energy expenditure due to physical activity (kj/kg/d), lipid-lowering medication (Yes, No), anti-hypertensive medication (Yes, No), hormone-replacement therapy (Yes, No, Men).; Model 3: Model 2 + intakes (g/d) of fruit, vegetables (not including potatoes), potatoes, legumes, processed cereals, whole-grain cereals, poultry and eggs, red meat, processed meat, fish, sauces, margarine, nuts, sweet snacks, sugar-sweetened beverages, artificially sweetened beverages, fruit juice, regular coffee, decaffeinated coffee, tea, alcoholic beverages, plasma vitamin C levels (μmol/l), dietary supplement use (Yes, No).

Table A.12 Associations of the change in total and types of dairy consumption with the change in markers of body weight and composition from baseline to the first follow-up after a mean of 3.7 years in the EPIC-Norfolk study

	Weight (kg)			**BMI (kg/m^2) †**			**Waist (cm)**			**Waist / Hip circumference**		
Mean (SD) of change	1.3 (4)			0.6 (1.4)			0.8 (5.5)			0 (0)		
Participants (N)	14,044			14,134			14,227			14,213		
Dairy consumption (servings/d) ‡	b	95% CI		b	95% CI		b	95% CI		b	95% CI	
Milk												
Demographic, lifestyle, energy §	0.05	-0.06	0.16	0.02	-0.02	0.06	-0.03	-0.18	0.13	-0.001	-0.002	0.001
+ Diet ‖	0.05	-0.06	0.17	0.02	-0.02	0.06	-0.03	-0.2	0.13	-0.001	-0.002	0.001
+ BMI #								-0.19	0.06	-0.001	-0.002	0.001
Yoghurt												
Demographic, lifestyle, energy §	-0.27	-0.49	-0.05	-0.11	-0.19	-0.02	-0.41	-0.71	-0.1	-0.002	-0.004	0.001
+ Diet ‖	-0.23	-0.46	-0.01	-0.09	-0.18	0	-0.38	-0.68	-0.07	-0.002	-0.004	0.001
+ BMI #								-0.41	0.14	-0.001	-0.003	0.002
Cheese												
Demographic, lifestyle, energy §	-0.18	-0.44	0.08	-0.07	-0.17	0.03	-0.36	-0.72	-0.01	-0.002	-0.005	0
+ Diet ‖	-0.14	-0.39	0.12	-0.06	-0.15	0.04	-0.34	-0.7	0.02	-0.002	-0.005	0
+ BMI #								-0.52	0.14	-0.002	-0.005	0.001
Butter												
Demographic, lifestyle, energy §	0.08	-0.03	0.18	0.02	-0.02	0.06	0.16	0.002	0.32	0.001	0	0.002
+ Diet ‖	0.08	-0.03	0.19	0.02	-0.02	0.06	0.16	-0.01	0.33	0.001	0	0.002
+ BMI #								-0.03	0.24	0.001	0	0.002
Full-fat milk												
Demographic, lifestyle, energy §	0.25	0.03	0.46	0.09	0.01	0.18	0.27	-0.02	0.57	0	-0.002	0.002
+ Diet ‖	0.23	0.02	0.45	0.09	0.01	0.17	0.25	-0.04	0.55	0	-0.002	0.002
+ BMI #								-0.19	0.28	-0.001	-0.003	0.002
Low-fat milk												
Demographic, lifestyle, energy §	0.05	-0.06	0.15	0.02	-0.02	0.06	0.01	-0.15	0.17	0	-0.002	0.001
+ Diet ‖	0.05	-0.06	0.16	0.02	-0.02	0.06	-0.01	-0.16	0.15	0	-0.002	0.001
+ BMI #								-0.17	0.08	-0.001	-0.002	0.001

Table A.12 (continued)

	Weight (kg)			BMI (kg/m^2) †			**Waist (cm)**			**Waist / Hip circumference**		
Mean (SD) of change	1.3 (4)			0.6 (1.4)			0.8 (5.5)			0 (0)		
Participants (N)	14,044			14,134			14,227			14,213		
Dairy consumption (servings/d) ‡	b	95% CI		b	95% CI		b	95% CI		b	95% CI	
Full-fat yoghurt												
Demographic, lifestyle, energy §	-0.07	-0.9	0.75	-0.03	-0.31	0.26	0.02	-1.06	1.1	0.002	-0.01	0.01
+ Diet ‖	-0.04	-0.86	0.78	-0.01	-0.3	0.27	0.02	-1.06	1.1	0.002	-0.01	0.01
+ BMI #								-0.93	1.02	0.002	-0.01	0.01
Low-fat yoghurt												
Demographic, lifestyle, energy §	-0.34	-0.59	-0.1	-0.14	-0.23	-0.05	-0.44	-0.79	-0.09	-0.002	-0.01	0.001
+ Diet ‖	-0.31	-0.56	-0.06	-0.13	-0.22	-0.04	-0.41	-0.77	-0.05	-0.002	-0.005	0.001
+ BMI #								-0.38	0.19	-0.001	-0.004	0.002
High-fat cheese												
Demographic, lifestyle, energy §	0.51	0.16	0.86	0.18	0.06	0.3	0.48	-0.06	1.01	0.002	-0.003	0.01
+ Diet ‖	0.48	0.13	0.84	0.17	0.05	0.29	0.41	-0.11	0.94	0.001	-0.003	0.01
+ BMI #								-0.44	0.42	0	-0.004	0.004
Low-fat cheese												
Demographic, lifestyle, energy §	-0.73	-1.07	-0.4	-0.27	-0.39	-0.15	-1.08	-1.57	-0.59	-0.01	-0.01	-0.001
+ Diet ‖	-0.64	-0.97	-0.31	-0.23	-0.36	-0.11	-1.01	-1.5	-0.51	-0.01	-0.01	-0.001
+ BMI #								-0.82	0.1	-0.003	-0.01	0.001
Ice-cream												
Demographic, lifestyle, energy §	0.24	-0.07	0.56	0.08	-0.04	0.19	0.2	-0.35	0.74	0	-0.004	0.004
+ Diet ‖	0.23	-0.08	0.55	0.07	-0.04	0.19	0.16	-0.37	0.7	0	-0.004	0.004
+ BMI #								-0.44	0.47	-0.001	-0.004	0.003
Fermented dairy products												
Demographic, lifestyle, energy §	-0.23	-0.37	-0.08	-0.09	-0.14	-0.03	-0.31	-0.58	-0.04	-0.002	-0.004	0
+ Diet ‖	-0.2	-0.34	-0.06	-0.07	-0.13	-0.02	-0.3	-0.56	-0.04	-0.002	-0.004	0
+ BMI #								-0.3	0.11	-0.001	-0.003	0.001

Table A.12 (continued)

	Weight (kg)			BMI (kg/m^2) †			**Waist (cm)**			**Waist / Hip circumference**		
Mean (SD) of change	1.3 (4)			0.6 (1.4)			0.8 (5.5)			0 (0)		
Participants (N)	14,044			14,134			14,227			14,213		
Dairy consumption (servings/d) ‡	b	95% CI		b	95% CI		b	95% CI		b	95% CI	
High-fat dairy products												
Demographic, lifestyle, energy §	0.13	0.05	0.21	0.04	0.02	0.07	0.12	0.01	0.24	0	-0.001	0.001
+ Diet ‖	0.13	0.05	0.21	0.04	0.01	0.07	0.11	-0.01	0.22	0	-0.001	0.001
+ BMI #								-0.09	0.13	0	-0.001	0.001
Low-fat dairy products												
Demographic, lifestyle, energy §	-0.08	-0.16	-0.01	-0.03	-0.06	-0.002	-0.15	-0.28	-0.02	-0.001	-0.002	0
+ Diet ‖	-0.07	-0.14	0.01	-0.02	-0.05	0.004	-0.15	-0.29	-0.02	-0.001	-0.002	0
+ BMI #								-0.2	0.05	-0.001	-0.002	0
Total dairy products												
Demographic, lifestyle, energy §	0.01	-0.06	0.07	0	-0.02	0.03	-0.01	-0.1	0.09	0	-0.001	0
+ Diet ‖	0.02	-0.05	0.09	0	-0.02	0.03	-0.01	-0.12	0.09	0	-0.001	0.001
+ BMI #								-0.08	0.08	0	-0.001	0.001

† BMI: Body mass index

‡ Servings as defined by Food Standards Agency 2002citeFSA20021: Milk- 1 average glass (200g); Yoghurt- 125g carton; Cheese- medium serving (40g); Single cream- 1 tablespoon (15g); Double cream- 1 tablespoon (30g); Butter- 1 teaspoon (10g); Ice-cream- 1 average scoop/tub (60g)

§ Linear regression model 1: age (years), sex, educational level (low, medium, high), age at completion of full-time education (years), marital status (single, married, widowed or separated), socio-economic status based on occupation (low: technical/semi-routine and routine occupationsl medium: lower managerial / intermediate occupationsl; high: professional / higher managerial occupations), individual follow-up time (years), physical activity level (inactive, moderately inactive, moderately active, active), smoking status (never, former and current smoker), lipid-lowering medication (Yes, No), anti-hypertensive medication (Yes, No), hormone-replacement therapy (Yes, No, Men), total energy intake (kcal/day)

‖ Linear regression model 2: Model 1 + intakes (g/d) of fruit, vegetables, potatoes, legumes, nuts, processed cereals, whole-grain cereals, poultry and eggs, red meat, processed meat, fish, sauces, margarine, sweet snacks, sugar-sweetened beverages, artificially sweetened beverages, fruit juice, coffee, tea and alcoholic beverages and dietary supplement use (Yes, No)

Linear regression model 3: Model 2 + BMI (kg/m^2)

Table A.13 Associations of the change in total and types of dairy consumption with the change in lipid markers from baseline to the first follow-up after a mean of 3.7 years in the EPIC-Norfolk study

	Total / HDL-C †			**Total cholesterol (mmol/l)**			**HDL-C (mmol/l) †**			**LDL-C (mmol/l) †**			**Triglycerides (mmol/l)**		
Mean (SD) of change	-0.2 (1.2)			-0.1 (1)			0.1 (0.3)			-0.2 (0.9)			0.1 (0.9)		
Participants (N)	12,959			13,350			12,993			12,963			13,302		
Dairy consumption (servings/d) ‡	b	95% CI		b	95% CI		b	95% CI		b	95% CI		b	95% CI	
Milk															
Demographic, lifestyle, energy §	-0.02	-0.05	0.01	-0.03	-0.06	0.01	0.002	-0.01	0.01	-0.04	-0.06	-0.01	0.02	-0.01	0.05
+ Diet ‖	-0.02	-0.05	0.02	-0.03	-0.06	0.01	0.003	-0.01	0.01	-0.04	-0.07	-0.01	0.02	-0.01	0.05
+ BMI #	-0.02	-0.05	0.01	-0.03	-0.06	0.01	0.003	-0.01	0.01	-0.04	-0.07	-0.01	0.02	-0.01	0.04
Yoghurt															
Demographic, lifestyle, energy §	-0.01	-0.07	0.05	-0.08	-0.14	-0.03	-0.02	-0.04	-0.01	-0.02	-0.07	0.02	-0.06	-0.12	-0.01
+ Diet ‖	-0.01	-0.07	0.05	-0.08	-0.14	-0.02	-0.02	-0.04	-0.004	-0.02	-0.07	0.03	-0.06	-0.11	-0.01
+ BMI #	0.01	-0.05	0.07	-0.06	-0.12	-0.01	-0.02	-0.04	-0.01	-0.01	-0.06	0.04	-0.04	-0.09	0.005
Cheese															
Demographic, lifestyle, energy §	-0.02	-0.09	0.05	0.04	-0.01	0.1	0.02	-0.005	0.04	0.02	-0.03	0.08	0.01	-0.05	0.06
+ Diet ‖	-0.02	-0.09	0.05	0.04	-0.01	0.1	0.02	-0.004	0.04	0.02	-0.03	0.07	0.01	-0.04	0.06
+ BMI #	-0.01	-0.08	0.05	0.05	-0.01	0.1	0.02	-0.005	0.04	0.02	-0.03	0.08	0.02	-0.03	0.07
Butter															
Demographic, lifestyle, energy §	0.03	0.003	0.07	0.06	0.04	0.08	0.003	-0.01	0.01	0.05	0.03	0.07	0.004	-0.02	0.03
+ Diet ‖	0.03	-0.01	0.06	0.06	0.03	0.08	0.004	-0.01	0.02	0.05	0.02	0.07	0.002	-0.03	0.03
+ BMI #	0.02	-0.01	0.06	0.05	0.03	0.08	0.01	-0.01	0.02	0.05	0.02	0.07	-0.001	-0.03	0.03
Full-fat milk															
Demographic, lifestyle, energy §	0.03	-0.02	0.09	0.03	-0.02	0.08	0	-0.02	0.02	0.04	-0.004	0.09	0.002	-0.04	0.04
+ Diet ‖	0.03	-0.02	0.09	0.03	-0.02	0.08	0	-0.02	0.02	0.04	-0.005	0.09	-0.001	-0.04	0.04
+ BMI #	0.02	-0.04	0.07	0.02	-0.03	0.07	0.002	-0.01	0.02	0.03	-0.01	0.08	-0.01	-0.06	0.03
Low-fat milk															
Demographic, lifestyle, energy §	-0.02	-0.05	0.01	-0.02	-0.05	0.02	0.001	-0.01	0.01	-0.03	-0.06	-0.01	0.02	-0.001	0.05
+ Diet ‖	-0.01	-0.04	0.01	-0.01	-0.04	0.02	0.001	-0.01	0.01	-0.03	-0.06	-0.01	0.03	0.001	0.05
+ BMI #	-0.02	-0.04	0.01	-0.02	-0.05	0.01	0.002	-0.01	0.01	-0.03	-0.06	-0.01	0.02	0	0.05

Table A.13 (continued)

	Total / HDL-C †			**Total cholesterol (mmol/l)**			**HDL-C (mmol/l) †**			**LDL-C (mmol/l) †**			**Triglycerides (mmol/l)**		
Mean (SD) of change	-0.2 (1.2)			-0.1 (1)			0.1 (0.3)			-0.2 (0.9)			0.1 (0.9)		
Participants (N)	12,959			13,350			12,993			12,963			13,302		
Dairy consumption (servings/d) ‡	b	95% CI		b	95% CI		b	95% CI		b	95% CI		b	95% CI	
Full-fat yoghurt															
Demographic, lifestyle, energy §	0.19	-0.03	0.41	0.1	-0.08	0.29	-0.04	-0.11	0.03	0.13	-0.05	0.3	-0.01	-0.18	0.15
+ Diet ‖	0.18	-0.04	0.4	0.08	-0.1	0.27	-0.04	-0.11	0.03	0.11	-0.07	0.28	-0.01	-0.17	0.16
+ BMI #	0.17	-0.05	0.4	0.08	-0.1	0.26	-0.04	-0.11	0.03	0.1	-0.07	0.28	-0.01	-0.17	0.15
Low-fat yoghurt															
Demographic, lifestyle, energy §	-0.01	-0.08	0.05	-0.09	-0.14	-0.03	-0.02	-0.04	0	-0.03	-0.08	0.02	-0.07	-0.12	-0.02
+ Diet ‖	-0.02	-0.08	0.05	-0.08	-0.14	-0.02	-0.02	-0.04	0.002	-0.03	-0.08	0.03	-0.07	-0.12	-0.01
+ BMI #	0.01	-0.05	0.07	-0.06	-0.12	-0.01	-0.02	-0.04	0	-0.01	-0.06	0.04	-0.05	-0.1	0
High-fat cheese															
Demographic, lifestyle, energy §	0.03	-0.07	0.13	0.16	0.06	0.25	0.03	0.01	0.06	0.11	0.04	0.19	0.02	-0.05	0.1
+ Diet ‖	0.02	-0.07	0.12	0.15	0.06	0.24	0.03	0.01	0.06	0.11	0.04	0.18	0.02	-0.06	0.09
+ BMI #	-0.01	-0.1	0.08	0.12	0.04	0.21	0.04	0.01	0.07	0.09	0.02	0.16	-0.01	-0.09	0.06
Low-fat cheese															
Demographic, lifestyle, energy §	-0.08	-0.18	0.02	-0.07	-0.17	0.03	0.001	-0.03	0.03	-0.05	-0.15	0.05	-0.03	-0.11	0.04
+ Diet ‖	-0.08	-0.18	0.03	-0.06	-0.16	0.04	0.003	-0.03	0.03	-0.05	-0.15	0.06	-0.02	-0.1	0.06
+ BMI #	-0.04	-0.14	0.06	-0.03	-0.13	0.07	-0.003	-0.03	0.03	-0.03	-0.13	0.08	0.02	-0.06	0.09
Ice-cream															
Demographic, lifestyle, energy §	0	-0.08	0.08	0.04	-0.03	0.12	-0.01	-0.03	0.02	-0.005	-0.07	0.06	0.08	0.02	0.15
+ Diet ‖	0.01	-0.07	0.1	0.05	-0.03	0.13	-0.01	-0.03	0.02	0.002	-0.06	0.07	0.08	0.02	0.15
+ BMI #	0.01	-0.08	0.09	0.04	-0.03	0.12	0	-0.03	0.02	-0.001	-0.07	0.07	0.07	0.01	0.14
Fermented dairy products															
Demographic, lifestyle, energy §	-0.03	-0.07	0.01	-0.02	-0.06	0.02	0.002	-0.01	0.01	-0.003	-0.04	0.03	-0.03	-0.07	0
+ Diet ‖	-0.03	-0.07	0.01	-0.02	-0.06	0.02	0.004	-0.01	0.02	-0.004	-0.04	0.03	-0.03	-0.06	0.002
+ BMI #	-0.02	-0.06	0.02	-0.01	-0.04	0.03	0.002	-0.01	0.02	0.005	-0.03	0.04	-0.02	-0.05	0.01

Table A.13 (continued)

	Total / HDL-C †			**Total cholesterol (mmol/l)**			**HDL-C (mmol/l) †**			**LDL-C (mmol/l) †**			**Triglycerides (mmol/l)**		
Mean (SD) of change	-0.2 (1.2)			-0.1 (1)			0.1 (0.3)			-0.2 (0.9)			0.1 (0.9)		
Participants (N)	12,959			13,350			12,993			12,963			13,302		
Dairy consumption (servings/d) ‡	b	95% CI		b	95% CI		b	95% CI		b	95% CI		b	95% CI	
High-fat dairy products															
Demographic, lifestyle, energy §	0.03	0.004	0.05	0.06	0.03	0.08	0.004	-0.002	0.01	0.04	0.02	0.06	0.01	-0.01	0.03
+ Diet ‖	0.02	-0.002	0.05	0.05	0.03	0.08	0.005	-0.002	0.01	0.04	0.02	0.06	0.01	-0.01	0.03
+ BMI #	0.02	-0.01	0.04	0.05	0.02	0.07	0.01	-0.001	0.01	0.04	0.02	0.06	0.001	-0.02	0.02
Low-fat dairy products															
Demographic, lifestyle, energy §	-0.02	-0.05	-0.002	-0.03	-0.06	-0.01	-0.002	-0.01	0.004	-0.03	-0.05	-0.01	0.002	-0.02	0.02
+ Diet ‖	-0.02	-0.05	0.001	-0.03	-0.05	-0.01	-0.001	-0.01	0.01	-0.03	-0.05	-0.01	0.004	-0.02	0.02
+ BMI #	-0.02	-0.04	0.01	-0.03	-0.05	-0.003	-0.002	-0.01	0.005	-0.03	-0.05	-0.01	0.01	-0.01	0.03
Total dairy products															
Demographic, lifestyle, energy §	0.003	-0.02	0.02	0.02	0	0.04	0.002	-0.004	0.01	0.01	-0.01	0.02	0.01	-0.003	0.03
+ Diet ‖	0.001	-0.02	0.02	0.02	0	0.03	0.002	-0.004	0.01	0.01	-0.01	0.02	0.01	-0.002	0.03
+ BMI #	0.002	-0.02	0.02	0.02	0	0.04	0.003	-0.003	0.01	0.01	-0.01	0.02	0.01	-0.001	0.03

† HDL-C: High density lipoprotein cholesterol; LDL-C: Low density lipoprotein cholesterol

‡ Servings as defined by Food Standards Agency 2002citeFSA20021: Milk- 1 average glass (200g); Yoghurt- 125g carton; Cheese- medium serving (40g); Single cream- 1 tablespoon (15g); Double cream- 1 tablespoon (30g); Butter- 1 teaspoon (10g); Ice-cream- 1 average scoop/tub (60g)

§ Linear regression model 1: age (years), sex, educational level (low, medium, high), age at completion of full-time education (years), marital status (single, married, widowed or separated), socio-economic status based on occupation (low: technical/semi-routine and routine occupationsl medium: lower managerial / intermediate occupationsl; high: professional / higher managerial occupations), individual follow-up time (years), physical activity level (inactive, moderately inactive, moderately active, active), smoking status (never, former and current smoker), lipid-lowering medication (Yes, No), anti-hypertensive medication (Yes, No), hormone-replacement therapy (Yes, No, Men), total energy intake (kcal/day)

‖ Linear regression model 2: Model 1 + intakes (g/d) of fruit, vegetables, potatoes, legumes, nuts, processed cereals, whole-grain cereals, poultry and eggs, red meat, processed meat, fish, sauces, margarine, sweet snacks, sugar-sweetened beverages, artificially sweetened beverages, fruit juice, coffee, tea and alcoholic beverages and dietary supplement use (Yes, No)

Linear regression model 3: Model 2 + BMI (kg/m^2)

Table A.14 Associations of the change in total and types of dairy consumption with the change in HbA1c, blood pressure and a z-score for metabolic risk from baseline to the first follow-up after a mean of 3.7 years in the EPIC-Norfolk study

	HbA1c (mmol/mol) †			**Systolic blood pressure (mmHg)**			**Diastolic blood pressure (mmHg)**			**Metabolic risk z-score**		
Mean (SD) of change	1.6 (6.1)			0.5 (14.9)			-0.1 (10.5)			0.01 (0.34)		
Participants (N)	6,224			14,210			14,231			6,033		
Dairy consumption (servings/d) ‡	b	95% CI		b	95% CI		b	95% CI		b	95% CI	
Milk												
Demographic, lifestyle, energy §	0.07	-0.18	0.31	-0.08	-0.52	0.36	0.11	-0.21	0.42	0.002	-0.02	0.02
+ Diet ‖	0.1	-0.14	0.35	-0.1	-0.55	0.35	0.1	-0.22	0.43	0.003	-0.02	0.02
+ BMI #	0.1	-0.14	0.34	-0.13	-0.56	0.31	0.09	-0.24	0.41			
Yoghurt												
Demographic, lifestyle, energy §	0.1	-0.4	0.61	-0.16	-1.03	0.7	-0.38	-1.03	0.27	-0.01	-0.05	0.02
+ Diet ‖	0.16	-0.35	0.66	-0.09	-0.98	0.81	-0.29	-0.96	0.38	-0.01	-0.04	0.03
+ BMI #	0.21	-0.29	0.71	0.14	-0.71	0.98	-0.14	-0.78	0.5			
Cheese												
Demographic, lifestyle, energy §	-0.11	-0.67	0.46	-0.04	-0.94	0.86	-0.15	-0.79	0.49	-0.03	-0.06	0.01
+ Diet ‖	-0.08	-0.64	0.48	-0.01	-0.92	0.89	-0.13	-0.78	0.51	-0.02	-0.06	0.01
+ BMI #	-0.04	-0.6	0.52	0.08	-0.8	0.96	-0.07	-0.69	0.56			
Butter												
Demographic, lifestyle, energy §	0.02	-0.22	0.27	0.37	-0.03	0.76	0.23	-0.08	0.53	-0.002	-0.02	0.01
+ Diet ‖	0.01	-0.24	0.25	0.35	-0.07	0.76	0.22	-0.09	0.53	-0.001	-0.02	0.02
+ BMI #	0	-0.24	0.24	0.3	-0.12	0.72	0.19	-0.12	0.51			
Full-fat milk												
Demographic, lifestyle, energy §	0.54	0.08	1	0.1	-0.79	0.99	0.24	-0.4	0.87	0.02	-0.01	0.05
+ Diet ‖	0.55	0.09	1.01	0.07	-0.83	0.96	0.2	-0.43	0.83	0.02	-0.02	0.05
+ BMI #	0.52	0.06	0.97	-0.1	-0.98	0.78	0.07	-0.54	0.69			
Low-fat milk												
Demographic, lifestyle, energy §	0	-0.26	0.27	-0.05	-0.62	0.52	0.12	-0.25	0.49	0.002	-0.02	0.02
+ Diet ‖	0.04	-0.23	0.31	-0.05	-0.62	0.51	0.12	-0.25	0.49	0.003	-0.02	0.02
+ BMI #	0.04	-0.23	0.31	-0.09	-0.65	0.48	0.1	-0.26	0.47			
Full-fat yoghurt												
Demographic, lifestyle, energy §	-0.06	-1.7	1.59	1.33	-1.77	4.43	0.14	-1.95	2.23	0.02	-0.08	0.13
+ Diet ‖	-0.02	-1.67	1.64	1.32	-1.8	4.44	0.08	-2.04	2.2	0.03	-0.07	0.14
+ BMI #	-0.01	-1.68	1.66	1.29	-1.75	4.32	0.05	-2.02	2.13			
Low-fat yoghurt												
Demographic, lifestyle, energy §	0.22	-0.36	0.8	-0.2	-1.04	0.65	-0.37	-0.95	0.22	-0.02	-0.06	0.02
+ Diet ‖	0.26	-0.3	0.82	-0.12	-0.97	0.72	-0.27	-0.87	0.33	-0.01	-0.05	0.03
+ BMI #	0.33	-0.23	0.89	0.16	-0.66	0.99	-0.07	-0.66	0.51			
High-fat cheese												
Demographic, lifestyle, energy §	-0.08	-0.81	0.65	1.33	0	2.65	0.75	-0.21	1.7	0.01	-0.05	0.06
+ Diet ‖	-0.05	-0.76	0.67	1.27	-0.05	2.6	0.67	-0.28	1.62	0.01	-0.05	0.06
+ BMI #	-0.11	-0.81	0.6	0.83	-0.44	2.1	0.36	-0.54	1.26			
Low-fat cheese												
Demographic, lifestyle, energy §	-0.13	-1	0.73	-1.26	-2.84	0.31	-1.06	-2.23	0.1	-0.04	-0.1	0.03
+ Diet ‖	-0.09	-0.96	0.79	-1.19	-2.78	0.39	-0.98	-2.18	0.22	-0.03	-0.1	0.03
+ BMI #	0.02	-0.86	0.9	-0.63	-2.17	0.9	-0.58	-1.77	0.61			
Ice-cream												
Demographic, lifestyle, energy §	0.12	-0.58	0.82	-0.26	-1.41	0.89	-0.36	-1.21	0.5	-0.002	-0.05	0.05
+ Diet ‖	0.14	-0.55	0.84	-0.35	-1.5	0.81	-0.35	-1.22	0.52	-0.003	-0.05	0.05
+ BMI #	0.14	-0.55	0.84	-0.48	-1.62	0.65	-0.43	-1.27	0.41			
Fermented dairy products												
Demographic, lifestyle, energy §	-0.03	-0.38	0.32	-0.27	-0.95	0.41	-0.37	-0.83	0.08	-0.03	-0.05	-0.004
+ Diet ‖	0	-0.36	0.35	-0.24	-0.94	0.47	-0.33	-0.79	0.14	-0.02	-0.04	0.001
+ BMI #	0.04	-0.32	0.39	-0.06	-0.74	0.63	-0.2	-0.64	0.24			
High-fat dairy products												
Demographic, lifestyle, energy §	0.06	-0.15	0.26	0.22	-0.11	0.55	0.18	-0.03	0.39	-0.002	-0.01	0.01
+ Diet ‖	0.06	-0.15	0.27	0.19	-0.15	0.54	0.16	-0.06	0.38	-0.001	-0.01	0.01
+ BMI #	0.04	-0.17	0.25	0.11	-0.23	0.46	0.11	-0.12	0.34			
Low-fat dairy products												
Demographic, lifestyle, energy §	0.01	-0.17	0.19	-0.11	-0.43	0.21	-0.04	-0.27	0.2	-0.01	-0.02	0.01
+ Diet ‖	0.04	-0.14	0.23	-0.09	-0.42	0.24	-0.01	-0.24	0.23	-0.004	-0.02	0.01
+ BMI #	0.06	-0.12	0.25	-0.01	-0.33	0.31	0.05	-0.18	0.28			

Table A.14 (continued)

	HbA1c (mmol/mol) †			**Systolic blood pressure (mmHg)**			**Diastolic blood pressure (mmHg)**			**Metabolic risk z-score**		
Mean (SD) of change	1.6 (6.1)			0.5 (14.9)			-0.1 (10.5)			0.01 (0.34)		
Participants (N)	6,224			14,210			14,231			6,033		
Dairy consumption (servings/d) ‡	b	95% CI		b	95% CI		b	95% CI		b	95% CI	
Total dairy products												
Demographic, lifestyle, energy §	0.005	-0.14	0.15	-0.01	-0.23	0.22	0	-0.16	0.16	-0.002	-0.01	0.01
+ Diet ‖	0.03	-0.12	0.17	-0.01	-0.24	0.22	0.003	-0.16	0.17	0	-0.01	0.01
+ BMI #	0.03	-0.12	0.17	-0.001	-0.23	0.22	0.01	-0.15	0.18			

† HbA1c: Haemoglobin A1c

‡ Servings as defined by Food Standards Agency 2002[197]: Milk- 1 average glass (200g); Yoghurt- 125g carton; Cheese- medium serving (40g); Single cream- 1 tablespoon (15g); Double cream- 1 tablespoon (30g); Butter- 1 teaspoon (10g); Ice-cream- 1 average scoop/tub (60g)

§ Linear regression model 1: age (years), sex, educational level (low, medium, high), age at completion of full-time education (years), marital status (single, married, widowed or separated), socio-economic status based on occupation (low: technical/semi-routine and routine occupationsl medium: lower managerial / intermediate occupationsl; high: professional / higher managerial occupations), individual follow-up time (years), physical activity level (inactive, moderately inactive, moderately active, active), smoking status (never, former and current smoker), lipid-lowering medication (Yes, No), anti-hypertensive medication (Yes, No), hormone-replacement therapy (Yes, No, Men), total energy intake (kcal/day)

‖ Linear regression model 2: Model 1 + intakes (g/d) of fruit, vegetables, potatoes, legumes, nuts, processed cereals, whole-grain cereals, poultry and eggs, red meat, processed meat, fish, sauces, margarine, sweet snacks, sugar-sweetened beverages, artificially sweetened beverages, fruit juice, coffee, tea and alcoholic beverages and dietary supplement use (Yes, No)

Linear regression model 3: Model 2 + BMI (kg/m^2)

Table A.15 Results from the complete-case analyses for the associations of the change in total and types of dairy consumption with the change in markers of body weight and composition from baseline to the first follow-up after a mean of 3.7 years in the EPIC-Norfolk study †

	Weight (kg)			**BMI (kg/m^2)**			**Waist (cm)**			**Waist / Hip circumference**		
Mean (SD) of change	1.3 (4)			0.6 (1.4)			0.8 (5.5)			0 (0.05)		
Participants range (N) †	5,540 - 5,699			5,536 - 5,698			5,564 - 5,724			5,561 - 5,718		
Dairy consumption (servings/d) §	b	95% CI		b	95% CI		b	95% CI		b	95% CI	
Milk	0.02	-0.14	0.19	-0.002	-0.06	0.06	0.03	-0.17	0.22	0	-0.002	0.002
Yoghurt	-0.47	-0.77	-0.17	-0.19	-0.3	-0.08	-0.28	-0.65	0.08	-0.003	-0.01	0.001
Cheese	-0.3	-0.64	0.03	-0.14	-0.27	-0.02	-0.16	-0.56	0.23	-0.001	-0.004	0.003
Butter	0.02	-0.13	0.18	0.002	-0.06	0.06	0.2	0.01	0.39	0.002	0	0.004
Full-fat milk	0.25	-0.06	0.56	0.07	-0.04	0.19	0.25	-0.12	0.62	0.002	-0.002	0.01
Low-fat milk	0.04	-0.12	0.2	0.001	-0.06	0.06	-0.04	-0.23	0.15	-0.001	-0.003	0.001
Full-fat yoghurt	-0.12	-1.14	0.89	-0.1	-0.47	0.28	-0.2	-1.41	1.01	-0.002	-0.01	0.01
Low-fat yoghurt	-0.54	-0.86	-0.23	-0.22	-0.33	-0.1	-0.28	-0.66	0.09	-0.003	-0.01	0.001
High-fat cheese	0.5	0.04	0.96	0.19	0.02	0.36	-0.05	-0.61	0.5	0	-0.01	0.01
Low-fat cheese	-0.97	-1.45	-0.49	-0.38	-0.55	-0.2	-0.37	-0.94	0.2	-0.002	-0.01	0.003
Ice-cream	-0.1	-0.54	0.33	-0.09	-0.25	0.07	-0.21	-0.72	0.31	-0.004	-0.01	0.001
Fermented dairy products	-0.34	-0.56	-0.13	-0.14	-0.22	-0.06	-0.15	-0.41	0.1	-0.001	-0.004	0.001
High-fat dairy products	0.13	0.002	0.27	0.04	-0.01	0.09	0.07	-0.08	0.23	0.001	-0.001	0.002
Low-fat dairy products	-0.09	-0.22	0.04	-0.04	-0.09	0.004	-0.12	-0.27	0.04	-0.001	-0.003	0
Total dairy products	-0.04	-0.15	0.07	-0.03	-0.07	0.01	0.04	-0.09	0.17	0	-0.001	0.001

† Associations from the maximally adjusted linear regression models are presented, which include: age (years), sex, educational level (low, medium, high), age at completion of full-time education (years), marital status (single, married, widowed or separated), socio-economic status based on occupation (low: technical/semi-routine and routine occupationsl medium: lower managerial / intermediate occupationsl; high: professional / higher managerial occupations), individual follow-up time (years), physical activity level (inactive, moderately inactive, moderately active, active), smoking status (never, former and current smoker), lipid-lowering medication (Yes, No), anti-hypertensive medication (Yes, No), hormone-replacement therapy (Yes, No, Men), total energy intake (kcal/day), intakes (g/d) of fruit, vegetables, potatoes, legumes, nuts, processed cereals, whole-grain cereals, poultry and eggs, red meat, processed meat, fish, sauces, margarine, sweet snacks, sugar-sweetened beverages, artificially sweetened beverages, fruit juice, coffee, tea and alcoholic beverages, dietary supplement use (Yes, No) and BMI (kg/m^2; in associations other than those for weight and BMI). When repeated measures of the covariates were available at baseline and first follow-up, their change was also included.

† A sample size range is given instead of a single sample size due to the exclusion of participants outside the range of mean±3SD of the change in dairy consumption

§ Servings as defined by Food Standards Agency 2002[197]: Milk- 1 average glass (200g); Yoghurt- 125g carton; Cheese- medium serving (40g); Single cream- 1 tablespoon (15g); Double cream- 1 tablespoon (30g); Butter- 1 teaspoon (10g); Ice-cream- 1 average scoop/tub (60g)

BMI: Body mass index

Table A.16 Results from the complete-case analyses for the associations of the change in total and types of dairy consumption with the change in lipid markers from baseline to the first follow-up after a mean of 3.7 years in the EPIC-Norfolk study †

	Total / HDL-C			**Total cholesterol (mmol/l)**			**HDL-C (mmol/l)**			**LDL-C (mmol/l)**			**Triglycerides (mmol/l)**		
Mean (SD) of change	-0.3 (1.2)			-0.1 (1)			0.1 (0.3)			-0.2 (0.9)			0.1 (0.9)		
Participants range (N) ‡	4,910 - 5,047			5,033 - 5,177			4,918 - 5,058			4,906 - 5,044			5,032 - 5,180		
Dairy consumption (servings/d) §	b	95% CI		b	95% CI		b	95% CI		b	95% CI		b	95% CI	
Milk	-0.03	-0.08	0.01	-0.05	-0.09	-0.01	0.002	-0.01	0.02	-0.05	-0.09	-0.01	0.01	-0.03	0.04
Yoghurt	0.05	-0.04	0.13	-0.01	-0.09	0.06	-0.02	-0.05	0.002	0.02	-0.05	0.09	-0.01	-0.08	0.06
Cheese	0.01	-0.09	0.1	0.05	-0.04	0.13	0.01	-0.02	0.04	0.02	-0.05	0.1	0.05	-0.03	0.12
Butter	-0.002	-0.06	0.06	0.001	-0.05	0.05	-0.001	-0.02	0.02	0.01	-0.04	0.06	0	-0.05	0.05
Full-fat milk	-0.01	-0.1	0.08	-0.03	-0.11	0.04	0.01	-0.02	0.03	0	-0.07	0.07	-0.07	-0.14	-0.001
Low-fat milk	-0.04	-0.08	0.01	-0.04	-0.08	0	0.002	-0.01	0.02	-0.05	-0.09	-0.01	0.01	-0.02	0.05
Full-fat yoghurt	0.33	0.04	0.63	0.2	-0.05	0.46	-0.04	-0.12	0.05	0.21	-0.03	0.45	0.03	-0.19	0.26
Low-fat yoghurt	0.01	-0.08	0.1	-0.03	-0.11	0.04	-0.02	-0.04	0.01	-0.001	-0.07	0.07	-0.03	-0.1	0.05
High-fat cheese	0.003	-0.13	0.14	0.15	0.04	0.27	0.03	-0.01	0.07	0.11	0.01	0.22	-0.01	-0.12	0.09
Low-fat cheese	0.01	-0.12	0.15	0	-0.12	0.12	-0.003	-0.04	0.04	-0.02	-0.13	0.09	0.12	0.02	0.23
Ice-cream	0.03	-0.01	0.08	0.05	0.01	0.09	0.001	-0.01	0.01	0.04	0	0.07	0.02	-0.02	0.05
Fermented dairy products	0.06	-0.06	0.19	0.01	-0.09	0.12	-0.02	-0.06	0.02	-0.06	-0.16	0.04	0.13	0.03	0.23
High-fat dairy products	0.02	-0.02	0.06	0.04	0.003	0.07	0.01	-0.01	0.02	0.03	-0.001	0.06	0	-0.03	0.03
Low-fat dairy products	-0.03	-0.07	0.01	-0.03	-0.06	0.001	0	-0.01	0.01	-0.04	-0.07	-0.01	0.002	-0.03	0.03
Total dairy products	0.01	-0.03	0.04	0.01	-0.02	0.04	0.001	-0.01	0.01	-0.003	-0.03	0.02	0.02	-0.003	0.05

† Associations from the maximally adjusted linear regression models are presented, which include: age (years), sex, educational level (low, medium, high), age at completion of full-time education (years), marital status (single, married, widowed or separated), socio-economic status based on occupation (low: technical/semi-routine and routine occupationsl medium: lower managerial / intermediate occupationsl; high: professional / higher managerial occupations), individual follow-up time (years), physical activity level (inactive, moderately inactive, moderately active, active), smoking status (never, former and current smoker), lipid-lowering medication (Yes, No), anti-hypertensive medication (Yes, No), hormone-replacement therapy (Yes, No, Men), total energy intake (kcal/day), intakes (g/d) of fruit, vegetables, potatoes, legumes, nuts, processed cereals, whole-grain cereals, poultry and eggs, red meat, processed meat, fish, sauces, margarine, sweet snacks, sugar-sweetened beverages, artificially sweetened beverages, fruit juice, coffee, tea and alcoholic beverages, dietary supplement use (Yes, No) and BMI (kg/m^2). When repeated measures of the covariates were available at baseline and first follow-up, their change was also included.

‡ A sample size range is given instead of a single sample size due to the exclusion of participants outside the range of mean±3SD of the change in dairy consumption

§ Servings as defined by Food Standards Agency 2002[197]: Milk- 1 average glass (200g); Yoghurt- 125g carton; Cheese- medium serving (40g); Single cream- 1 tablespoon (15g); Double cream- 1 tablespoon (30g); Butter- 1 teaspoon (10g); Ice-cream- 1 average scoop/tub (60g)

HDL-C: High density lipoprotein cholesterol; LDL-C: Low density lipoprotein cholesterol

Table A.17 Results from the complete-case analyses for the associations of the change in total and types of dairy consumption with the change in HbA1c, blood pressure and a z-score for metabolic risk from baseline to the first follow-up after a mean of 3.7 years in the EPIC-Norfolk study †

	HbA1c (mmol/mol)			**Systolic blood pressure (mmHg)**			**Diastolic blood pressure (mmHg)**			**Metabolic risk z-score**		
Mean (SD) of change	1.3 (6.3)			0.5 (14.9)			-0.1 (10.5)			-0.03 (0.4)		
Participants range (N) ‡	3,292 - 3,378			5,547 - 5,707			5,556 - 5,714			3,135 - 3,216		
Dairy consumption (servings/d) §	b	95% CI		b	95% CI		b	95% CI		b	95% CI	
Milk	0.12	-0.16	0.39	-0.39	-1.01	0.23	0.01	-0.41	0.44	0.001	-0.02	0.02
Yoghurt	0.11	-0.39	0.6	-0.3	-1.45	0.84	-0.27	-1.06	0.52	-0.01	-0.06	0.03
Cheese	0.32	-0.21	0.84	0.77	-0.48	2.01	0.73	-0.12	1.58	-0.01	-0.06	0.03
Butter	0.08	-0.18	0.33	-0.07	-0.66	0.52	-0.02	-0.43	0.39	-0.01	-0.03	0.02
Full-fat milk	0.49	-0.03	1.02	-0.63	-1.79	0.53	-0.28	-1.07	0.52	-0.002	-0.05	0.04
Low-fat milk	0.14	-0.12	0.41	-0.43	-1.02	0.17	-0.02	-0.42	0.39	-0.001	-0.02	0.02
Full-fat yoghurt	-0.51	-2.08	1.07	2.52	-1.32	6.37	0.99	-1.65	3.63	0.05	-0.09	0.18
Low-fat yoghurt	0.13	-0.39	0.65	-0.26	-1.45	0.93	-0.19	-1	0.63	-0.03	-0.07	0.02
High-fat cheese	0.3	-0.45	1.05	1.22	-0.54	2.98	1.05	-0.15	2.26	-0.01	-0.07	0.06
Low-fat cheese	0.26	-0.51	1.03	0.81	-0.98	2.61	0.97	-0.26	2.21	0.02	-0.04	0.09
Ice-cream	-0.16	-0.87	0.55	-0.77	-2.41	0.86	0.38	-0.74	1.5	0.003	-0.06	0.06
Fermented dairy products	0.17	-0.17	0.51	-0.26	-1.07	0.54	-0.09	-0.65	0.46	-0.03	-0.06	0.002
High-fat dairy products	0.15	-0.06	0.37	-0.03	-0.53	0.47	0.04	-0.3	0.38	-0.01	-0.02	0.01
Low-fat dairy products	0.17	-0.05	0.38	-0.38	-0.87	0.11	-0.04	-0.37	0.3	-0.01	-0.03	0.01
Total dairy products	0.13	-0.05	0.3	-0.17	-0.57	0.24	0.04	-0.24	0.32	-0.002	-0.02	0.01

† Associations from the maximally adjusted linear regression models are presented, which include: age (years), sex, educational level (low, medium, high), age at completion of full-time education (years), marital status (single, married, widowed or separated), socio-economic status based on occupation (low: technical/semi-routine and routine occupationsl medium: lower managerial / intermediate occupationsl; high: professional / higher managerial occupations), individual follow-up time (years), physical activity level (inactive, moderately inactive, moderately active, active), smoking status (never, former and current smoker), lipid-lowering medication (Yes, No), anti-hypertensive medication (Yes, No), hormone-replacement therapy (Yes, No, Men), total energy intake (kcal/day), intakes (g/d) of fruit, vegetables, potatoes, legumes, nuts, processed cereals, whole-grain cereals, poultry and eggs, red meat, processed meat, fish, sauces, margarine, sweet snacks, sugar-sweetened beverages, artificially sweetened beverages, fruit juice, coffee, tea and alcoholic beverages, dietary supplement use (Yes, No) and BMI (kg/m^2; in associations other than those for metabolic risk z-score). When repeated measures of the covariates were available at baseline and first follow-up, their change was also included.

‡ A sample size range is given instead of a single sample size due to the exclusion of participants outside the range of mean±3SD of the change in dairy consumption

§ Servings as defined by Food Standards Agency 2002[197]: Milk- 1 average glass (200g); Yoghurt- 125g carton; Cheese- medium serving (40g); Single cream- 1 tablespoon (15g); Double cream- 1 tablespoon (30g); Butter- 1 teaspoon (10g); Ice-cream- 1 average scoop/tub (60g)

HbA1c: Haemoglobin A1c

A.4 Chapter 6

Table A.18 Metabolite concentrations in the discovery and internal validation sets of the Fenland study and the external validation set of the EPIC Norfolk diabetes case-cohort study

Metabolite class	Metabolite abbreviation	Matched metabolite name ‡	Discovery set*		Internal validation set*		External validation set †	
			Mean	SD	Mean	SD	Mean	SD
Amino acid	Ala	alanine	336.1	78.8	335.2	77.6	1	0.1
Amino acid	Arg	arginine	83.2	20.8	83.5	20.6	1	0.2
Amino acid	Asn	asparagine	44.9	10.4	44.9	10.4	1	0.4
Amino acid	Asp	aspartate	3.5	3.1	3.4	2.8	1.1	0.3
Amino acid	Cit	citrulline	32.6	9.1	32.7	9	1	0.2
Amino acid	Gln	glutamine	610.3	108	610.4	103.3	1	0.3
Amino acid	Glu	glutamate	48.5	31.7	47.1	31.8	1	0.2
Amino acid	Gly	glycine	250.3	79.3	252.2	80.8	1	0.3
Amino acid	His	histidine	83.1	12.9	83	13.1	1	0.2
Amino acid	Ile	isoleucine	66.8	16.6	66.6	15.9	1	0.2
Amino acid	Leu	leucine	128.5	34.8	128.5	33.4	1	0.3
Amino acid	Lys	lysine	213.3	38.1	213.3	37.9	1	0.2
Amino acid	Met	methionine	21.9	4.2	21.9	4.1	1	0.3
Amino acid	Orn	ornithine	60.5	15.7	60.4	15.6	1	0.2
Amino acid	Phe	phenylalanine	59.1	9.1	59.3	9	1	0.2
Amino acid	Pro	proline	162.7	50.6	162.3	49.5	1.1	0.3
Amino acid	Ser	serine	96.8	24.8	97.1	24.8	1	0.2
Amino acid	Thr	threonine	158.4	885	147.7	102	1.1	0.7
Amino acid	Trp	tryptophan	59.2	10.5	59.2	10.4	1	0.3
Amino acid	Tyr	tyrosine	62.1	13.9	62.4	13.7	1	0.2
Amino acid	Val	valine	194.1	35.3	193.8	34.6	1	0.2
Biogenic amines	AcOrn	N-delta-acetylornithine	0.6	0.6	0.6	0.6	1	0.2
Biogenic amines	SDMA	dimethylarginine (SDMA + ADMA)	1.1	0.3	1.1	0.3	1	0.2
Biogenic amines	alpha-AAA	2-aminoadipate	0.5	0.3	0.6	0.3	1	0.2
Biogenic amines	Creatinine	creatinine §	68.1	15	67.8	14.5	1.4	1.7
Biogenic amines	Kynurenine	kynurenine ‖	2	0.8	2	0.7	1.1	0.7
Biogenic amines	Met-SO	methionine sulfoxide	0.4	0.2	0.4	0.2	1	0.3
Biogenic amines	t4OHPro	trans-4-hydroxyproline	9.2	4.8	9.2	4.8	1	0.2
Biogenic amines	Sarcosine	sarcosine (N-Methylglycine)	2.3	3.3	2.2	2.9	1.1	0.4
Biogenic amines	Serotonin	serotonin	0.2	0.2	0.2	0.2	1	0.2
Biogenic amines	Taurine	taurine	55.1	20.2	55.2	20	1.1	0.5
Acylcarnitines	C0	carnitine	34.9	8.3	34.9	8.3	1	0.3
Acylcarnitines	C2	acetylcarnitine	6.7	2.3	6.7	2.2	1.1	0.3
Acylcarnitines	C3	propionylcarnitine	0.4	0.1	0.4	0.1	1.1	0.4

Table A.18 (continued)

Metabolite class	Metabolite abbreviation	Matched metabolite name ‡	Discovery set*		Internal validation set*		External validation set †	
			Mean	SD	Mean	SD	Mean	SD
Acylcarnitines	C4	butyrylcarnitine	0.2	0.1	0.2	0.1	1.2	0.8
Acylcarnitines	C3DCC4OH	hydroxybutyrylcarnitine	0.1	0	0.1	0	1.2	0.8
Acylcarnitines	C51	tiglyl carnitine	0.1	0	0.1	0	1.1	0.4
Acylcarnitines	C5MDC	3-methylglutarylcarnitine (2)	0.1	0	0.1	0	1	0.2
Acylcarnitines	C6C41DC	hexanoylcarnitine	0.1	0	0.1	0	1	0.3
Acylcarnitines	C8	octanoylcarnitine	0.3	0.2	0.3	0.2	1	0.3
Acylcarnitines	C10	decanoylcarnitine	0.5	0.3	0.5	0.3	1	0.3
Acylcarnitines	C12	laurylcarnitine	0.2	0.1	0.2	0.1	1.1	0.5
Acylcarnitines	C14	myristoylcarnitine ‖	0.1	0	0.1	0.1	1	0.4
Acylcarnitines	C141	myristoleoylcarnitine	0.2	0.1	0.2	0.1	1	0.2
Acylcarnitines	C16	palmitoylcarnitine	0.1	0	0.1	0	1	0.2
Acylcarnitines	C18	stearoylcarnitine	0.1	0	0.1	0	1	0.3
Acylcarnitines	C182	linoleoylcarnitine	0.1	0	0.1	0	1	0.2
Lysophosphatidylcholines	LPC a C14:0	1-myristoyl-GPC (14:0)	3.5	0.6	3.4	0.6	1	0.3
Lysophosphatidylcholines	LPC a C16:1	1-palmitoleoyl-GPC (16:1)	3	1.1	2.9	1	1	0.2
Lysophosphatidylcholines	LPC a C17:0	1-margaroyl-GPC (17:0)	1.5	0.5	1.4	0.5	1.2	0.9
Lysophosphatidylcholines	LPC a C18:0	1-stearoyl-GPC (18:0) §	21.8	6.3	21.6	6.1	1.2	0.8
Lysophosphatidylcholines	LPC a C18:2	1-linoleoyl-GPC (18:2)	26.7	9.5	26.5	9.2	1	0.3
Lysophosphatidylcholines	LPC a C20:4	1-arachidonoyl-GPC (20:4)	5.4	1.8	5.3	1.7	1.2	0.8
Lysophosphatidylcholines	LPC a C16:0	1-palmitoyl-GPC (16:0)	79.9	20.5	78.9	19.5	1.1	0.4
Lysophosphatidylcholines	LPC a C18:1	1-oleoyl-GPC (18:1)	17.8	5.6	17.6	5.4	1.1	0.5
Lysophosphatidylcholines	LPC a C20:3	1-eicosatrienoyl-GPC (20:3) #	1.9	0.7	1.8	0.6	1.3	2.9
Phosphatidylcholines	PC aa C30:0	1-myristoyl-2-palmitoyl-GPC (14:0/16:0)	2.5	0.9	2.5	0.9	1	0.2
Phosphatidylcholines	PC aa C32:0	1,2-dipalmitoyl-GPC (16:0/16:0)	309.6	80.3	311.4	80.1	1.1	0.6
Phosphatidylcholines	PC aa C32:1	1-palmitoyl-2-palmitoleoyl-GPC (16:0/16:1)	531.5	294.4	530.8	288.1	1.2	0.7
Phosphatidylcholines	PC aa C32:2	1,2-dipalmitoleoyl-GPC (16:1/16:1)	172	67.3	173.9	67.3	1.2	0.7
Phosphatidylcholines	PC aa C34:1	1-palmitoyl-2-oleoyl-GPC (16:0/18:1)	3,305.50	1,060.70	3,323.60	1,040.40	1.2	0.8
Phosphatidylcholines	PC aa C34:2	1-palmitoyl-2-linoleoyl-GPC (16:0/18:2)	9,299.60	2,408.60	9,384.50	2,390.10	1	0.3
Phosphatidylcholines	PC aa C36:1	1-stearoyl-2-oleoyl-GPC (18:0/18:1)	240.9	85.4	243.1	85.3	1	0.3
Phosphatidylcholines	PC aa C36:2	1-stearoyl-2-linoleoyl-GPC (18:0/18:2)	2,108.70	567.2	2,141.80	567.1	1	0.3
Phosphatidylcholines	PC aa C36:3	1-palmitoyl-2-dihomo-linolenoyl-GPC (16:0/20:3)	1,451.60	411.2	1,472.00	415.7	1.1	0.6
		1-oleoyl-2-linoleoyl-GPC (18:1/18:2)					1	0.3
Phosphatidylcholines	PC aa C36:4	1-palmitoyl-2-arachidonoyl-GPC (16:0/20:4)	2,309.20	708.1	2,322.00	693.3	1	0.3
Phosphatidylcholines	PC aa C36:5	1-palmitoyl-2-eicosapentaenoyl-GPC (16:0/20:5)	503.8	280.9	502.7	271.4	1.1	0.4
Phosphatidylcholines	PC aa C38:4	1-stearoyl-2-arachidonoyl-GPC (18:0/20:4)	475.2	151.3	480.6	151.2	1	0.2

Table A.18 (continued)

Metabolite class	Metabolite abbreviation	Matched metabolite name ‡	Discovery set*		Internal validation set*		External validation set †	
			Mean	SD	Mean	SD	Mean	SD
Phosphatidylcholines	PC aa C38:5	1-palmitoyl-2-docosapentaenoyl-GPC (16:0/22:5n6) ||	356.8	112.8	358.9	112	1.1	0.7
		1-palmitoyl-2-docosapentaenoyl-GPC (16:0/22:5n3)					1.1	0.4
Phosphatidylcholines	PC ae C34:1	1-pentadecanoyl-2-oleoyl-GPC (15:0/18:1) ||	166.2	45.6	167.7	45.4	1	0.3
Phosphatidylcholines	PC ae C34:2	1-pentadecanoyl-2-linoleoyl-GPC (15:0/18:2)	241	67.8	242.5	67.1	1.1	0.4
Phosphatidylcholines	PC ae C36:1	1-margaroyl-2-oleoyl-GPC (17:0/18:1)	88	28.9	89.4	28.7	1	0.3
Phosphatidylcholines	PC ae C36:2	1-margaroyl-2-linoleoyl-GPC (17:0/18:2)	160.9	48.7	162.9	48.6	1	0.3
Sphingomyelins	SM C16:0	palmitoyl sphingomyelin (d18:1/16:0)	29.5	6.9	29.5	6.6	1.2	0.7
Sphingomyelins	SM C16:1	sphingomyelin (d18:2/16:0, d18:1/16:1)	4.8	1.2	4.8	1.1	1.1	0.6
Sphingomyelins	SM C18:0	stearoyl sphingomyelin (d18:1/18:0)	4	1.2	4	1.2	1.1	0.4
Sphingomyelins	SM C18:1	sphingomyelin (d18:1/18:1, d18:2/18:0)	2.2	0.6	2.2	0.6	1	0.3
Sphingomyelins	SM C24:1	sphingomyelin (d18:1/24:1, d18:2/24:0)	2.9	1.1	2.9	1	1.1	0.6
Sphingomyelins	SM-OH C14:1	sphingomyelin (d18:1/15:0, d16:1/17:0)	2	0.6	2	0.6	1	0.4
Sphingomyelins	SM-OH C22:1	tricosanoyl sphingomyelin (d18:1/23:0)	0.9	0.3	0.9	0.3	1.1	0.5
Sphingomyelins	SM -OH C22:2	sphingomyelin (d18:2/23:0, d18:1/23:1, d17:1/24:1)	0.8	0.5	0.7	0.4	1	0.4
Hexose	Hexose	mannose	5,557.80	891.1	5,547.30	861.5	1.1	0.4
		fructose					1.1	0.4
		glucose					1	0.3

*Mean and SD presented for metabolites after quality control and batch correction and before log-transformation, mean centring and standardisation

†Mean values are close to 1 due to normalisation by "block correction" i.e. setting median to 1 and adjusting each data point proportionately. Mean values presented after quality control and normalisation and before log-transformation, mean centring and standardisation.

‡Missing values for all the metabolites in the Fenland study and most of the metabolites in the EPIC Norfolk diabetes case-cohort sub-study were < 5% unless otherwise stated

§Missing values: 25-50%

||Missing values: 5-15%

#Missing values: 15-25%

Table A.19 Coefficients of predictors in decreasing order of magnitude of the absolute values* for total and types of dairy products as derived from elastic net logistic prediction models † in the discovery set of the Fenland study with all 174 metabolites included

	Milk		Yoghurt		Cheese		Butter		Total dairy products	
Order #	Predictor	Coefficient	Predictor	Coefficient	Predictor	Coefficient	Predictor	Coefficient	Predictor	Coefficient
1	SM-OH C14:1	0.463	c4OHPro	-0.836	SM-OH C14:1	4.648	SM-OH C14:1	0.292	SM-OH C14:1	1.128
2	SM C16:1	-0.291	SM-OH C14:1	0.806	PC ae C36:4	-4.288	PC aa C32:0	-0.238	SM C16:1	-0.486
3	PC aa C28:1	0.237	SM C24:0	-0.755	C0	-3.568	C9	0.236	LPC a C18:0	-0.383
4	C141	-0.208	SM C16:0	-0.747	PC aa C42:4	3.363	PC ae C34:1	0.218	LPC a C17:0	0.312
5	Ile	-0.203	PC aa C28:1	0.715	Tyr	3.335	LPC a C17:0	0.193	SM C24:1	-0.241
6	LPC a C18:0	-0.192	Taurine	0.709	LPC a C17:0	3.257	LPC a C14:0	0.189	SM C16:0	-0.213
7	LPC a C20:3	0.184	PC ae C40:4	0.624	C51	3.213	SM-OH C16:1	0.157	PC aa C28:1	0.182
8	Thr	-0.145	C9	-0.602	PC ae C36:2	2.873	PC ae C34:0	0.147	LPC a C14:0	0.143
9	LPC a C17:0	0.133	PC aa C36:1	-0.593	LPC a C18:0	-2.78	C102	-0.13	SM C24:0	-0.134
10	C142	-0.126	C141	-0.561	C5MDC	-2.423	SM C16:1	-0.124	LPC a C24:0	-0.109
11	PC aa C40:5	0.125	LPC a C24:0	0.55	Glu	2.239	C182	-0.113	PC aa C40:3	-0.089
12	PC aa C38:3	0.122	C161	0.548	PC aa C38:0	-2.232	LPC a C16:0	-0.11	PC ae C30:0	0.088
13	PC aa C38:1	-0.119	Asp	-0.527	Arg	2.21	PC ae C30:0	0.107	PC ae C34:0	0.087
14	PC ae C38:5	-0.111	C18	-0.509	PC ae C30:0	2.171	LPC a C18:0	-0.092	C3DCMC5OH	-0.075
15	Creatinine	0.109	PC aa C36:2	-0.507	SM C18:1	-2.151	PC ae C44:6	-0.083	PC aa C32:1	0.071
16	SM C24:0	-0.109	C102	0.492	PC aa C28:1	2.086	C141	0.077	Orn	-0.067
17	t4OHPro	-0.108	Ile	-0.482	PC aa C36:1	-2.063	PC ae C36:1	0.076	Arg	0.066
18	SM C16:0	-0.106	C2	0.458	C142	-2.048	PC ae C42:1	-0.074	t4OHPro	-0.066
19	Cit	0.105	PC ae C42:4	0.457	Creatinine	-1.918	PC ae C36:2	0.073	PC aa C40:5	0.051
20	C18	-0.102	C16	-0.454	C162	1.795	SM C24:1	-0.072	c4OHPro	-0.049
21	Orn	-0.1	PC ae C44:6	0.44	c4OHPro	-1.732	LPC a C20:3	-0.07	PC ae C34:3	-0.047
22	Pro	0.099	PC aa C36:0	0.411	Gln	1.723	Taurine	-0.069	PC ae C38:0	-0.045
23	LPC a C18:1	-0.098	PC aa C30:0	0.408	C10	1.626	PC ae C38:4	0.064	PC aa C32:3	-0.04
24	PC aa C36:4	-0.097	LPC a C17:0	0.407	PC aa C32:3	1.616	LPC a C26:0	-0.059	PC aa C36:2	-0.01
25	C161	0.094	His	0.403	C101	-1.572	Glu	0.058	Ala	0
26	PC ae C30:2	-0.09	C182	0.389	PC ae C40:2	-1.425	C142OH	-0.058	Asn	0
27	SM-OH C22:1	-0.088	C5	-0.388	C142OH	-1.172	PC aa C42:0	-0.056	Asp	0
28	His	-0.085	LPC a C16:1	0.376	SM-OH C16:1	1.156	Lys	-0.052	Cit	0
29	PC aa C36:1	-0.084	PC ae C30:1	-0.362	Thr	1.112	PC aa C36:1	0.05	Gln	0
30	C3	0.082	LPC a C18:0	-0.358	C102	-1.112	C3DCC4OH	-0.049	Glu	0
31	PC ae C30:1	-0.082	SM C24:1	-0.354	PC ae C34:0	1.029	LPC a C24:0	-0.049	Gly	0
32	Taurine	0.08	C14	0.353	SM C24:1	-0.982	C31	-0.048	His	0
33	SM C18:1	-0.08	PC ae C40:1	0.343	PC ae C44:3	-0.909	PC ae C40:1	-0.048	Ile	0
34	PC aa C34:4	-0.077	Met	0.339	SDMA	0.898	PC aa C42:5	-0.045	Leu	0

Table A.19 (continued)

Order #	Milk Predictor	Coefficient	Yoghurt Predictor	Coefficient	Cheese Predictor	Coefficient	Butter Predictor	Coefficient	Total dairy products Predictor	Coefficient
35	C51	0.07	Glu	-0.335	PC ae C34:3	-0.862	PC aa C38:1	-0.044	Lys	0
36	PC ae C42:5	0.069	PEA	-0.332	SM C18:0	-0.845	PC aa C40:6	-0.043	Met	0
37	alpha-AAA	0.068	Kynurenine	0.331	SM C16:1	-0.841	Trp	-0.041	Phe	0
38	Serotonin	-0.066	PC aa C36:3	0.327	C16:1OH	0.84	Val	0.039	Pro	0
39	C14:1OH	0.064	Gln	-0.319	C18:1OH	0.738	PC ae C38:0	-0.039	Ser	0
40	Ser	-0.061	PC ae C44:5	-0.313	Spermidine	-0.716	Arg	0.036	Thr	0
41	C0	0.06	PC ae C44:3	-0.312	t4OHPro	-0.693	Orn	0.036	Trp	0
42	PC aa C32:3	-0.06	t4OHPro	-0.311	PC ae C42:1	-0.689	Asn	-0.034	Tyr	0
43	Leu	0.059	Sarcosine	-0.309	C14:1	-0.623	PC ae C42:4	-0.034	Val	0
44	PC aa C40:4	0.056	LPC a C14:0	0.306	PC aa C42:5	0.588	PC aa C40:2	-0.025	PEA	0
45	C12:1	0.055	PC aa C40:1	-0.294	C12:1	-0.556	C12DC	-0.024	AcOrn	0
46	Gln	0.054	PC ae C36:5	-0.29	C3DCC4OH	-0.53	PC ae C42:2	-0.024	SDMA	0
47	C6C41DC	-0.053	PC ae C40:2	-0.278	PC aa C32:1	0.52	C18:1	0.02	alpha-AAA	0
48	SM C18:0	0.052	C4	0.272	SM C24:0	-0.483	Gly	-0.019	Creatinine	0
49	C4:1	-0.051	C5:1	0.271	PC ae C36:5	-0.459	Serotonin	-0.019	Kynurenine	0
50	Hexose	0.051	Lys	0.267	PC ae C38:2	0.456	C5:1	-0.017	Met-SO	0
51	PC aa C42:0	0.05	LPC a C28:0	-0.252	C3	-0.386	C16:2OH	-0.017	Sarcosine	0
52	C18:2	-0.046	PC ae C36:2	0.242	LPC a C16:1	0.308	C18	0.017	Serotonin	0
53	PC aa C42:4	0.044	PC aa C40:3	0.24	PC aa C36:0	-0.28	SM C24:0	-0.016	Spermidine	0
54	Lys	0.042	PC ae C42:5	-0.237	PC ae C36:0	-0.268	Ile	0.015	Taurine	0
55	LPC a C16:1	0.042	PC ae C36:4	-0.23	PC ae C30:2	0.265	t4OHPro	0.015	C0	0
56	SM -OH C22:2	0.042	PC ae C42:3	-0.23	PC ae C38:5	-0.239	PC aa C40:3	-0.015	C2	0
57	PC ae C44:6	0.041	PC ae C34:3	-0.227	PC ae C38:6	-0.195	C16	0.014	C3	0
58	C18:1OH	0.04	Val	0.221	C6:1	-0.122	PC aa C38:6	-0.012	C3:1	0
59	PC ae C38:4	0.04	Ser	-0.205	PC aa C42:6	0.082	PC ae C36:0	0.012	C3OH	0
60	PC ae C42:4	-0.039	PC aa C38:4	-0.202	C3DCMC5OH	0.079	Spermidine	-0.011	C4	0
61	PC aa C32:0	0.038	PC ae C40:3	-0.183	LPC a C28:1	0.056	C3DCMC5OH	-0.011	C4:1	0
62	PC ae C44:3	-0.038	C16OH	-0.182	Ala	0	PC aa C40:1	-0.005	C3DCC4OH	0
63	PC aa C24:0	-0.037	C6:1	-0.179	Asn	0	Ala	0	C5	0
64	PC ae C42:2	0.036	Asn	0.172	Asp	0	Asp	0	C5:1	0
65	PC aa C36:0	0.035	LPC a C20:3	0.167	Cit	0	Cit	0	C5:1DC	0
66	PC ae C36:0	-0.034	Serotonin	0.158	Gly	0	Gln	0	C5DCC6OH	0
67	Val	0.033	PC ae C42:1	0.157	His	0	His	0	C5MDC	0
68	PC ae C34:1	0.032	Cit	-0.155	Ile	0	Leu	0	C6C4:1DC	0

Table A.19 (continued)

Order #	Milk Predictor	Coefficient	Yoghurt Predictor	Coefficient	Cheese Predictor	Coefficient	Butter Predictor	Coefficient	Total dairy products Predictor	Coefficient
69	PC aa C40:3	-0.031	PC ae C36:3	-0.155	Leu	0	Met	0	C61	0
70	Arg	0.029	Ala	-0.149	Lys	0	Phe	0	C7DC	0
71	PC ae C36:4	-0.029	PC ae C40:6	0.141	Met	0	Pro	0	C8	0
72	Spermidine	0.028	PC aa C34:2	-0.133	Orn	0	Ser	0	C9	0
73	PC aa C40:2	0.028	PC aa C40:5	0.133	Phe	0	Thr	0	C10	0
74	Tyr	0.027	AcOrn	0.131	Pro	0	Tyr	0	C101	0
75	SM C20:2	0.027	LPC a C26:1	0.122	Ser	0	PEA	0	C102	0
76	PC ae C40:1	0.025	PC ae C38:5	-0.122	Trp	0	AcOrn	0	C12	0
77	Asp	0.024	LPC a C20:4	-0.114	Val	0	SDMA	0	C12DC	0
78	SDMA	-0.024	PC ae C30:0	0.111	PEA	0	alpha-AAA	0	C121	0
79	C3DCC4OH	0.024	PC aa C34:1	-0.106	AcOrn	0	Creatinine	0	C14	0
80	SM-OH C16:1	0.023	LPC a C28:1	0.097	alpha-AAA	0	Kynurenine	0	C141	0
81	PC aa C42:1	0.022	PC aa C42:2	-0.095	Kynurenine	0	Met-SO	0	C141OH	0
82	C8	-0.021	SM C18:1	0.076	Met-SO	0	c4OHPro	0	C142	0
83	PC ae C44:5	0.021	PC aa C34:3	0.062	Sarcosine	0	Sarcosine	0	C142OH	0
84	Ala	-0.02	C51DC	-0.058	Serotonin	0	C0	0	C16	0
85	PC ae C40:5	-0.02	PC aa C32:3	-0.058	Taurine	0	C2	0	C161	0
86	Trp	0.019	PC aa C42:5	-0.054	C2	0	C3	0	C161OH	0
87	C16OH	0.018	C101	-0.034	C31	0	C3OH	0	C162	0
88	LPC a C18:2	-0.018	Hexose	-0.032	C3OH	0	C4	0	C162OH	0
89	SM C24:1	-0.018	C3	0.031	C4	0	C41	0	C16OH	0
90	PC aa C38:5	0.013	C10	-0.015	C41	0	C5	0	C18	0
91	PC aa C42:6	0.012	PC aa C38:6	0.01	C5	0	C51DC	0	C181	0
92	PC aa C40:1	-0.011	PC ae C42:0	0.009	C51DC	0	C5DCC6OH	0	C181OH	0
93	PC ae C36:1	0.011	C142	-0.005	C5DCC6OH	0	C5MDC	0	C182	0
94	c4OHPro	-0.01	Arg	0	C6C41DC	0	C6C41DC	0	LPC a C16:1	0
95	LPC a C20:4	-0.01	Gly	0	C7DC	0	C61	0	LPC a C18:2	0
96	C161OH	0.009	Leu	0	C8	0	C7DC	0	LPC a C20:4	0
97	Gly	-0.008	Orn	0	C9	0	C8	0	LPC a C26:1	0
98	C16	0.008	Phe	0	C12	0	C10	0	LPC a C16:0	0
99	C61	0.007	Pro	0	C12DC	0	C101	0	LPC a C18:1	0
100	PC aa C38:6	-0.007	Thr	0	C14	0	C12	0	LPC a C20:3	0
101	PC ae C42:0	0.007	Trp	0	C141OH	0	C121	0	LPC a C26:0	0
102	Sarcosine	-0.006	Tyr	0	C16	0	C14	0	LPC a C28:0	0

Table A.19 (continued)

Order #	Milk Predictor	Coefficient	Yoghurt Predictor	Coefficient	Cheese Predictor	Coefficient	Butter Predictor	Coefficient	Total dairy products Predictor	Coefficient
103	LPC a C14:0	0.005	SDMA	0	C161	0	C141OH	0	LPC a C28:1	0
104	Asn	0	alpha-AAA	0	C162OH	0	C142	0	PC aa C24:0	0
105	Glu	0	Creatinine	0	C16OH	0	C161	0	PC aa C26:0	0
106	Met	0	Met-SO	0	C18	0	C161OH	0	PC aa C30:0	0
107	Phe	0	Spermidine	0	C181	0	C162	0	PC aa C32:0	0
108	PEA	0	C0	0	C182	0	C16OH	0	PC aa C32:2	0
109	AcOrn	0	C31	0	LPC a C14:0	0	C181OH	0	PC aa C34:1	0
110	Kynurenine	0	C3OH	0	LPC a C18:2	0	LPC a C16:1	0	PC aa C34:2	0
111	Met-SO	0	C41	0	LPC a C20:4	0	LPC a C18:2	0	PC aa C34:3	0
112	C2	0	C3DCC4OH	0	LPC a C26:1	0	LPC a C20:4	0	PC aa C34:4	0
113	C31	0	C5DCC6OH	0	LPC a C16:0	0	LPC a C26:1	0	PC aa C36:0	0
114	C3OH	0	C5MDC	0	LPC a C18:1	0	LPC a C18:1	0	PC aa C36:1	0
115	C4	0	C3DCMC5OH	0	LPC a C20:3	0	LPC a C28:0	0	PC aa C36:3	0
116	C5	0	C6C41DC	0	LPC a C24:0	0	LPC a C28:1	0	PC aa C36:4	0
117	C51DC	0	C7DC	0	LPC a C26:0	0	PC aa C24:0	0	PC aa C36:5	0
118	C5DCC6OH	0	C8	0	LPC a C28:0	0	PC aa C26:0	0	PC aa C36:6	0
119	C5MDC	0	C12	0	PC aa C24:0	0	PC aa C28:1	0	PC aa C38:0	0
120	C3DCMC5OH	0	C12DC	0	PC aa C26:0	0	PC aa C30:0	0	PC aa C38:1	0
121	C7DC	0	C121	0	PC aa C30:0	0	PC aa C32:1	0	PC aa C38:3	0
122	C9	0	C141OH	0	PC aa C32:0	0	PC aa C32:2	0	PC aa C38:4	0
123	C10	0	C142OH	0	PC aa C32:2	0	PC aa C32:3	0	PC aa C38:5	0
124	C101	0	C161OH	0	PC aa C34:1	0	PC aa C34:1	0	PC aa C38:6	0
125	C102	0	C162	0	PC aa C34:2	0	PC aa C34:2	0	PC aa C40:1	0
126	C12	0	C162OH	0	PC aa C34:3	0	PC aa C34:3	0	PC aa C40:2	0
127	C12DC	0	C181	0	PC aa C34:4	0	PC aa C34:4	0	PC aa C40:4	0
128	C14	0	C181OH	0	PC aa C36:2	0	PC aa C36:0	0	PC aa C40:6	0
129	C142OH	0	LPC a C18:2	0	PC aa C36:3	0	PC aa C36:2	0	PC aa C42:0	0
130	C162	0	LPC a C16:0	0	PC aa C36:4	0	PC aa C36:3	0	PC aa C42:1	0
131	C162OH	0	LPC a C18:1	0	PC aa C36:5	0	PC aa C36:4	0	PC aa C42:2	0
132	C181	0	LPC a C26:0	0	PC aa C36:6	0	PC aa C36:5	0	PC aa C42:4	0
133	LPC a C26:1	0	PC aa C24:0	0	PC aa C38:1	0	PC aa C36:6	0	PC aa C42:5	0
134	LPC a C16:0	0	PC aa C26:0	0	PC aa C38:3	0	PC aa C38:0	0	PC aa C42:6	0
135	LPC a C24:0	0	PC aa C32:0	0	PC aa C38:4	0	PC aa C38:3	0	PC ae C30:1	0
136	LPC a C26:0	0	PC aa C32:1	0	PC aa C38:5	0	PC aa C38:4	0	PC ae C30:2	0

Table A.19 (continued)

Order #	Milk Predictor	Coefficient	Yoghurt Predictor	Coefficient	Cheese Predictor	Coefficient	Butter Predictor	Coefficient	Total dairy products Predictor	Coefficient
137	LPC a C28:0	0	PC aa C32:2	0	PC aa C38:6	0	PC aa C38:5	0	PC ae C32:1	0
138	LPC a C28:1	0	PC aa C34:4	0	PC aa C40:1	0	PC aa C40:4	0	PC ae C32:2	0
139	PC aa C26:0	0	PC aa C36:4	0	PC aa C40:2	0	PC aa C40:5	0	PC ae C34:1	0
140	PC aa C30:0	0	PC aa C36:5	0	PC aa C40:3	0	PC aa C42:1	0	PC ae C34:2	0
141	PC aa C32:1	0	PC aa C36:6	0	PC aa C40:4	0	PC aa C42:2	0	PC ae C36:0	0
142	PC aa C32:2	0	PC aa C38:0	0	PC aa C40:5	0	PC aa C42:4	0	PC ae C36:1	0
143	PC aa C34:1	0	PC aa C38:1	0	PC aa C40:6	0	PC aa C42:6	0	PC ae C36:2	0
144	PC aa C34:2	0	PC aa C38:3	0	PC aa C42:0	0	PC ae C30:1	0	PC ae C36:3	0
145	PC aa C34:3	0	PC aa C38:5	0	PC aa C42:1	0	PC ae C30:2	0	PC ae C36:4	0
146	PC aa C36:2	0	PC aa C40:2	0	PC aa C42:2	0	PC ae C32:1	0	PC ae C36:5	0
147	PC aa C36:3	0	PC aa C40:4	0	PC ae C30:1	0	PC ae C32:2	0	PC ae C38:1	0
148	PC aa C36:5	0	PC aa C40:6	0	PC ae C32:1	0	PC ae C34:2	0	PC ae C38:2	0
149	PC aa C36:6	0	PC aa C42:0	0	PC ae C32:2	0	PC ae C34:3	0	PC ae C38:3	0
150	PC aa C38:0	0	PC aa C42:1	0	PC ae C34:1	0	PC ae C36:3	0	PC ae C38:4	0
151	PC aa C38:4	0	PC aa C42:4	0	PC ae C34:2	0	PC ae C36:4	0	PC ae C38:5	0
152	PC aa C40:6	0	PC aa C42:6	0	PC ae C36:1	0	PC ae C36:5	0	PC ae C38:6	0
153	PC aa C42:2	0	PC ae C30:2	0	PC ae C36:3	0	PC ae C38:1	0	PC ae C40:1	0
154	PC aa C42:5	0	PC ae C32:1	0	PC ae C38:0	0	PC ae C38:2	0	PC ae C40:2	0
155	PC ae C30:0	0	PC ae C32:2	0	PC ae C38:1	0	PC ae C38:3	0	PC ae C40:3	0
156	PC ae C32:1	0	PC ae C34:0	0	PC ae C38:3	0	PC ae C38:5	0	PC ae C40:4	0
157	PC ae C32:2	0	PC ae C34:1	0	PC ae C38:4	0	PC ae C38:6	0	PC ae C40:5	0
158	PC ae C34:0	0	PC ae C34:2	0	PC ae C40:1	0	PC ae C40:2	0	PC ae C40:6	0
159	PC ae C34:2	0	PC ae C36:0	0	PC ae C40:3	0	PC ae C40:3	0	PC ae C42:0	0
160	PC ae C34:3	0	PC ae C36:1	0	PC ae C40:4	0	PC ae C40:4	0	PC ae C42:1	0
161	PC ae C36:2	0	PC ae C38:0	0	PC ae C40:5	0	PC ae C40:5	0	PC ae C42:2	0
162	PC ae C36:3	0	PC ae C38:1	0	PC ae C40:6	0	PC ae C40:6	0	PC ae C42:3	0
163	PC ae C36:5	0	PC ae C38:2	0	PC ae C42:0	0	PC ae C42:0	0	PC ae C42:4	0
164	PC ae C38:0	0	PC ae C38:3	0	PC ae C42:2	0	PC ae C42:3	0	PC ae C42:5	0
165	PC ae C38:1	0	PC ae C38:4	0	PC ae C42:3	0	PC ae C42:5	0	PC ae C44:3	0
166	PC ae C38:2	0	PC ae C38:6	0	PC ae C42:4	0	PC ae C44:3	0	PC ae C44:5	0
167	PC ae C38:3	0	PC ae C40:5	0	PC ae C42:5	0	PC ae C44:5	0	PC ae C44:6	0
168	PC ae C38:6	0	PC ae C42:2	0	PC ae C44:5	0	SM C16:0	0	SM C18:0	0
169	PC ae C40:2	0	SM C16:1	0	PC ae C44:6	0	SM C18:0	0	SM C18:1	0
170	PC ae C40:3	0	SM C18:0	0	SM C16:0	0	SM C18:1	0	SM C20:2	0

Table A.19 (continued)

Order #	Milk Predictor	Coefficient	Yoghurt Predictor	Coefficient	Cheese Predictor	Coefficient	Butter Predictor	Coefficient	Total dairy products Predictor	Coefficient
171	PC ae C40:4	0	SM C20:2	0	SM C20:2	0	SM C20:2	0	SM-OH C16:1	0
172	PC ae C40:6	0	SM-OH C16:1	0	SM-OH C22:1	0	SM-OH C22:1	0	SM-OH C22:1	0
173	PC ae C42:1	0	SM-OH C22:1	0	SM -OH C22:2	0	SM -OH C22:2	0	SM -OH C22:2	0
174	PC ae C42:3	0	SM -OH C22:2	0	Hexose	0	Hexose	0	Hexose	0
	age	-0.017	age	0	age	-0.43	age	-0.003	age	-0.002
	sex	-0.294	sex	0.088	sex	2.326	sex	0	sex	0
	test site	0.243	test site	-0.027	test site	-0.024	test site	-0.272	test site	0
	smoking status	-0.016	smoking status	-1.678	smoking status	0	smoking status	0.108	smoking status	0
	physical activity	0.012	physical activity	0	physical activity	0	physical activity	0	physical activity	0.009
	lipid-lowering medication	-0.011	lipid-lowering medication	-0.129	lipid-lowering medication	0	lipid-lowering medication	0	lipid-lowering medication	0
	HRT	0	HRT	-1.312	HRT	-6.561	HRT	0	HRT	0
	BMI	0	BMI	0	BMI	0.332	BMI	0	BMI	0.05
	intercept	0.062	intercept	-1.655	intercept	0.284	intercept	-0.713	intercept	0.438

*The metabolites are presented in order of decreasing absolute value of the elastic net coefficients, whereas the rest of the predictors used in the models are presented in the end, along with the model intercept

†Coefficients are derived from the main models, which include age, sex, test site, smoking status, physical activity, lipid lowering medication, hormone-replacement therapy, body mass index and the metabolites

Table A.20 Coefficients of predictors in decreasing order of magnitude of the absolute values* for milk and total dairy products as derived from elastic net linear prediction models † in the discovery set of the Fenland study with all 174 metabolites included

	Milk		**Total dairy products**	
Order #	**Predictor**	**Coefficient**	**Predictor**	**Coefficient**
1	SM-OH C14:1	0.078	SM-OH C14:1	0.11
2	SM C16:1	-0.055	LPC a C18:0	-0.073
3	LPC a C18:0	-0.034	LPC a C17:0	0.072
4	PC aa C28:1	0.033	SM C16:1	-0.053
5	LPC a C17:0	0.026	PC aa C28:1	0.037
6	Pro	0.023	SM C24:0	-0.026
7	SM C24:0	-0.022	SM C16:0	-0.024
8	LPC a C20:3	0.02	PC aa C32:1	0.017
9	C0	0.019	t4OHPro	-0.015
10	LPC a C18:2	-0.019	Pro	0.014
11	Orn	-0.018	PC ae C34:1	0.013
12	C141	-0.017	PC ae C38:4	0.013
13	PC aa C40:5	0.017	C51	0.012
14	Ile	-0.016	C141	-0.012
15	His	-0.014	C182	-0.012
16	PC aa C38:1	-0.014	PC aa C38:1	-0.012
17	Cit	0.013	PC ae C34:3	-0.012
18	SM C16:0	-0.013	His	-0.011
19	Creatinine	0.012	LPC a C14:0	0.011
20	t4OHPro	-0.012	PC aa C40:5	0.011
21	C51	0.012	Ile	-0.01
22	C6C41DC	-0.012	PC ae C36:4	-0.01
23	PC aa C38:3	0.012	PC aa C26:0	-0.009
24	Thr	-0.011	PC aa C34:4	-0.009
25	PC aa C32:3	-0.011	PC aa C40:4	0.009
26	PC aa C34:4	-0.011	PC ae C38:2	-0.009
27	Trp	0.01	Ala	-0.008
28	SDMA	-0.01	alpha-AAA	0.008
29	Spermidine	0.009	PC ae C36:2	0.008
30	LPC a C16:1	0.009	PC ae C44:3	-0.008
31	PC aa C24:0	-0.009	SM C24:1	-0.008
32	PC aa C40:4	0.009	AcOrn	0.007
33	PC ae C30:2	-0.009	Kynurenine	0.006
34	PC ae C34:3	-0.009	PC aa C42:1	-0.006

Table A.20 (continued)

	Milk		Total dairy products	
Order #	Predictor	Coefficient	Predictor	Coefficient
35	SM-OH C22:1	-0.009	PC ae C30:2	-0.006
36	C181OH	0.008	C8	-0.005
37	Lys	0.007	PC aa C32:3	-0.005
38	alpha-AAA	0.007	PC ae C30:0	0.005
39	PC aa C32:0	0.007	PC ae C38:0	-0.005
40	PC aa C42:0	0.007	Arg	0.004
41	SM -OH C22:2	0.007	Trp	0.004
42	Taurine	0.006	c4OHPro	-0.004
43	C161	0.006	C51DC	-0.004
44	C16OH	0.006	C16	0.004
45	PC ae C40:2	-0.006	SM-OH C16:1	0.004
46	PC ae C40:5	-0.006	Orn	-0.003
47	PC ae C44:3	-0.006	Val	0.003
48	C12	-0.005	Sarcosine	-0.003
49	PC aa C40:3	-0.005	Spermidine	0.003
50	PC ae C38:4	0.005	C5MDC	-0.003
51	PC ae C44:5	0.005	LPC a C16:1	0.003
52	SM C20:2	0.005	LPC a C26:1	-0.003
53	Sarcosine	-0.004	PC aa C42:5	-0.003
54	C142	-0.004	SDMA	-0.002
55	Ala	-0.003	C3DCMC5OH	-0.002
56	Asp	0.003	C142	-0.002
57	Ser	-0.003	C18	-0.002
58	C41	-0.003	LPC a C26:0	-0.002
59	C51DC	-0.003	PC ae C34:0	0.002
60	C3DCMC5OH	-0.003	Taurine	0.001
61	C18	-0.003	C0	0.001
62	PC aa C40:2	0.003	C4	0.001
63	PC ae C42:5	0.003	C61	-0.001
64	Arg	0.002	C102	-0.001
65	Glu	-0.002	C142OH	-0.001
66	Gly	-0.002	LPC a C18:2	-0.001
67	Phe	-0.002	LPC a C28:0	-0.001
68	C16	0.002	PC ae C36:1	0.001

Table A.20 (continued)

	Milk		Total dairy products	
Order #	**Predictor**	**Coefficient**	**Predictor**	**Coefficient**
69	C182	-0.002	PC ae C42:4	-0.001
70	PC aa C36:2	-0.002	SM -OH C22:2	0.001
71	PC ae C34:1	0.002	Asn	0
72	PC ae C44:6	0.002	Asp	0
73	Gln	0.001	Cit	0
74	Leu	0.001	Gln	0
75	Kynurenine	0.001	Glu	0
76	Serotonin	-0.001	Gly	0
77	C2	-0.001	Leu	0
78	PC aa C36:6	-0.001	Lys	0
79	PC aa C42:1	-0.001	Met	0
80	PC aa C42:4	0.001	Phe	0
81	PC ae C38:0	-0.001	Ser	0
82	PC ae C40:6	-0.001	Thr	0
83	PC ae C42:4	-0.001	Tyr	0
84	Asn	0	PEA	0
85	Met	0	Creatinine	0
86	Tyr	0	Met-SO	0
87	Val	0	Serotonin	0
88	PEA	0	C2	0
89	AcOrn	0	C3	0
90	Met-SO	0	C31	0
91	c4OHPro	0	C3OH	0
92	C3	0	C41	0
93	C31	0	C3DCC4OH	0
94	C3OH	0	C5	0
95	C4	0	C5DCC6OH	0
96	C3DCC4OH	0	C6C41DC	0
97	C5	0	C7DC	0
98	C5DCC6OH	0	C9	0
99	C5MDC	0	C10	0
100	C61	0	C101	0
101	C7DC	0	C12	0
102	C8	0	C12DC	0

Table A.20 (continued)

	Milk		Total dairy products	
Order #	Predictor	Coefficient	Predictor	Coefficient
103	C9	0	C121	0
104	C10	0	C14	0
105	C101	0	C141OH	0
106	C102	0	C161	0
107	C12DC	0	C161OH	0
108	C121	0	C162	0
109	C14	0	C162OH	0
110	C141OH	0	C16OH	0
111	C142OH	0	C181	0
112	C161OH	0	C181OH	0
113	C162	0	LPC a C20:4	0
114	C162OH	0	LPC a C16:0	0
115	C181	0	LPC a C18:1	0
116	LPC a C14:0	0	LPC a C20:3	0
117	LPC a C20:4	0	LPC a C24:0	0
118	LPC a C26:1	0	LPC a C28:1	0
119	LPC a C16:0	0	PC aa C24:0	0
120	LPC a C18:1	0	PC aa C30:0	0
121	LPC a C24:0	0	PC aa C32:0	0
122	LPC a C26:0	0	PC aa C32:2	0
123	LPC a C28:0	0	PC aa C34:1	0
124	LPC a C28:1	0	PC aa C34:2	0
125	PC aa C26:0	0	PC aa C34:3	0
126	PC aa C30:0	0	PC aa C36:0	0
127	PC aa C32:1	0	PC aa C36:1	0
128	PC aa C32:2	0	PC aa C36:2	0
129	PC aa C34:1	0	PC aa C36:3	0
130	PC aa C34:2	0	PC aa C36:4	0
131	PC aa C34:3	0	PC aa C36:5	0
132	PC aa C36:0	0	PC aa C36:6	0
133	PC aa C36:1	0	PC aa C38:0	0
134	PC aa C36:3	0	PC aa C38:3	0
135	PC aa C36:4	0	PC aa C38:4	0
136	PC aa C36:5	0	PC aa C38:5	0

Table A.20 (continued)

	Milk		Total dairy products	
Order #	**Predictor**	**Coefficient**	**Predictor**	**Coefficient**
137	PC aa C38:0	0	PC aa C38:6	0
138	PC aa C38:4	0	PC aa C40:1	0
139	PC aa C38:5	0	PC aa C40:2	0
140	PC aa C38:6	0	PC aa C40:3	0
141	PC aa C40:1	0	PC aa C40:6	0
142	PC aa C40:6	0	PC aa C42:0	0
143	PC aa C42:2	0	PC aa C42:2	0
144	PC aa C42:5	0	PC aa C42:4	0
145	PC aa C42:6	0	PC aa C42:6	0
146	PC ae C30:0	0	PC ae C30:1	0
147	PC ae C30:1	0	PC ae C32:1	0
148	PC ae C32:1	0	PC ae C32:2	0
149	PC ae C32:2	0	PC ae C34:2	0
150	PC ae C34:0	0	PC ae C36:0	0
151	PC ae C34:2	0	PC ae C36:3	0
152	PC ae C36:0	0	PC ae C36:5	0
153	PC ae C36:1	0	PC ae C38:1	0
154	PC ae C36:2	0	PC ae C38:3	0
155	PC ae C36:3	0	PC ae C38:5	0
156	PC ae C36:4	0	PC ae C38:6	0
157	PC ae C36:5	0	PC ae C40:1	0
158	PC ae C38:1	0	PC ae C40:2	0
159	PC ae C38:2	0	PC ae C40:3	0
160	PC ae C38:3	0	PC ae C40:4	0
161	PC ae C38:5	0	PC ae C40:5	0
162	PC ae C38:6	0	PC ae C40:6	0
163	PC ae C40:1	0	PC ae C42:0	0
164	PC ae C40:3	0	PC ae C42:1	0
165	PC ae C40:4	0	PC ae C42:2	0
166	PC ae C42:0	0	PC ae C42:3	0
167	PC ae C42:1	0	PC ae C42:5	0
168	PC ae C42:2	0	PC ae C44:5	0
169	PC ae C42:3	0	PC ae C44:6	0
170	SM C18:0	0	SM C18:0	0

Table A.20 (continued)

Order #	Milk Predictor	Coefficient	Total dairy products Predictor	Coefficient
171	SM C18:1	0	SM C18:1	0
172	SM C24:1	0	SM C20:2	0
173	SM-OH C16:1	0	SM-OH C22:1	0
174	Hexose	0	Hexose	0
	age	0.001	age	0.001
	sex	0	sex	0
	test site	0.04	test site	0
	smoking status	0	smoking status	0
	physical activity	0	physical activity	0
	lipid-lowering medication	0	lipid-lowering medication	0
	HRT	0	HRT	-0.027
	BMI	0.001	BMI	0.003
	intercept	0.738	intercept	1.236

*The metabolites are presented in order of decreasing absolute value of the elastic net coefficients, whereas the rest of the predictors used in the models are presented in the end, along with the model intercept

†Coefficients are derived from the main models, which include age, sex, test site, smoking status, physical activity, lipid lowering medication, hormone-replacement therapy, body mass index and the metabolites

Table A.21 Coefficients of predictors for total and types of dairy products as derived from logistic and linear prediction models without penalisation* in the discovery set of the Fenland study with all 174 metabolites included

Predictor	Milk	Yoghurt	Cheese	Butter	Total dairy products	Milk (continuous)	Total dairy products (continuous)
Ala	-114.667	-13.475	-272.273	-29.816	-254.369	-0.005	-0.011
Arg	138.858	-134.948	133.374	163.449	293.77	0.005	0.008
Asn	-68.203	101.301	265.649	-114.776	-95.209	0.002	-0.003
Asp	79.938	-143.259	-143.113	-167.316	6.22	0.008	0.003
Cit	217.879	-30.699	-120.486	-55.147	185.831	0.019	0.005
Gln	120.646	-107.698	257.596	-88.471	-120.86	0.006	0.003
Glu	-146.368	-174.856	238.954	110.061	-110.022	-0.007	0.003
Gly	-162.413	-41.122	232.249	-99.369	-3.678	-0.007	-0.003
His	-116.733	162.942	112.922	125.97	-112.583	-0.025	-0.019
Ile	-305.317	-287.471	-199.094	164.266	-192.457	-0.029	-0.011
Leu	119.675	54.409	19.436	-92.748	231.49	0.009	0
Lys	52.997	108.234	-6.18	-231.907	134.562	0.009	-0.001
Met	-50.335	16.108	-58.387	4.079	165.469	-0.001	0.004
Orn	-212.742	-56.118	-136.096	208.394	-319.858	-0.024	-0.01
Phe	8.1	-14.929	-45.55	-68.665	-111.789	-0.004	-0.004
Pro	288.195	61.108	33.729	33.32	301.759	0.025	0.019
Ser	-239.987	-105.311	116.561	87.977	58.985	-0.007	0.001
Thr	-337.071	-134.345	390.399	39.275	-97.674	-0.014	-0.003
Trp	126.052	34.491	-80.467	-22.419	-6.987	0.016	0.01
Tyr	158.529	67.617	324.211	90.692	204.321	-0.009	-0.005
Val	109.365	202.219	-32.321	153.94	103.31	0.005	0.011
PEA	-112.661	-97.367	15.244	81.432	-93.432	0.005	-0.006
AcOrn	-13.053	60.063	-3.605	9.254	84.733	0.005	0.011
SDMA	-52.077	107.967	265.523	138.199	-118.346	-0.016	-0.008
alpha-AAA	211.429	112.383	-67.363	93.254	309.625	0.01	0.013
Creatinine	370.359	-186.038	-269.782	-114.552	-49.829	0.019	0.001
Kynurenine	41.504	264.978	411.408	-32.441	223.1	0.003	0.009
Met-SO	-17.544	49.816	-39.717	-61.007	-106.667	-0.002	0.001
c4OHPro	-159.829	-218.793	-340.114	15.958	-320.429	0.007	-0.003
t4OHPro	-192.465	-190.873	-368.287	109.232	-349.233	-0.02	-0.017
Sarcosine	-64.447	-178.76	-119.204	-35.498	-159.711	-0.007	-0.006
Serotonin	-110.115	180.739	118.909	-20.343	100.979	-0.009	-0.004
Spermidine	24.932	24.054	-214.372	-81.657	15.86	0.01	0.006
Taurine	158.915	274.595	-34.42	-179.552	-47.456	0.014	0.006
C0	203.499	61.137	-400.74	-186.54	40.207	0.03	0.009
C2	-52.768	158.584	204.657	60.924	-170.337	-0.015	-0.001
C3	240.648	163.589	-124.265	78.88	201.873	0	-0.002
C31	-58.48	49.34	-112.066	-126.674	-16.199	0.003	0.001
C3OH	-26.434	-25.083	-44.613	-15.015	-83.802	0.004	0.007
C4	-22.514	65.431	176.318	-66.388	124.846	-0.001	0.004
C41	-106.797	36.085	-123.297	31.886	169.077	-0.011	-0.005
C3DCC4OH	74.47	-103.906	-69.564	-58.939	-18.041	0.007	0.005
C5	-16.339	-200.036	-51.192	101.783	-13.522	-0.005	-0.005
C51	101.55	152.446	337.695	-149.948	175.579	0.02	0.019
C51DC	34.714	-51.445	86.093	3.683	-174.873	-0.007	-0.009
C5DCC6OH	-31.554	8.747	-15.95	38.715	107.51	-0.002	0.008
C5MDC	-28.313	94.217	-252.746	-11.629	1.949	0	-0.008
C3DCMC5OH	-25.762	148.113	5.736	-127.635	-365.39	-0.007	-0.006

Table A.21 (continued)

Predictor	Milk	Yoghurt	Cheese	Butter	Total dairy products	Milk (continuous)	Total dairy products (continuous)
C6C41DC	-183.242	-71.93	-97.666	51.848	-18.284	-0.015	-0.004
C61	177.754	-72.791	38.026	5.182	47.496	-0.003	-0.007
C7DC	103.013	70.146	-160.22	100.094	6.215	-0.004	-0.004
C8	-76.435	19.912	-180.556	3.4	-27.735	-0.001	-0.02
C9	-55.864	-267.99	106.371	449.273	117.885	-0.005	0.003
C10	-121.246	-129.112	95.936	86.786	63.638	0.008	0.02
C101	7.525	-76.313	-329.793	-61.522	-80.278	-0.003	-0.004
C102	15.399	281.207	-219.422	-307.257	-30.574	0.001	-0.002
C12	-78.086	-103.467	174.781	65.868	-27.406	-0.016	0.001
C12DC	-27.037	51.681	62.917	-84.705	-50.048	-0.002	-0.007
C121	101.936	-13.754	-69.486	94.568	166.088	0.024	0.017
C14	-16.288	137.686	63.978	-28.281	48.656	-0.005	-0.005
C141	-321.08	-212.4	18.972	138.301	-202.158	-0.038	-0.043
C141OH	58.149	124.764	112.809	29.759	179.609	0.009	0.008
C142	-214.237	-11.128	-48.01	-171.086	-200.098	-0.008	0.002
C142OH	-36.288	69.588	-94.904	-204.579	-76.891	-0.003	-0.008
C16	-13.425	-159.109	111.708	115.849	188.154	0.008	0.014
C161	171.775	169.786	-35.09	-114.5	86.347	0.014	0.005
C161OH	2.733	33.538	184.432	130.001	43.964	-0.003	-0.003
C162	47.027	11.172	148.796	4.889	-25.603	0.008	0.006
C162OH	80.154	-87.243	-184.01	-64.985	-40.119	-0.005	-0.006
C16OH	111.623	-98.32	7.052	60.284	23.019	0.012	0.007
C18	-186.578	-247.864	-109.864	94.934	51.461	-0.013	-0.016
C181	27.052	40.092	-86.832	110.078	21.614	0.022	0.014
C181OH	127.515	-120.905	159.677	-100.554	-29.075	0.014	0.008
C182	-213.787	212.664	-68.011	-289.607	-164.108	-0.013	-0.017
LPC a C14:0	49.831	156.808	82.745	315.916	349.517	0.003	0.023
LPC a C16:1	131.846	261.269	19.095	-106.093	91.787	0.023	0.014
LPC a C17:0	301.307	212.685	319.973	529.748	627.877	0.047	0.078
LPC a C18:0	-245.195	-181.951	-272.12	-146.079	-288.527	-0.084	-0.109
LPC a C18:2	-66.423	46.646	23.73	-104.068	-295.201	-0.004	-0.01
LPC a C20:4	-82.388	-62.959	-23.984	6.551	-159.621	-0.024	-0.019
LPC a C26:1	4.977	50.625	-19.131	-73.184	-116.458	0.008	-0.006
LPC a C16:0	-152.474	14.741	-132.106	-191.498	-133.411	0.031	0.019
LPC a C18:1	-150.626	-59.249	-6.834	2.424	-212.825	-0.039	-0.002
LPC a C20:3	288.567	174.476	20.038	-221.398	52.742	0.053	0.035
LPC a C24:0	52.847	217.899	-50.295	-149.671	-207.996	0.003	0.006
LPC a C26:0	7.631	-39.894	-96.265	-218.713	5.238	-0.016	-0.02
LPC a C28:0	-16.184	-204.872	32.396	-84.411	-32.691	0.005	-0.008
LPC a C28:1	49.144	-14.277	133.169	39.566	108.733	0.005	0.001
PC aa C24:0	-162.25	2.989	-107.473	147.446	-98.112	-0.016	0.01
PC aa C26:0	-136.585	-80.869	-119.218	-128.985	-138.136	0.004	-0.006
PC aa C28:1	503.434	432.294	495.201	244.064	684.642	0.049	0.045
PC aa C30:0	74.702	99.178	136.299	107.856	308.67	-0.004	0.007
PC aa C32:0	66.485	-44.203	-35.785	-405.311	31.768	0.03	0.002
PC aa C32:1	116.232	-34.698	64.232	114.284	241.624	-0.019	0.006
PC aa C32:2	50.379	55.974	65.367	64.852	84.372	-0.012	-0.01
PC aa C32:3	-236.695	-99.975	187.832	174.746	-103.076	-0.017	-0.012

Table A.21 (continued)

Predictor	Milk	Yoghurt	Cheese	Butter	Total dairy products	Milk (continuous)	Total dairy products (continuous)
PC aa C34:1	-64.81	-41.998	-15.677	134.115	-4.386	-0.077	-0.06
PC aa C34:2	-8.147	-95.577	59.575	-51.906	-151.657	0.056	0.053
PC aa C34:3	39.807	108.114	43.255	-97.771	-73.787	0.015	-0.001
PC aa C34:4	-256.17	-2.866	-79.285	-86.082	-2.704	-0.009	-0.023
PC aa C36:0	95.409	293.322	-24.181	-49.858	53.782	0.003	0.013
PC aa C36:1	-168.328	-273.277	-190.801	126.253	-147.337	-0.002	0.034
PC aa C36:2	-83.421	-231.993	-113.024	23.365	-314.384	-0.032	-0.076
PC aa C36:3	-3.038	150.776	12.887	-44.812	-16.271	-0.097	-0.025
PC aa C36:4	-205.502	36.497	100.489	-28.916	-55.046	0.025	0.006
PC aa C36:5	26.752	-77.053	206.543	-49.924	-71.22	-0.006	0.007
PC aa C36:6	-11.028	27.215	57.816	-87.771	39.911	-0.009	-0.011
PC aa C38:0	24.387	24.711	-275.554	3.704	-171.602	0.001	0.004
PC aa C38:1	-318.283	32.438	-304.206	-112.167	-165.86	-0.016	-0.015
PC aa C38:3	150.75	63.725	-119.523	-182.357	38.953	0.123	0.056
PC aa C38:4	29.089	-133.196	-36.61	1.106	-82.059	-0.019	-0.003
PC aa C38:5	101.731	-84.287	64.526	8.775	16.641	0.011	0.057
PC aa C38:6	-106.984	119.779	-1.532	-79.14	-73.677	0.04	0.002
PC aa C40:1	-96.367	-148.197	-11.904	-10.189	-92.649	0.001	-0.002
PC aa C40:2	37.183	48.834	75.215	-136.209	-34.436	0.007	0.002
PC aa C40:3	-88.189	123.852	208.157	-32.901	-309.669	-0.01	-0.002
PC aa C40:4	124.818	-56.752	104.129	-131.08	-53.615	0.01	0.014
PC aa C40:5	241.932	119.175	-192.569	154.818	216.519	0.018	0.018
PC aa C40:6	-64.232	-86.319	-270.365	-75.467	-129.512	0.015	-0.002
PC aa C42:0	199.518	-2.94	129.749	-210.786	79.599	0.01	0.003
PC aa C42:1	188.373	-128.814	-4.48	-89.575	-175.1	-0.005	-0.01
PC aa C42:2	-6.744	-66.036	-82.987	-1.306	42.784	0.004	0.002
PC aa C42:4	131.745	64.288	274.578	58.033	-145.028	0.006	0.004
PC aa C42:5	-29.754	50.557	-49.462	-94.453	-144.499	-0.004	-0.007
PC aa C42:6	159.578	-136.132	90.825	-74.674	182.958	0.001	-0.002
PC ae C30:0	65.42	67.919	311.435	268.234	477.881	-0.003	0.009
PC ae C30:1	-205.506	-224.57	-34.223	28.164	-168.72	0.003	0.008
PC ae C30:2	-324.466	-60.325	17.856	224.694	76.93	-0.018	-0.013
PC ae C32:1	76.428	-59.248	42.164	-179.459	71.961	0.006	0.009
PC ae C32:2	17.128	41.098	-29.929	92.802	-20.842	-0.013	-0.013
PC ae C34:0	-51.378	-161.654	193.323	235.898	443.687	-0.011	-0.001
PC ae C34:1	75.654	29.526	141.95	191.758	297.453	0.005	0.003
PC ae C34:2	84.873	16.964	43.723	137.104	112.915	0.011	0.011
PC ae C34:3	-148.926	-96.512	-326.246	-70.832	-284.456	-0.017	-0.017
PC ae C36:0	-116.249	1.657	9.972	29.567	-34.523	0	0.003
PC ae C36:1	122.628	-25.564	-12.09	193.576	295.701	-0.001	-0.002
PC ae C36:2	96.48	95.98	379.403	323.72	394.67	0.001	0.021
PC ae C36:3	-130.428	-112.485	-134.584	-92.281	-78.435	-0.002	-0.013
PC ae C36:4	-94.951	-152.295	-300.519	108.909	-115.575	-0.03	-0.037
PC ae C36:5	-96.162	-104.031	-209.859	94.049	-96.563	0.024	0.022
PC ae C38:0	-128.577	17.302	88.267	-129.921	-258.796	-0.013	-0.016
PC ae C38:1	-42.561	46.616	27.752	86.113	13.483	0.085	0.066
PC ae C38:2	9.927	-40.816	191.757	-72.577	-123.012	-0.067	-0.064
PC ae C38:3	53.421	-6.092	-137.04	155.775	255.896	0.006	0.004

Table A.21 (continued)

Predictor	Milk	Yoghurt	Cheese	Butter	Total dairy products	Milk (continuous)	Total dairy products (continuous)
PC ae C38:4	223.922	-5.959	183.265	292.494	291.573	0.031	0.035
PC ae C38:5	-172.989	-96.414	-181.112	20.863	-136.315	-0.013	-0.003
PC ae C38:6	35.019	-86.387	-259.037	-97.569	-32.46	0.01	-0.001
PC ae C40:1	175.288	14.325	-8.019	-174.137	-127.142	0.006	-0.001
PC ae C40:2	-122.843	-219.348	-139.464	94.137	-111.175	-0.011	-0.042
PC ae C40:3	-45.62	-168.171	-45.024	13.639	-260.18	0.053	0.095
PC ae C40:4	26.389	189.489	28.392	-61.577	-6.592	0.071	0.02
PC ae C40:5	-141.58	25.599	170.397	-9.498	-39.109	-0.041	-0.01
PC ae C40:6	-89.874	125.544	92.175	1.295	16.143	-0.024	-0.01
PC ae C42:0	91.107	52.725	230.646	95.71	11.647	0.001	0.002
PC ae C42:1	62.131	123.29	-152.916	-174.634	1.722	-0.115	-0.069
PC ae C42:2	120.295	-23.409	109.642	-0.829	15.268	0.019	-0.047
PC ae C42:3	22.882	-155.173	-59.759	44.614	-110.443	0.02	0.009
PC ae C42:4	-106.958	235.781	61.978	-72.419	-87.545	-0.051	-0.003
PC ae C42:5	214.603	-93.557	134.274	-5.532	135.241	0.005	0.001
PC ae C44:3	-200.426	-139.425	55.915	-96.164	-169.005	-0.01	-0.012
PC ae C44:5	3.199	-207.978	-63.355	25.178	-152.543	0.009	0.002
PC ae C44:6	230.697	285.709	-124.623	-189.553	-6.388	0.005	0
SM C16:0	-157.731	-268.587	-177.076	-205.661	-303.774	-0.022	-0.032
SM C16:1	-382.606	60.633	-260.1	-205.841	-338.261	-0.061	-0.062
SM C18:0	236.856	-26.366	-94.665	-103.763	31.364	0.009	0.006
SM C18:1	-297.83	130.968	-275.4	-6.182	-142.573	-0.006	-0.004
SM C20:2	190.243	-28.957	89.175	57.587	-19.563	0.012	0.002
SM C24:0	-341.698	-303.44	-63.071	-119.424	-452.158	-0.024	-0.027
SM C24:1	-103.48	-177.096	-142.692	-283.462	-409.743	-0.002	-0.01
SM-OH C14:1	678	409.681	602.433	472.175	941.878	0.083	0.118
SM-OH C16:1	163.775	82.308	312.934	365.193	480.346	-0.005	0.007
SM-OH C22:1	-245.526	-50.245	74.389	49.342	-51.791	-0.017	-0.006
SM -OH C22:2	185.849	64.066	134.661	107.042	293.475	0.013	0.007
Hexose	200.258	-76.054	168.219	-124.193	-97.978	0.007	-0.001
age	-4.633	33.647	-105.29	-11.999	28.849	0	0.001
sex	-153.416	192.655	246.222	-1.679	9.28	-0.064	0.052
test site	526.453	-33.063	-308.135	-619.958	61.863	0.043	0.003
smoking status	-137.922	-749.571	86.539	149.909	-97.312	-0.005	0
physical activity	-16.901	-13.011	8.995	-5.743	-4.103	0	0
lipid-lowering	-107.525	-85.375	30.8	-55.882	-33.292	-0.026	-0.036
HRT	209.915	-609.862	-378.867	86.267	44.105	-0.039	-0.015
BMI	-7.934	26.87	137.865	-59.713	114.371	-0.001	0.003
intercept	4.998	-50.365	27.672	14.464	12.11	0.964	1.163

*Coefficients are derived from the main models, which include age, sex, test site, smoking status, physical activity, lipid lowering medication, hormone-replacement therapy, body mass index and the metabolites

Table A.22 Coefficients of predictors in decreasing order of magnitude of the absolute values* for total and types of dairy products as derived from elastic net logistic prediction models† in the discovery set of the Fenland study with the 82 overlapping metabolites with the diabetes case-cohort set nested within the EPIC Norfolk study

	Milk		Yoghurt		Cheese		Butter		Total dairy products	
Order #	**Predictor**	**Coefficient**	**Predictor**	**Coefficient**	**Predictor**	**Coefficient**	**Predictor**	**Coefficient**	**Predictor**	**Coefficient**
1	SM-OH C14:1	0.587	SM-OH C14:1	0.271	SM-OH C14:1	0.475	SM-OH C14:1	0.407	SM-OH C14:1	1.185
2	LPC a C20:3	0.324	trans-OH-Pro	-0.166	trans-OH-Pro	-0.263	PC ae C34:1	0.283	SM C16:1	-0.462
3	SM C16:1	-0.279	LPC a C17:0	0.13	LPC a C18:0	-0.215	LPC a C17:0	0.267	LPC a C18:0	-0.424
4	Ile	-0.238	taurine	0.128	C0	-0.189	LPC a C14:0	0.233	LPC a C17:0	0.4
5	PC aa C36:4	-0.188	C16	-0.127	kynurenine	0.164	PC aa C32:0	-0.223	SM C24:1	-0.273
6	C14:1	-0.182	C14:1	-0.118	SM C18:1	-0.16	C18:2	-0.183	SM C16:0	-0.217
7	LPC a C17:0	0.166	PC aa C36:2	-0.118	SM C16:0	-0.152	LPC a C16:0	-0.156	LPC a C14:0	0.172
8	SM C18:1	-0.163	C18	-0.117	creatinine	-0.105	LPC a C18:0	-0.141	trans-OH-Pro	-0.12
9	LPC a C18:0	-0.159	Ile	-0.109	Ala	-0.104	PC ae C36:2	0.138	PC aa C32:1	0.098
10	SM C16:0	-0.157	SM C16:0	-0.108	Gly	0.1	PC aa C38:5	-0.127	kynurenine	0.096
11	Ser	-0.153	Met	0.105	Glu	0.091	SM C16:1	-0.126	Orn	-0.083
12	SM C18:0	0.144	PC aa C36:1	-0.103	PC ae C36:2	0.09	PC ae C36:1	0.113	LPC a C18:2	-0.077
13	PC aa C38:5	0.139	C2	0.094	LPC a C17:0	0.086	Lys	-0.099	SM C18:1	-0.074
14	trans-OH-Pro	-0.137	LPC a C16:1	0.086	PC aa C30:0	0.074	SM C24:1	-0.098	Arg	0.057
15	LPC a C20:4	-0.122	PC aa C30:0	0.086	LPC a C14:0	0.071	LPC a C20:3	-0.093	C18:2	-0.051
16	creatinine	0.12	LPC a C20:4	-0.076	C5-M-DC	-0.067	taurine	-0.074	Phe	0.034
17	His	-0.118	C18:2	0.075	Asn	0.059	Glu	0.066	Ile	-0.026
18	SM-OH C22:1	-0.117	sarcosine	-0.07	Gln	0.051	Orn	0.058	C6	-0.024
19	Cit	0.113	Asp	-0.066	Trp	0.043	C5:1	-0.057	Ala	-0.022
20	Orn	-0.106	Glu	-0.063	C5:1	0.037	C18	0.051	alpha-AAA	0.019
21	C18	-0.105	SM C24:1	-0.062	hexose	0.034	C12	0.05	PC aa C36:2	-0.014
22	LPC a C18:1	-0.104	C5:1	0.061	LPC a C16:1	0.032	C4-OH	-0.049	Asn	0
23	Phe	0.103	PC aa C36:3	0.058	SM C18:0	-0.028	PC aa C34:2	-0.046	Asp	0
24	taurine	0.1	PC aa C34:2	-0.053	C8	-0.027	Asn	-0.044	Cit	0
25	C0	0.097	PC aa C36:5	0.047	Ser	0.023	C14:1	0.037	Gln	0
26	LPC a C16:1	0.097	PC aa C38:4	-0.043	SDMA	0.022	Tyr	0.036	Glu	0
27	PC aa C36:1	-0.097	C14	0.039	Thr	-0.01	PC aa C36:1	0.031	Gly	0
28	C16	0.094	Gly	-0.035	PC aa C38:4	-0.005	Gly	-0.03	His	0
29	C3	0.091	Lys	0.034	Arg	0	Arg	0.026	Leu	0
30	C5:1	0.089	Asn	0.032	Asp	0	trans-OH-Pro	0.023	Lys	0
31	Gln	0.085	creatinine	-0.026	Cit	0	C16	0.022	Met	0
32	Leu	0.084	Ser	-0.024	His	0	serotonin	-0.017	Pro	0
33	SM-OH C22:2	0.082	Ac-Orn	0.024	Ile	0	PC aa C36:5	-0.015	Ser	0
34	serotonin	-0.078	serotonin	0.023	Leu	0	Ile	0.01	Thr	0

Table A.22 (continued)

Order #	Milk Predictor	Coefficient	Yoghurt Predictor	Coefficient	Cheese Predictor	Coefficient	Butter Predictor	Coefficient	Total dairy products Predictor	Coefficient
35	LPC a C16:0	-0.077	Gln	-0.021	Lys	0	C0	-0.006	Trp	0
36	C18:2	-0.075	Met-SO	0.018	Met	0	PC aa C36:3	-0.002	Tyr	0
37	PC aa C32:0	0.073	C3	0.018	Orn	0	SDMA	0.001	Val	0
38	PC aa C36:5	-0.073	Val	-0.017	Phe	0	Ala	0	Ac-Orn	0
39	Pro	-0.071	SM-OH C22:2	0.016	Pro	0	Asp	0	SDMA	0
40	LPC a C18:2	-0.069	His	0.014	Tyr	0	Cit	0	creatinine	0
41	PC aa C38:4	0.068	Phe	0.014	Val	0	Gln	0	Met-SO	0
42	alpha-AAA	0.062	Pro	-0.014	Ac-Orn	0	His	0	sarcosine	0
43	hexose	0.059	Trp	0.014	alpha-AAA	0	Leu	0	serotonin	0
44	SM C24:1	-0.057	kynurenine	0.013	Met-SO	0	Met	0	taurine	0
45	Tyr	0.055	hexose	-0.012	sarcosine	0	Phe	0	C0	0
46	Asp	0.05	PC aa C32:0	0.011	serotonin	0	Pro	0	C2	0
47	C6	-0.048	C4	0.01	taurine	0	Ser	0	C3	0
48	C2	-0.044	Arg	-0.006	C2	0	Thr	0	C4	0
49	PC aa C30:0	-0.043	C6	-0.006	C3	0	Trp	0	C4-OH	0
50	PC ae C34:1	0.042	Ala	-0.005	C4	0	Val	0	C5:1	0
51	Lys	0.041	LPC a C14:0	0.004	C4-OH	0	Ac-Orn	0	C5-M-DC	0
52	SDMA	-0.04	Leu	0.003	C6	0	alpha-AAA	0	C8	0
53	C8	-0.039	Cit	0	C10	0	creatinine	0	C10	0
54	C4	-0.037	Orn	0	C12	0	kynurenine	0	C12	0
55	PC aa C34:2	0.037	Thr	0	C14	0	Met-SO	0	C14	0
56	Thr	0.034	Tyr	0	C14:1	0	sarcosine	0	C14:1	0
57	C4-OH	0.033	SDMA	0	C16	0	C2	0	C16	0
58	PC ae C34:2	-0.033	alpha-AAA	0	C18	0	C3	0	C18	0
59	Arg	0.032	C0	0	C18:2	0	C4	0	LPC a C16:1	0
60	PC aa C32:2	-0.03	C4-OH	0	LPC a C18:2	0	C5-M-DC	0	LPC a C20:4	0
61	Ala	-0.028	C5-M-DC	0	LPC a C20:4	0	C6	0	LPC a C16:0	0
62	kynurenine	0.027	C8	0	LPC a C16:0	0	C8	0	LPC a C18:1	0
63	Glu	-0.024	C10	0	LPC a C18:1	0	C10	0	LPC a C20:3	0
64	PC aa C32:1	0.023	C12	0	LPC a C20:3	0	C14	0	PC aa C30:0	0
65	sarcosine	-0.02	LPC a C18:0	0	PC aa C32:0	0	LPC a C16:1	0	PC aa C32:0	0
66	Trp	0.017	LPC a C18:2	0	PC aa C32:1	0	LPC a C18:2	0	PC aa C32:2	0
67	C14	0.017	LPC a C16:0	0	PC aa C32:2	0	LPC a C20:4	0	PC aa C34:1	0
68	Gly	-0.014	LPC a C18:1	0	PC aa C34:1	0	LPC a C18:1	0	PC aa C34:2	0

Table A.22 (continued)

Order #	Milk Predictor	Coefficient	Yoghurt Predictor	Coefficient	Cheese Predictor	Coefficient	Butter Predictor	Coefficient	Total dairy products Predictor	Coefficient
69	Met	-0.014	LPC a C20:3	0	PC aa C34:2	0	PC aa C30:0	0	PC aa C36:1	0
70	Met-SO	0.014	PC aa C32:1	0	PC aa C36:1	0	PC aa C32:1	0	PC aa C36:3	0
71	Ac-Orn	0.012	PC aa C32:2	0	PC aa C36:2	0	PC aa C32:2	0	PC aa C36:4	0
72	PC ae C36:1	0.007	PC aa C34:1	0	PC aa C36:3	0	PC aa C34:1	0	PC aa C36:5	0
73	Asn	0	PC aa C36:4	0	PC aa C36:4	0	PC aa C36:2	0	PC aa C38:4	0
74	Val	0	PC aa C38:5	0	PC aa C36:5	0	PC aa C36:4	0	PC aa C38:5	0
75	C5-M-DC	0	PC ae C34:1	0	PC aa C38:5	0	PC aa C38:4	0	PC ae C34:1	0
76	C10	0	PC ae C34:2	0	PC ae C34:1	0	PC ae C34:2	0	PC ae C34:2	0
77	C12	0	PC ae C36:1	0	PC ae C34:2	0	SM C16:0	0	PC ae C36:1	0
78	LPC a C14:0	0	PC ae C36:2	0	PC ae C36:1	0	SM C18:0	0	PC ae C36:2	0
79	PC aa C34:1	0	SM C16:1	0	SM C16:1	0	SM C18:1	0	SM C18:0	0
80	PC aa C36:2	0	SM C18:0	0	SM C24:1	0	SM-OH C22:1	0	SM-OH C22:1	0
81	PC aa C36:3	0	SM C18:1	0	SM-OH C22:1	0	SM-OH C22:2	0	SM-OH C22:2	0
82	PC ae C36:2	0	SM-OH C22:1	0	SM-OH C22:2	0	hexose	0	hexose	0
	age	-0.015	age	0.026	age	-0.026	age	-0.004	age	0
	sex	-0.31	sex	0	sex	0	sex	0	sex	0
	test site	0.268	test site	-0.023	test site	-0.094	test site	-0.262	test site	0
	smoking status	-0.027	smoking status	-0.299	smoking status	0	smoking status	0.1	smoking status	0
	physical activity	0.012	physical activity	0.012	physical activity	0	physical activity	0	physical activity	0.014
	lipid-lowering	-0.129	lipid-lowering	-0.006	lipid-lowering	0	lipid-lowering	0	lipid-lowering	0
	HRT	0	HRT	-0.337	HRT	-0.204	HRT	0	HRT	0
	BMI	0	BMI	0.011	BMI	0.014	BMI	0	BMI	0.053
	intercept	-0.036	intercept	-1.594	intercept	0.396	intercept	-0.689	intercept	0.262

*The metabolites are presented in order of decreasing absolute value of the elastic net coefficients, whereas the rest of the predictors used in the models are presented in the end, along with the model intercept

†Coefficients are derived from the main models, which include age, sex, test site, smoking status, physical activity, lipid lowering medication, hormone-replacement therapy, body mass index and the metabolites

Table A.23 Coefficients of predictors in decreasing order of magnitude of the absolute values* for milk and total dairy products as derived from elastic net linear prediction models† in the discovery set of the Fenland study with the 82 overlapping metabolites with the diabetes case-cohort sub-study of the EPIC Norfolk study

	Milk		Total dairy products	
Order #	Predictor	Coefficient	Predictor	Coefficient
1	SM-OH C14:1	0.104	SM-OH C14:1	0.141
2	SM C16:1	-0.058	LPC a C18:0	-0.072
3	LPC a C20:3	0.041	LPC a C17:0	0.071
4	LPC a C18:0	-0.034	SM C16:1	-0.058
5	Phe	0.026	SM C16:0	-0.038
6	SM C16:0	-0.026	trans-OH-Pro	-0.024
7	C0	0.025	C18:2	-0.018
8	PC aa C38:5	0.025	Phe	0.017
9	LPC a C17:0	0.024	PC ae C34:1	0.016
10	Ile	-0.021	His	-0.015
11	His	-0.02	PC aa C32:1	0.015
12	LPC a C18:2	-0.02	C14:1	-0.014
13	Orn	-0.019	PC ae C36:2	0.014
14	LPC a C16:1	0.019	SM C24:1	-0.013
15	SM-OH C22:1	-0.019	LPC a C14:0	0.012
16	trans-OH-Pro	-0.017	C16	0.011
17	creatinine	0.016	PC aa C34:2	-0.011
18	C6	-0.016	C5:1	0.01
19	C14:1	-0.016	C8	-0.01
20	PC aa C32:0	0.016	LPC a C18:2	-0.01
21	PC aa C36:5	-0.016	Ala	-0.009
22	Cit	0.015	kynurenine	0.009
23	C16	0.015	Ile	-0.008
24	Ser	-0.014	Ac-Orn	0.008
25	SDMA	-0.014	alpha-AAA	0.008
26	LPC a C20:4	-0.014	LPC a C20:3	0.008
27	PC aa C32:2	-0.014	SM-OH C22:1	-0.008
28	C5:1	0.013	PC ae C34:2	-0.007
29	LPC a C18:1	-0.013	Tyr	0.006
30	Thr	0.012	C5-M-DC	-0.006
31	taurine	0.011	PC aa C32:2	-0.006
32	SM-OH C22:2	0.01	Arg	0.005
33	PC aa C36:4	-0.009	Thr	0.005
34	SM C18:1	-0.009	Val	-0.005
35	Lys	0.008	SDMA	-0.005
36	PC aa C30:0	-0.008	sarcosine	-0.005
37	PC aa C36:1	-0.008	C18	-0.005
38	PC ae C34:1	0.008	LPC a C16:1	0.005
39	SM C18:0	0.008	PC aa C36:5	-0.004
40	Asp	0.007	C0	0.003
41	Gln	0.007	Asp	0.002
42	alpha-AAA	0.007	Cit	0.002
43	C12	-0.007	Orn	-0.002
44	Leu	0.006	taurine	0.002
45	sarcosine	-0.006	SM-OH C22:2	0.002
46	C2	-0.006	hexose	-0.002
47	C18:2	-0.006	Ser	-0.001
48	Gly	-0.005	Asn	0

Table A.23 (continued)

	Milk		Total dairy products	
Order #	**Predictor**	**Coefficient**	**Predictor**	**Coefficient**
49	Trp	-0.005	Gln	0
50	kynurenine	0.005	Glu	0
51	PC ae C34:2	-0.005	Gly	0
52	Ala	-0.004	Leu	0
53	Glu	-0.004	Lys	0
54	Pro	-0.004	Met	0
55	C18	-0.004	Pro	0
56	Arg	0.003	Trp	0
57	Tyr	0.003	creatinine	0
58	Ac-Orn	0.003	Met-SO	0
59	serotonin	-0.003	serotonin	0
60	C4-OH	0.002	C2	0
61	SM C24:1	-0.002	C3	0
62	C4	-0.001	C4	0
63	C5-M-DC	-0.001	C4-OH	0
64	Asn	0	C6	0
65	Met	0	C10	0
66	Val	0	C12	0
67	Met-SO	0	C14	0
68	C3	0	LPC a C20:4	0
69	C8	0	LPC a C16:0	0
70	C10	0	LPC a C18:1	0
71	C14	0	PC aa C30:0	0
72	LPC a C14:0	0	PC aa C32:0	0
73	LPC a C16:0	0	PC aa C34:1	0
74	PC aa C32:1	0	PC aa C36:1	0
75	PC aa C34:1	0	PC aa C36:2	0
76	PC aa C34:2	0	PC aa C36:3	0
77	PC aa C36:2	0	PC aa C36:4	0
78	PC aa C36:3	0	PC aa C38:4	0
79	PC aa C38:4	0	PC aa C38:5	0
80	PC ae C36:1	0	PC ae C36:1	0
81	PC ae C36:2	0	SM C18:0	0
82	hexose	0	SM C18:1	0
	age	0.001	age	0.001
	sex	0	sex	0
	test site	0.042	test site	0
	smoking status	-0.004	smoking status	0
	physical activity	0	physical activity	0
	lipid-lowering	0	lipid-lowering	0
	HRT	-0.003	HRT	-0.028
	BMI	0.001	BMI	0.004
	intercept	0.753	intercept	1.229

*The metabolites are presented in order of decreasing absolute value of the elastic net coefficients, whereas the rest of the predictors used in the models are presented in the end, along with the model intercept

†Coefficients are derived from the main models, which include age, sex, test site, smoking status, physical activity, lipid lowering medication, hormone-replacement therapy, body mass index and the metabolites

Table A.24 Coefficients of predictors for total and types of dairy consumption as derived from logistic and linear prediction models without penalisation* in the discovery set of the Fenland study with the 82 overlapping metabolites with the diabetes case-cohort sub-study of the EPIC Norfolk study

Predictor	Milk	Yoghurt	Cheese	Butter	Total dairy products	Milk (continuous)	Total dairy products (continuous)
Ala	-143.198	-64.385	-222.506	-13.584	-180.01	-0.005	-0.01
Arg	102.863	-32.195	109.613	181.474	272.256	0.004	0.008
Asn	24.06	114.728	266.29	-66.098	-123.062	0.001	-0.002
Asp	185.378	-234.109	-25.095	-102.595	19.501	0.009	0.004
Cit	203.809	-96.813	-98.783	-25.375	54.178	0.019	0.005
Gln	274.874	-109.97	165.836	-62.224	-201.428	0.009	0.005
Glu	-145.642	-199.445	279.685	120.973	-20.849	-0.007	0.001
Gly	-136.595	-51.062	285.891	-23.52	-42.357	-0.007	-0.004
His	-135.65	285.559	60.437	-66.114	-147.783	-0.025	-0.021
Ile	-355.023	-251.954	-186.261	119.865	-179.133	-0.028	-0.009
Leu	200.616	29.299	39.634	-93.706	166.572	0.009	0
Lys	82.517	212.537	-36.287	-301.793	88.642	0.009	-0.003
Met	-119.524	151.609	-38.632	-70.469	157.095	-0.002	0.001
Orn	-186.071	-67.515	-65.998	233.687	-320.898	-0.02	-0.005
Phe	287.902	90.389	35.253	74.043	399.086	0.028	0.022
Pro	-141.136	-206.124	172.771	178.437	5.531	-0.006	0.002
Ser	-419.58	-101.194	314.448	41.986	-66.766	-0.016	-0.004
Thr	108.146	47.888	-124.655	-106.31	-38.739	0.015	0.008
Trp	173.413	73.394	288.358	100.846	155.143	-0.012	-0.008
Tyr	134.041	162.386	-96.591	239.259	112.402	0.008	0.014
Val	-149.039	-142.483	0.252	37.567	-113.412	0.003	-0.009
Ac-Orn	-6.873	107.434	24.685	-20.06	62.816	0.005	0.011
SDMA	-35.102	50.987	154.119	166.84	-106.907	-0.017	-0.009
alpha-AAA	165.264	132.012	-58.213	116.278	297.413	0.009	0.011
creatinine	268.785	-198.651	-365.098	-122.46	-57.686	0.018	0.001
kynurenine	57.473	309.143	424.232	0.471	289.037	0.006	0.012
Met-SO	59.006	19.278	-6.195	-22.181	-50.095	0	0.001
trans-OH-Pro	-449.838	-513.987	-629.298	118.232	-454.737	-0.017	-0.025
sarcosine	-56.738	-183.022	-1.957	13.748	-82.831	-0.007	-0.007
serotonin	-155.792	130.609	39.352	-76.589	48.062	-0.007	-0.003
taurine	132.586	301.053	-146.507	-184.242	-58.585	0.015	0.006
C0	213.612	4.828	-466.513	-151.6	41.214	0.032	0.01
C2	18.769	187.212	58.19	71.748	-162.489	-0.01	0.006
C3	220.007	79.017	-117.944	112.887	165.978	-0.004	-0.007
C4	-36.715	1.236	80.884	-90.735	118.254	-0.003	0.001
C4-OH	65.767	-73.698	-19.459	-140.085	-83.43	0.005	0.006
C5:1	207.869	236.072	305.677	-234.622	234.399	0.016	0.014
C5-M-DC	3.706	55.736	-275.6	-52.23	-85.592	-0.003	-0.01
C6	-79.542	-135.688	-58.323	81.349	-29.21	-0.017	-0.005
C8	-69.302	74.544	-251.173	7.656	-24.392	-0.01	-0.025
C10	-99.185	-71.754	7.051	39.404	52.412	0.016	0.023
C12	-69.399	-69.705	95.968	20.611	-18.649	-0.013	0.003
C14	102.44	114.726	91.443	-62.646	76.546	-0.003	-0.005
C14:1	-281.027	-241.69	-64.148	84.54	-197.973	-0.019	-0.028
C16	114.085	-222.309	76.383	186.718	179.984	0.021	0.022
C18	-157.098	-271.604	-114.565	212.189	66.967	-0.007	-0.012
C18:2	-223.682	291.298	-57.803	-430.611	-329.483	-0.008	-0.018
LPC a C14:0	56.283	195.572	248.358	396.581	405.686	0	0.023

Table A.24

Predictor	Milk	Yoghurt	Cheese	Butter	Total dairy products	Milk (continuous)	Total dairy products (continuous)
LPC a C16:1	97.78	261.993	115.53	-93.16	105.055	0.024	0.003
LPC a C17:0	345.707	274.542	458.102	631.874	742.721	0.032	0.076
LPC a C18:0	-233.945	-164.559	-340.655	-224.129	-354.051	-0.067	-0.097
LPC a C18:2	-138.403	10.938	-49.501	-110.814	-325.018	-0.011	-0.013
LPC a C20:4	-116.933	-184.359	-107.993	-0.758	-232.628	-0.027	-0.016
LPC a C16:0	-205.259	-4.856	-151.217	-264.12	-209.627	0.022	0.009
LPC a C18:1	-177.67	-109.484	-10.173	30.141	-249.642	-0.026	0.01
LPC a C20:3	435.634	127.195	13.56	-178.954	146.221	0.056	0.023
PC aa C30:0	22.171	153.285	297.543	142.761	301.5	-0.013	0.001
PC aa C32:0	79.444	50.057	-51.247	-530.323	-42.55	0.024	0.003
PC aa C32:1	91.947	-3.814	77.889	101.316	286.851	-0.004	0.017
PC aa C32:2	-36.134	54.99	109.41	40.356	106.996	-0.019	-0.023
PC aa C34:1	-101.395	-115.201	-52.799	63.029	-76.25	0.007	0.003
PC aa C34:2	-100.014	-163.653	-3.562	-162.961	-241.636	-0.001	-0.01
PC aa C36:1	-154.716	-329.106	-227.963	123.277	-162.61	-0.013	-0.012
PC aa C36:2	-123.584	-310.97	-249.56	-127.575	-442.529	0.013	0.009
PC aa C36:3	73.226	40.598	-3.917	-105.874	-75.227	-0.014	-0.012
PC aa C36:4	-289.719	-99.454	-24.099	-51.021	-211.335	-0.021	-0.004
PC aa C36:5	-106.71	166.754	29.327	-203.104	-182.154	-0.028	-0.021
PC aa C38:4	28.582	-242.468	-179.69	-104.857	-224.233	0.012	0.006
PC aa C38:5	107.769	26.643	-69.764	-183.344	-114.367	0.043	0.025
PC ae C34:1	151.597	25.273	166.8	358.062	373.816	0.009	0.013
PC ae C34:2	-18.566	-76.865	-110.281	217.646	103.343	-0.011	-0.013
PC ae C36:1	107.664	16.167	79.216	300.626	394.359	-0.005	-0.001
PC ae C36:2	115.588	93.876	428.374	399.754	523.118	0.005	0.028
SM C16:0	-162.654	-263.869	-160.722	-120.393	-379.806	-0.032	-0.044
SM C16:1	-339.491	57.665	-188.616	-195.507	-414.27	-0.055	-0.059
SM C18:0	245.664	13.04	-39.94	24.299	87.156	0.02	0.013
SM C18:1	-256.47	173.658	-191.415	82.302	-131.605	-0.022	-0.013
SM C24:1	-236.518	-191.344	-106.072	-376.263	-526.657	-0.003	-0.013
SM-OH C14:1	817.199	494.282	742.892	750.773	1,174.93	0.109	0.148
SM-OH C22:1	-172.573	-24.445	87.84	85.343	-32.746	-0.024	-0.015
SM-OH C22:2	243.775	138.598	184.16	72.305	243.677	0.013	0.006
hexose	125.309	-112.988	189.883	-179.92	-164.593	0.003	-0.006
age	-4.34	37.505	-68.865	-32.526	11.867	0.001	0.001
sex	-213.742	213.438	268.874	-47.195	-67.027	-0.067	0.029
test site	609.148	-154.403	-313.191	-519.754	47.965	0.044	0.002
smoking status	-99.732	-792.127	-9.686	186.147	-133.895	-0.006	0
physical activity	-14.39	-3.683	0.278	-2.883	7.42	0	0
lipid-lowering	-68.158	-89.035	41.098	-44.65	-9.875	-0.024	-0.027
HRT	207.045	-633.632	-432.651	104.984	104.299	-0.041	-0.02
BMI	6.672	30.794	142.298	-27.835	55.053	0	0.004
intercept	-22.918	-44.009	18.185	4.086	-7.308	0.941	1.17

*Coefficients are derived from the main models, which include age, sex, test site, smoking status, physical activity, lipid lowering medication, hormone-replacement therapy, body mass index and the metabolites

Table A.25 Coefficients of predictors in decreasing order of magnitude of the absolute values* for total and types of dairy consumption as derived from elastic net logistic prediction models † in the discovery set of the Fenland study with the 82 overlapping metabolites and the odd-chain saturated fatty acids with the diabetes case-cohort set nested within the EPIC Norfolk study

	Milk		Yoghurt		Cheese		Butter		Total dairy products	
Order #	**Predictor**	**Coefficient**	**Predictor**	**Coefficient**	**Predictor**	**Coefficient**	**Predictor**	**Coefficient**	**Predictor**	**Coefficient**
1	SM-OH C14:1	0.632	LPC a C18:2	8.603	trans-OH-Pro	-71,749.07	SM-OH C14:1	0.471	SM-OH C14:1	1.067
2	SM C16:1	-0.552	SM-OH C14:1	6.805	C15:0	67,831.95	C15:0	0.327	SM C16:1	-0.577
3	LPC a C20:3	0.253	LPC a C20:4	-6.783	C3	-61,920.04	LPC a C17:0	0.248	C15:0	0.429
4	LPC a C14:0	-0.229	C10	-6.245	PC ae C34:2	-54,000.45	PC ae C34:1	0.228	SM C16:0	-0.213
5	LPC a C16:1	0.213	C8	5.773	SM-OH C14:1	53,010.37	SM C16:1	-0.22	LPC a C20:3	0.195
6	LPC a C18:1	-0.172	LPC a C18:1	-5.489	PC ae C36:2	48,878.57	SM C16:0	-0.188	PC aa C34:2	-0.185
7	PC aa C38:5	0.169	PC aa C36:4	5.156	creatinine	-48,270.06	Tyr	0.164	Thr	-0.183
8	Val	-0.144	PC aa C32:0	4.801	PC aa C30:0	46,074.96	PC aa C32:0	-0.149	Val	-0.146
9	PC ae C34:1	0.144	PC aa C34:2	-4.647	LPC a C18:2	41,844.58	PC aa C38:4	-0.137	Tyr	0.14
10	PC aa C36:1	-0.142	PC aa C34:1	4.374	PC aa C32:1	37,692.12	Trp	-0.136	LPC a C18:0	-0.121
11	Tyr	0.13	PC aa C38:5	-4.331	C4	35,138.13	PC aa C34:1	0.135	trans-OH-Pro	-0.117
12	C14:1	-0.129	LPC a C18:0	4.266	SM C18:0	-34,310.24	LPC a C16:1	-0.126	LPC a C14:0	0.114
13	PC aa C32:2	-0.128	PC aa C36:1	-3.549	Val	34,003.46	Gly	-0.119	Ser	-0.112
14	PC aa C36:5	-0.12	PC aa C32:1	-3.41	LPC a C18:1	33,251.37	PC ae C36:2	0.105	serotonin	0.112
15	SM-OH C22:1	-0.12	PC aa C36:5	3.215	serotonin	32,937.41	Glu	0.103	Asn	0.111
16	creatinine	0.116	LPC a C16:0	-3.116	SDMA	31,737.29	kynurenine	0.097	PC ae C34:1	0.11
17	Cit	0.111	PC ae C34:2	-3.104	His	30,986.57	Ala	0.084	C5:1	0.102
18	C15:0	0.107	PC ae C34:1	-3.099	Gln	-29,505.14	Orn	0.083	Asp	0.1
19	Ala	-0.102	LPC a C14:0	3.036	C12	28,608.95	C18:2	-0.077	Ala	-0.096
20	hexose	-0.096	PC aa C36:3	2.923	Tyr	26,289.54	Asp	0.072	C16	0.087
21	Ser	-0.095	LPC a C16:1	2.893	PC aa C34:2	24,860.14	Lys	-0.068	PC aa C34:1	0.074
22	His	0.094	SM C16:1	-2.841	Pro	24,397.63	alpha-AAA	-0.068	C18:2	-0.058
23	PC aa C32:1	0.093	C2	2.566	SM C16:0	-24,062.28	PC aa C32:2	-0.061	PC ae C36:2	0.048
24	taurine	0.09	PC aa C32:2	-2.519	taurine	-24,018.25	taurine	-0.06	hexose	-0.047
25	trans-OH-Pro	-0.088	Ser	-2.27	PC aa C36:3	-23,974.41	LPC a C16:0	-0.059	Cit	0.046
26	Orn	-0.082	SM-OH C22:1	-2.245	PC aa C34:1	23,853.24	SDMA	0.056	PC aa C32:0	-0.046
27	LPC a C18:2	-0.078	C6	-2.164	C18	-23,406.09	PC aa C30:0	0.052	C18	-0.043
28	LPC a C17:0	0.077	SM C16:0	-2.054	SM C16:1	-22,943.23	Asn	-0.051	Ac-Orn	-0.041
29	Trp	0.074	Met	-1.883	LPC a C17:0	21,824.61	LPC a C14:0	0.049	C14:1	-0.033
30	Phe	0.073	Tyr	1.845	Ser	-21,468.81	PC ae C34:2	0.049	taurine	0.031
31	C18	-0.073	C15:0	1.808	C18:2	-21,122.65	C5:1	-0.047	SM-OH C22:2	0.026
32	PC aa C34:2	-0.073	C3	-1.713	C2	20,959.70	PC ae C36:1	0.046	kynurenine	0.025
33	Gly	0.072	Asn	1.711	Cit	19,049.50	Cit	-0.04	Leu	-0.021
34	C0	0.067	Phe	-1.702	Met-SO	18,692.43	Met	-0.04	alpha-AAA	0.021

Table A.25 (continued)

Order #	Milk Predictor	Coefficient	Yoghurt Predictor	Coefficient	Cheese Predictor	Coefficient	Butter Predictor	Coefficient	Total dairy products Predictor	Coefficient
35	PC aa C32:0	-0.067	C4	1.616	PC aa C38:5	-18,395.79	C12	0.04	SDMA	0.02
36	Lys	0.061	hexose	1.539	Phe	-17,674.40	hexose	-0.039	PC aa C36:2	-0.015
37	C2	-0.054	LPC a C20:3	1.531	PC ae C36:1	17,246.21	Met-SO	-0.033	PC aa C32:1	0.013
38	Asn	-0.052	Glu	-1.517	PC aa C32:2	-16,831.55	C14	0.033	Orn	0.009
39	C14	0.052	Met-SO	-1.478	Orn	-16,633.26	C18	0.031	C6	0.009
40	LPC a C16:0	-0.052	SM C18:0	-1.463	Lys	-16,371.61	C3	-0.03	PC aa C38:4	-0.006
41	PC aa C36:2	0.049	kynurenine	-1.436	LPC a C18:0	-15,978.06	creatinine	0.026	Lys	0.004
42	Met	-0.046	serotonin	1.296	C5-M-DC	-15,961.77	Ile	0.025	C17:0	0
43	PC aa C38:4	0.046	C5:1	1.267	C6	-15,196.75	SM-OH C22:1	0.022	Arg	0
44	C5:1	0.045	C0	1.23	PC aa C36:5	-15,114.23	C4	-0.021	Gln	0
45	Pro	-0.041	SM C24:1	-1.187	C5:1	14,779.42	PC aa C38:5	-0.021	Glu	0
46	SM C18:1	-0.041	LPC a C17:0	-1.171	LPC a C16:1	-14,235.89	Val	-0.015	Gly	0
47	Leu	0.04	PC ae C36:1	1.126	alpha-AAA	13,559.03	serotonin	0.013	His	0
48	Arg	-0.036	Lys	1.108	Arg	-13,149.49	C8	0.008	Ile	0
49	Ile	-0.035	SM-OH C22:2	1.085	LPC a C16:0	-12,704.90	SM C18:0	0.004	Met	0
50	C8	-0.035	C17:0	1.055	C0	-12,552.60	PC aa C36:4	-0.003	Phe	0
51	SDMA	-0.034	Arg	-1.019	C14:1	-12,533.92	C17:0	0	Pro	0
52	PC aa C36:3	0.033	Ile	-1.013	Glu	-11,912.76	Arg	0	Trp	0
53	SM-OH C22:2	0.032	PC aa C38:4	0.946	Gly	11,871.15	Gln	0	creatinine	0
54	kynurenine	0.031	alpha-AAA	0.94	C10	11,848.89	His	0	Met-SO	0
55	C3	0.03	Ala	-0.919	Asn	11,812.71	Leu	0	sarcosine	0
56	Glu	-0.028	SM C18:1	0.877	C16	-11,785.66	Phe	0	C0	0
57	Thr	-0.027	Gly	0.826	LPC a C20:4	11,756.29	Pro	0	C2	0
58	C4-OH	0.025	PC ae C36:2	-0.822	PC aa C36:1	-11,735.54	Ser	0	C3	0
59	C5-M-DC	0.025	Thr	0.808	Ala	-11,515.33	Thr	0	C4	0
60	serotonin	0.024	SDMA	0.783	Ac-Orn	10,165.52	Ac-Orn	0	C4-OH	0
61	PC ae C36:1	-0.024	C18:2	-0.764	C14	-10,066.51	trans-OH-Pro	0	C5-M-DC	0
62	Gln	0.022	C12	-0.746	C17:0	-9,451.06	sarcosine	0	C8	0
63	Ac-Orn	-0.022	Leu	0.711	PC aa C36:4	8,976.24	C0	0	C10	0
64	C18:2	0.019	His	0.688	SM-OH C22:1	-8,401.53	C2	0	C12	0
65	PC aa C30:0	-0.017	C4-OH	0.683	SM-OH C22:2	7,860.08	C4-OH	0	C14	0
66	C4	0.016	C14:1	0.683	Met	-7,082.99	C5-M-DC	0	LPC a C16:1	0
67	Asp	0.011	Cit	-0.656	PC aa C32:0	6,766.54	C6	0	LPC a C17:0	0
68	C6	-0.008	Gln	-0.6	SM C18:1	5,603.44	C10	0	LPC a C18:2	0

Table A.25 (continued)

Order #	Milk Predictor	Coefficient	Yoghurt Predictor	Coefficient	Cheese Predictor	Coefficient	Butter Predictor	Coefficient	Total dairy products Predictor	Coefficient
69	C10	-0.006	sarcosine	-0.57	kynurenine	-5,558.14	C14:1	0	LPC a C20:4	0
70	C17:0	0	Ac-Orn	0.551	LPC a C20:3	-5,316.29	C16	0	LPC a C16:0	0
71	alpha-AAA	0	PC aa C36:2	-0.529	PC aa C38:4	-5,095.60	LPC a C18:0	0	LPC a C18:1	0
72	Met-SO	0	C5-M-DC	0.42	Thr	-4,502.30	LPC a C18:2	0	PC aa C30:0	0
73	sarcosine	0	trans-OH-Pro	-0.414	Trp	4,231.58	LPC a C20:4	0	PC aa C32:2	0
74	C12	0	C16	-0.36	Ile	3,776.47	LPC a C18:1	0	PC aa C36:1	0
75	C16	0	Asp	-0.286	sarcosine	-3,435.01	LPC a C20:3	0	PC aa C36:3	0
76	LPC a C18:0	0	creatinine	-0.253	PC aa C36:2	-2,545.58	PC aa C32:1	0	PC aa C36:4	0
77	LPC a C20:4	0	Val	-0.105	Asp	2,154.18	PC aa C34:2	0	PC aa C36:5	0
78	PC aa C34:1	0	C18	0.067	Leu	-1,537.65	PC aa C36:1	0	PC aa C38:5	0
79	PC aa C36:4	0	Orn	0	PC ae C34:1	1,015.68	PC aa C36:2	0	PC ae C34:2	0
80	PC ae C34:2	0	Pro	0	hexose	-817.477	PC aa C36:3	0	PC ae C36:1	0
81	PC ae C36:2	0	Trp	0	C8	509.189	PC aa C36:5	0	SM C18:0	0
82	SM C16:0	0	taurine	0	C4-OH	-509.127	SM C18:1	0	SM C18:1	0
83	SM C18:0	0	C14	0	LPC a C14:0	444.34	SM C24:1	0	SM C24:1	0
84	SM C24:1	0	PC aa C30:0	0	SM C24:1	-224.814	SM-OH C22:2	0	SM-OH C22:1	0
	age	0.026	age	0	age	-4,935.93	age	0	age	0.007
	sex	0	sex	8.776	sex	41,210.73	sex	0	sex	0
	test site	0.193	test site	0.602	test site	-66,489.21	test site	-0.194	test site	0
	smoking status	0.031	smoking status	-2.66	smoking status	11,028.38	smoking status	0.151	smoking status	0
	physical activity	0.009	physical activity	0	physical activity	1,604.98	physical activity	0	physical activity	0.008
	lipid-lowering	-0.087	lipid-lowering	-0.232	lipid-lowering	14,992.73	lipid-lowering	-0.044	lipid-lowering	0
	HRT	0.104	HRT	-0.372	HRT	-78,063.18	HRT	0	HRT	0
	BMI	0.051	BMI	0	BMI	5,585.89	BMI	0.019	BMI	0.055
	intercept	-3.675	intercept	-24.322	intercept	350.666	intercept	-1.612	intercept	-0.665

*The metabolites are presented in order of decreasing absolute value of the elastic net coefficients, whereas the rest of the predictors used in the models are presented in the end, along with the model intercept

†Coefficients are derived from the main models, which include age, sex, test site, smoking status, physical activity, lipid lowering medication, hormone-replacement therapy, body mass index and the metabolites

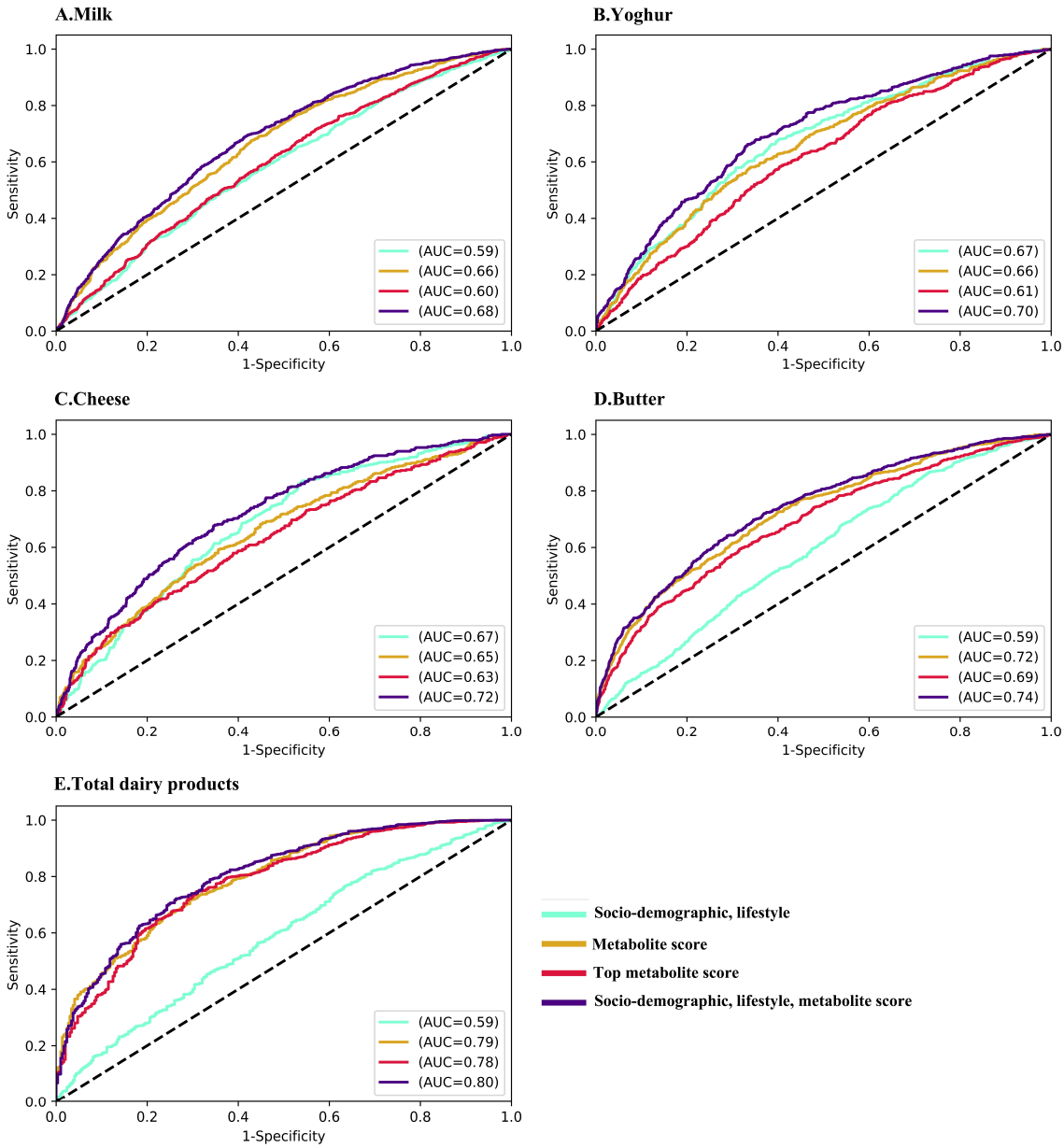


Fig. A.1 Receiver operating characteristic (ROC) curves and corresponding areas under the curve (AUC) for prediction models of metabolite scores reflecting (A) milk, (B) yoghurt, (C) cheese, (D) butter and (E) total dairy consumption in the internal validation set (n=4,246) using all the 174 metabolites. Socio-demographic, lifestyle model: age, sex, test site, smoking status, physical activity, lipid lowering medication, hormone-replacement therapy and body mass index; Top metabolite score: metabolites with coefficients with absolute values >mean+2*SD

Table A.26 Coefficients of predictors in decreasing order of magnitude of the absolute values* for milk and total dairy consumption as derived from elastic net linear prediction models † in the discovery set of the Fenland study with the 82 overlapping metabolites and the odd-chain saturated fatty acids with the diabetes case-cohort set nested within the EPIC Norfolk study

	Milk		Total dairy products	
Order #	Predictor	Coefficient	Predictor	Coefficient
1	SM-OH C14:1	0.108	SM-OH C14:1	0.162
2	SM C16:1	-0.077	SM C16:1	-0.091
3	LPC a C16:1	0.054	C15:0	0.045
4	LPC a C18:2	-0.037	LPC a C16:1	0.041
5	SM-OH C22:1	-0.023	LPC a C18:0	-0.034
6	LPC a C14:0	-0.021	PC aa C32:2	-0.028
7	creatinine	0.019	SM C16:0	-0.028
8	LPC a C20:3	0.018	LPC a C17:0	0.023
9	LPC a C18:0	-0.017	trans-OH-Pro	-0.019
10	Ala	-0.013	PC aa C34:1	0.018
11	Lys	0.012	Tyr	0.014
12	taurine	0.012	PC aa C30:0	0.013
13	C14:1	-0.012	Ser	-0.011
14	PC aa C34:2	-0.011	C5:1	0.011
15	Tyr	0.01	LPC a C16:0	-0.009
16	C5:1	0.01	PC aa C32:0	-0.008
17	Val	-0.009	SM-OH C22:1	-0.008
18	trans-OH-Pro	-0.009	hexose	-0.008
19	PC aa C32:2	-0.009	C14:1	-0.007
20	PC aa C38:4	0.009	PC aa C32:1	0.007
21	Orn	-0.008	PC aa C36:1	-0.007
22	Phe	0.008	creatinine	0.006
23	SDMA	-0.008	serotonin	0.006
24	hexose	-0.008	taurine	0.006
25	Asn	-0.007	LPC a C18:2	-0.006
26	Cit	0.007	PC ae C36:2	0.006
27	C0	0.007	Leu	-0.005
28	PC aa C32:0	-0.007	C3	-0.005
29	Gly	0.006	C18	-0.005
30	Pro	-0.006	PC aa C34:2	-0.005
31	C2	-0.006	Thr	-0.004
32	SM-OH C22:2	0.006	Ala	-0.003
33	C15:0	0.005	Asp	0.003
34	Glu	-0.005	SM C24:1	-0.003
35	Trp	0.004	SM-OH C22:2	0.003
36	serotonin	0.004	Glu	-0.002
37	C18:2	0.004	Phe	0.002
38	LPC a C17:0	0.004	sarcosine	0.002
39	Ser	-0.003	C18:2	-0.002
40	Thr	-0.003	SM C18:1	0.002
41	alpha-AAA	0.003	Asn	-0.001
42	C8	-0.003	Gln	-0.001
43	C18	-0.003	Ac-Orn	0.001
44	PC aa C36:5	-0.003	C0	0.001
45	PC ae C34:1	0.003	C4-OH	-0.001
46	Met-SO	0.002	PC aa C36:4	0.001
47	C3	0.002	C17:0	0
48	PC aa C36:1	-0.002	Arg	0

Table A.26 (continued)

	Milk		Total dairy products	
Order #	**Predictor**	**Coefficient**	**Predictor**	**Coefficient**
49	SM C18:0	-0.002	Cit	0
50	His	-0.001	Gly	0
51	Met	0.001	His	0
52	PC aa C38:5	0.001	Ile	0
53	C17:0	0	Lys	0
54	Arg	0	Met	0
55	Asp	0	Orn	0
56	Gln	0	Pro	0
57	Ile	0	Trp	0
58	Leu	0	Val	0
59	Ac-Orn	0	SDMA	0
60	kynurenine	0	alpha-AAA	0
61	sarcosine	0	kynurenine	0
62	C4	0	Met-SO	0
63	C4-OH	0	C2	0
64	C5-M-DC	0	C4	0
65	C6	0	C5-M-DC	0
66	C10	0	C6	0
67	C12	0	C8	0
68	C14	0	C10	0
69	C16	0	C12	0
70	LPC a C20:4	0	C14	0
71	LPC a C16:0	0	C16	0
72	LPC a C18:1	0	LPC a C14:0	0
73	PC aa C30:0	0	LPC a C20:4	0
74	PC aa C32:1	0	LPC a C18:1	0
75	PC aa C34:1	0	LPC a C20:3	0
76	PC aa C36:2	0	PC aa C36:2	0
77	PC aa C36:3	0	PC aa C36:3	0
78	PC aa C36:4	0	PC aa C36:5	0
79	PC ae C34:2	0	PC aa C38:4	0
80	PC ae C36:1	0	PC aa C38:5	0
81	PC ae C36:2	0	PC ae C34:1	0
82	SM C16:0	0	PC ae C34:2	0
83	SM C18:1	0	PC ae C36:1	0
84	SM C24:1	0	SM C18:0	0
	age	0.001	age	0.001
	sex	0	sex	0
	test site	0.027	test site	-0.006
	smoking status	0.005	smoking status	0.012
	physical activity	0	physical activity	0
	lipid-lowering	0	lipid-lowering	0
	HRT	0	HRT	-0.034
	BMI	0.002	BMI	0.005
	intercept	0.727	intercept	1.182

*The metabolites are presented in order of decreasing absolute value of the elastic net coefficients, whereas the rest of the predictors used in the models are presented in the end, along with the model intercept
†Coefficients are derived from the main models, which include age, sex, test site, smoking status, physical activity, lipid lowering medication, hormone-replacement therapy, body mass index and the metabolites

Table A.27 Coefficients of predictors for total and types of dairy consumption as derived from logistic and linear prediction models without penalisation* in the discovery set of the Fenland study with the 82 overlapping metabolites and the odd-chain saturated fatty acids with the diabetes case-cohort set nested within the EPIC Norfolk study

Predictor	Milk	Yoghurt	Cheese	Butter	Total dairy products	Milk (continuous)	Total dairy products (continuous)
C15:0	443.509	328.304	959.961	744.317	1,027.88	-0.001	0.039
C17:0	82.421	297.792	202.873	88.382	284.113	-0.002	-0.004
Ala	-271.581	55.321	-27.927	176.974	-204.174	-0.014	-0.003
Arg	-93.496	-151.176	-193.711	90.311	-80.403	-0.004	-0.004
Asn	-164.492	272.711	237.779	-257.337	142.176	-0.012	-0.004
Asp	210.722	-115.677	-32.838	254.939	327.843	0.004	0.005
Cit	283.941	-53.671	224.898	-120.623	168.491	0.013	0.002
Gln	126.655	-60.888	-241.004	-2.555	-94.849	0.003	-0.003
Glu	9.352	-257.356	-80.664	250.159	90.604	-0.011	-0.009
Gly	10.569	215.733	229.914	-290.172	-173.758	0.011	0
His	141	128.964	340.011	-38.737	-52.426	-0.008	-0.001
Ile	-126.546	-267.954	-112.663	91.123	-111.962	0.021	0.021
Leu	-14.68	-35.655	-148.194	-10.837	-137.034	0	-0.011
Lys	189.899	215.418	-302.813	-76.33	84.847	0.019	0.008
Met	-70.894	-212.273	5.316	-129.018	40.514	0.002	-0.009
Orn	-19.895	-172.792	-141.811	112.895	48.146	-0.015	0
Phe	220.827	-95.612	6.095	22.108	62.792	0.009	0.01
Pro	-291.827	-76.989	66.065	83.365	33.937	-0.008	0
Ser	-213.515	-124.87	-58.359	-112.416	-264.003	-0.009	-0.013
Thr	-137.723	141.796	94.078	53.264	-446.445	-0.006	-0.007
Trp	186.324	8.644	100.964	-166.863	135.661	0.007	-0.002
Tyr	374.589	251.808	147.439	251.673	235.258	0.018	0.023
Val	-239.251	-70.101	2.299	-107.552	-269.057	-0.029	-0.014
Ac-Orn	-146.654	178.755	117.933	35.782	-152.845	0	0.003
SDMA	15.088	149.699	72.049	206.072	-40.048	-0.014	-0.003
alpha-AAA	8.853	151.22	-26.065	-44.684	75.915	0.006	0.004
creatinine	313.395	-176.43	-451.026	106.048	77.394	0.026	0.014
kynurenine	245.017	-81.755	-7.354	206.474	137.66	-0.002	0.001
Met-SO	-65.114	-176.48	253.273	-168.553	-15.228	0.005	-0.001
trans-OH-Pro	-207.047	-80.291	-831.721	-5.997	-488.35	-0.013	-0.023
sarcosine	-25.742	111.835	105.101	-16.274	97.756	0.001	0.004
serotonin	20.651	420.964	298.005	-3.975	302.175	0.006	0.009
taurine	142.344	217.022	-267.702	-197.406	181.212	0.015	0.007
C0	58.644	110.72	-348.054	29.727	2.406	0.015	0.018
C2	-189.338	458.488	177.293	-19.166	9.646	-0.014	0
C3	78.535	24.581	-322.924	-64.308	58.143	0.001	-0.02
C4	13.803	333.057	292.267	-40.734	38.666	0.002	0.004
C4-OH	115.855	152.11	-174.505	-57.758	-9.273	0.002	-0.006
C5:1	111.504	159.056	63.937	-233.919	343.3	0.012	0.015
C5-M-DC	128.979	169.61	-74.594	-4.951	-45.817	-0.003	0
C6	8.41	-68.323	-136.294	-31.096	87.185	0.011	0
C8	-21.288	21.06	59.123	90.258	-12.881	-0.013	0.008
C10	-29.921	-112.19	73.522	46.566	51.98	-0.002	-0.017
C12	-101.84	-118.581	123.157	71.429	48.104	0.004	0.016
C14	13.207	-18.361	-17.051	91.612	97.219	-0.003	0.001
C14:1	-185.475	-40.923	-77.054	-121.889	-236.972	-0.023	-0.022

Table A.27 (continued)

Predictor	Milk	Yoghurt	Cheese	Butter	Total dairy products	Milk (continuous)	Total dairy products (continuous)
C16	22.23	-37.394	-75.075	9.309	220.947	0.014	0.01
C18	-138.952	-345.413	-104.466	276.222	-127.191	-0.012	-0.014
C18:2	-104.265	-71.765	-232.925	-261.754	-356.565	0.013	0
LPC a C14:0	-289.231	321.66	379.965	216.924	343.797	-0.038	0.002
LPC a C16:1	398.396	205.315	125.539	-215.482	126.214	0.076	0.045
LPC a C17:0	221.248	182.669	264.079	459.257	515.653	0.031	0.039
LPC a C18:0	-105.151	-16.45	-122.539	-124.504	-185.042	-0.039	-0.023
LPC a C18:2	-301.503	-51.143	175.266	-14.818	-154.686	-0.049	-0.014
LPC a C20:4	69.177	-161.103	-138.501	-138.073	12.114	-0.007	-0.017
LPC a C16:0	-150.954	36.484	-48.713	-151.12	-133.881	-0.016	-0.049
LPC a C18:1	-143.254	-121.772	160.814	66.655	69.307	0.004	0.031
LPC a C20:3	411.744	115.281	-10.041	-5.891	232.816	0.034	0.004
PC aa C30:0	-183.54	198.699	651.294	160.563	293.378	0.016	0.027
PC aa C32:0	-117.383	34.1	-3.259	-333.213	-200.051	-0.021	-0.02
PC aa C32:1	266.229	9.075	310.469	27.636	143.165	-0.024	0.005
PC aa C32:2	-191.157	-13.635	181.429	-11.19	2.195	-0.009	-0.044
PC aa C34:1	7.549	-179.861	79.568	76.592	37.707	0.012	0.035
PC aa C34:2	-297.787	-198.348	-1.19	-27.635	-410.006	-0.008	-0.008
PC aa C36:1	-167.039	-319.803	-49.372	23.865	-130.47	-0.014	-0.033
PC aa C36:2	-119.776	-120.775	-164.997	0.712	-343.533	0.016	0.006
PC aa C36:3	173.691	106.966	-173.268	-65.535	-82.697	-0.009	0
PC aa C36:4	57.851	79.675	-189.691	-150.504	-219.718	0.005	0.023
PC aa C36:5	-284.475	203.446	15.08	-69.217	-91.577	-0.018	-0.009
PC aa C38:4	208.445	-57.73	-327.062	-271.348	-255.354	0.012	0.001
PC aa C38:5	159.641	-45.419	-101.684	-149.425	-78.466	0.026	0.008
PC ae C34:1	308.353	-160.497	71.32	416.825	399.795	0.012	-0.001
PC ae C34:2	-124.343	-220.531	-320.647	148.243	-24.179	-0.001	-0.006
PC ae C36:1	75.656	7.887	97.735	404.559	183.971	-0.002	0.002
PC ae C36:2	134.963	-14.916	274.035	438.122	334.658	-0.015	0.01
SM C16:0	-50.588	-151.915	-95.85	-308.129	-231.74	0.012	-0.024
SM C16:1	-382.357	55.32	-150.801	-392.819	-415.522	-0.104	-0.119
SM C18:0	41.742	24.843	-301.395	134.233	76.575	-0.026	-0.02
SM C18:1	-180.361	199.658	-28.895	-1.177	-138.53	0.019	0.026
SM C24:1	-202.582	-225.907	79.561	4.045	-161.269	0.003	-0.003
SM-OH C14:1	837.69	453.076	459.712	615.332	1,018.34	0.127	0.18
SM-OH C22:1	-140.431	-279.242	-1.234	135.262	56.4	-0.032	-0.017
SM-OH C22:2	66.914	257.571	12.264	-3.195	313.29	0.01	0.006
hexose	-240.446	136.508	-136.433	-109.164	-132.01	-0.012	-0.009
age	40.018	9.26	-47.134	-35.111	54.457	0.001	0.001
sex	-313.638	357.837	355.415	-183.808	-105.009	0.073	0.205
test site	493.899	-149.587	-667.295	-425.189	-122.968	0.028	-0.011
smoking status	-5.387	-676.405	25.127	253.235	29.993	0.01	0.015
physical activity	57.504	-40.73	-60.443	48.692	62.196	0	0
lipid-lowering	-94.309	27.962	14.843	-118.142	-33.692	-0.051	-0.041
HRT	186.497	-837.976	-691.876	58.152	2.391	0.03	0.06
BMI	-19.891	-5.402	71.992	31.235	123.081	0.002	0.005
intercept	-84.536	-25.941	1.017	-55.642	-42.235	0.54	0.686

*Coefficients are derived from the main models, which include age, sex, test site, smoking status, physical activity, lipid lowering medication, hormone-replacement therapy, body mass index and the metabolites

Table A.28 Areas under the curve in the discovery and internal validation sets for the different dairy prediction models

	Discovery set	Internal validation set
N	6,035	4,246
Milk		
Socio-demographic, lifestyle factors*	0.59	0.59
Metabolite score †	0.71	0.66
Top metabolite score ‡	0.61	0.61
Socio-demographic, lifestyle factors, metabolite score*	0.73	0.68
Socio-demographic, lifestyle factors, top metabolite score* †	0.64	0.63
Yoghurt		
Socio-demographic, lifestyle factors*	0.67	0.67
Metabolite score †	0.7	0.66
Top metabolite score ‡	0.63	0.62
Socio-demographic, lifestyle factors, metabolite score*	0.72	0.7
Socio-demographic, lifestyle factors, top metabolite score* †	0.69	0.68
Cheese		
Socio-demographic, lifestyle factors*	0.62	0.67
Metabolite score †	0.73	0.69
Top metabolite score ‡	0.67	0.65
Socio-demographic, lifestyle factors, metabolite score*	0.75	0.73
Socio-demographic, lifestyle factors, top metabolite score* †	0.69	0.71
Butter		
Socio-demographic, lifestyle factors*	0.6	0.59
Metabolite score †	0.73	0.72
Top metabolite score ‡	0.68	0.67
Socio-demographic, lifestyle factors, metabolite score*	0.74	0.73
Socio-demographic, lifestyle factors, top metabolite score* †	0.7	0.69
Total dairy products		
Socio-demographic, lifestyle factors*	0.56	0.59
Metabolite score †	0.83	0.79
Top metabolite score ‡	0.78	0.77
Socio-demographic, lifestyle factors, metabolite score*	0.83	0.81
Socio-demographic, lifestyle factors, top metabolite score* †	0.79	0.78

*Socio-demographic and lifestyle factors include age, sex, test site, smoking status, physical activity, lipid lowering medication, hormone-replacement therapy and body mass index

†Metabolite score created from the 82 metabolites overlapping with the metabolites of the external validation set

‡Top metabolite scores were derived from the top mean+2*SD of the absolute values of the elastic net coefficients from the main models, which included age, sex, test site, smoking status, physical activity, lipid lowering medication, hormone-replacement therapy, body mass index and the metabolites

Table A.29 Areas under the curve (AUC) for models of logistic regression of the dairy outcomes against the metabolite scores adjusted for socio-demographic and lifestyle factors* in the internal and external validation sets (odd-chain saturated fatty acid subsample)

Table 2. Areas under the curve (AUC) for models of logistic regression of the dairy outcomes against the metabolite scores adjusted for socio-demographic and lifestyle factors* in the internal and external validation sets (odd-chain fatty acid subsample)

			Internal validation set			External validation set			
				Metabolite score †			Metabolite score †		
				No	Yes		No	Yes	
Dairy outcome	**Metabolite set ‡**	**Number of metabolites**	**N**	**AUC**	**AUC**	**N**	**AUC**	**AUC**	**p §**
Milk	overlapping ‖	82	2,758	0.59	0.68	362	0.61	0.63	0.01
	top §ignals from overlapping #	3		0.59	0.63	362	0.61	0.63	0.006
	overlapping, odd-chain fatty acids	84		0.59	0.72	362	0.61	0.62	0.14
	top Signals from overlapping, odd-chain fatty acids**	2		0.59	0.66	362	0.61	0.65	0.005
Yoghurt	overlapping ‖	82	2,089	0.67	0.7	391	0.7	0.7	0.71
	top §ignals from overlapping #	2		0.67	0.68	391	0.7	0.7	0.99
	overlapping, odd-chain fatty acids	84		0.67	0.75	391	0.7	0.72	0.94
	top §ignals from overlapping, odd-chain fatty acids**	5		0.67	0.68	391	0.7	0.72	0.98
Cheese	overlapping ‖	82	1,574	0.67	0.73	194	0.7	0.7	0.65
	top §ignals from overlapping #	4		0.67	0.71	194	0.7	0.7	0.56
	overlapping, odd-chain fatty acids	84		0.67	0.79	194	0.7	0.7	0.94
	top §ignals from overlapping, odd-chain fatty acids**	3		0.67	0.72	194	0.7	0.7	0.36
Butter	overlapping ‖	82	3,066	0.59	0.73	515	0.62	0.63	0.03
	top Signals from overlapping #	5		0.59	0.69	515	0.62	0.63	0.34
	overlapping, odd-chain fatty acids	84		0.59	0.76	515	0.62	0.64	0.05
	top §ignals from overlapping, odd-chain fatty acids**	5		0.59	0.72	515	0.62	0.66	0.001
Total dairy products	overlapping ‖	82	1,969	0.59	0.81	282	0.59	0.67	0.08
	top §ignals from overlapping #	4		0.59	0.78	282	0.59	0.71	0.006
	overlapping, odd-chain fatty acids	84		0.59	0.84	282	0.59	0.74	0.01
	top §ignals from overlapping, odd-chain fatty acids**	3		0.59	0.8	282	0.59	0.77	0.001

*Socio-demographic and lifestyle factors include age, sex, test site, smoking status, physical activity, lipid lowering medication, hormone-replacement therapy and body mass index

†The metabolite score-No model includes the socio-demographic and lifestyle factors; the metabolite score-Yes model includes the socio-demographic and lifestyle factors as well as the metabolite score

‡The metabolite set used in the discovery analysis to derive the metabolite scores

§*p* values of the coefficients of the metabolite score

parllelMetabolites of the discovery (targeted) set approximately matched with metabolites from the external validation (untargeted) set

#Top Signals defined as metabolites with absolute values of coefficients >mean+2*SD

**Top Signals derived from the discovery models which included the overlapping metabolites and the odd-chain fatty acids

Table A.30 Areas under the curve of the logistic models* with or without the fatty acids C15:0, C17:0, trans-16:1n-7 and the metabolite score (174 metabolites) for total and types of dairy products

Metabolite score				**Top metabolite score †**			
Milk		Fatty acids		**Milk**			Fatty acids
		No	Yes			No	Yes
Metabolite score	No	0.59	0.62	Metabolite score	No	0.59	0.62
	Yes	0.68	0.68		Yes	0.66	0.65
Yoghurt		Fatty acids		**Yoghurt**			Fatty acids
		No	Yes			No	Yes
Metabolite score	No	0.67	0.69	Metabolite score	No	0.67	0.69
	Yes	0.7	0.71		Yes	0.7	0.7
Cheese		Fatty acids		**Cheese**			Fatty acids
		No	Yes			No	Yes
Metabolite score	No	0.67	0.72	Metabolite score	No	0.67	0.72
	Yes	0.72	0.73		Yes	0.73	0.72
Butter		Fatty acids		**Butter**			Fatty acids
		No	Yes			No	Yes
Metabolite score	No	0.59	0.7	Metabolite score	No	0.59	0.7
	Yes	0.74	0.74		Yes	0.73	0.72
Total dairy products		Fatty acids		**Total dairy products**			Fatty acids
		No	Yes			No	Yes
Metabolite score	No	0.59	0.75	Metabolite score	No	0.59	0.75
	Yes	0.8	0.81		Yes	0.8	0.8

*All models include age, sex, test site, smoking status, physical activity, lipid lowering medication, hormone-replacement therapy and body mass index

†The top metabolite score includes only the top metabolite signals defined as the metabolites with absolute values of coefficients >mean+2*SD

Table A.31 Areas under the curve of the logistic models* with or without the odd-chain saturated fatty acids C15:0, C17:0 and the metabolite score (82 metabolites †) for total and types of dairy products in the external validation set of the EPIC Norfolk diabetes case-cohort study)

Metabolite score				**Top metabolite score ‡**			
Milk		Fatty acids		**Milk**		Fatty acids	
		No	Yes			No	Yes
Metabolite score	No	0.61	0.62	Metabolite score	No	0.61	0.62
	Yes	0.63	0.63		Yes	0.63	0.64
Yoghurt		Fatty acids		**Yoghurt**		Fatty acids	
		No	Yes			No	Yes
Metabolite score	No	0.70	0.70	Metabolite score	No	0.70	0.70
	Yes	0.70	0.70		Yes	0.70	0.70
Cheese		Fatty acids		**Cheese**		Fatty acids	
		No	Yes			No	Yes
Metabolite score	No	0.70	0.71	Metabolite score	No	0.70	0.71
	Yes	0.70	0.71		Yes	0.70	0.70
Butter		Fatty acids		**Butter**		Fatty acids	
		No	Yes			No	Yes
Metabolite score	No	0.62	0.77	Metabolite score	No	0.62	0.77
	Yes	0.63	0.78		Yes	0.63	0.77
Total dairy products		Fatty acids		**Total dairy products**		Fatty acids	
		No	Yes			No	Yes
Metabolite score	No	0.59	0.76	Metabolite score	No	0.59	0.76
	Yes	0.67	0.76		Yes	0.71	0.77

*All models include age, sex, test site, smoking status, physical activity, lipid lowering medication, hormone-replacement therapy and body mass index

†Metabolites approximately matched between the discovery (targeted) set and the external validation (untargeted) set

‡The top metabolite score includes only the top metabolite signals defined as the metabolites with absolute values of coefficients >mean+2*SD

Table A.32 Change in the ranking of top metabolite signals (>absolute mean+2SD) after adjusting for dietary factors in the set of 174 metabolites

Order # main model	Metabolite	Order # with diet	Difference
Milk			
1	SM(OH) C141	5	4
2	SM C161	34	32
3	PC aa C281	12	9
4	C141	11	7
5	Ile	120	115
6	LPCa C180	49	43
7	LPCa C203	21	14
Yoghurt			
1	c4OHPro	15	14
2	SM(OH) C141	2	0
3	SM C240	11	8
4	SM C160	90	86
5	PC aa C281	4	1
6	taurine	1	5
7	PC ae C404	94	87
8	C9	27	19
Cheese			
1	SM(OH) C141	2	1
2	PC ae C364	28	26
3	C0	21	18
4	PC aa C424	30	26
5	Trp	96	91
6	LPC a C170	32	26
7	C51	57	50
8	PC ae C362	10	2
9	LPC a C180	13	4
10	C5MDC	112	102
Butter			
1	SM(OH) C141	1	0
2	PC aa C320	40	38
3	C9	5	2
4	PC ae C341	9	5
5	LPC a C170	6	1
6	LPC a C140	15	9
7	SM(OH) C161	2	5
8	PC ae C340	4	4
9	C102	48	39
Total dairy products			
1	SM(OH) C141	1	0
2	SM C161	163	161
3	LPC a C180	35	32
4	LPC a C170	9	5
5	SM C241	18	13

Table A.33 Change in the ranking of top metabolite signals (>absolute mean+2SD) after adjusting for dietary factors in the set of 82 overlapping metabolites

Order # main model	Metabolite	Order # with diet	Difference
	Milk		
1	SM(OH) C141	5	-
2	LPCa C203	34	-
3	SM C161	12	-
	Yoghurt		
1	SM(OH) C141	1	0
2	t4OHPro	14	12
3	C18	16	13
4	LPC a C170	42	38
	Cheese		
1	SM(OH) C141	1	0
2	t4OHPro	28	26
3	LPC a C180	21	18
4	C0	30	26
	Butter		
1	SM(OH) C141	1	0
2	PC ae C341	5	3
3	LPC a C170	4	1
4	LPC a C140	7	3
5	PC aa C320	70	65
	Total dairy products		
1	SM(OH) C141	1	0
2	SM C161	163	161
3	LPC a C180	35	32
4	LPC a C170	9	5

A.5 Chapter 8

Table A.34 Associations of top hits from the genome-wide association studies on total dairy products, milk, yoghurt and cheese with all dairy phenotypes in the UK Biobank

SNP	Dairy phenotype	Chromosome	Base position	Gene position	Effect allele	Minor allele	EAF	INFO †	b ‡	se	*p*
rs1222762	Total dairy products §	1	97029325	1.26566	G	A	0.432208	0.996402	-0.01	0.003	0.08
	Milk as a drink and added to cereals ‖	1	97029325	1.26566	G	A	0.433602	0.996402	-0.001	0.001	0.06
	Milk added to coffee, tea or cereals ‖	1	97029325	1.26566	G	A	0.432208	0.996402	-0.001	0.001	0.38
	Milk as drink ‖	1	97029325	1.26566	G	A	0.432208	0.996402	0.001	0.001	0.5
	Yoghurt ‖	1	97029325	1.26566	G	A	0.432208	0.996402	-0.001	0.002	0.35
	Cheese ‖	1	97029325	1.26566	G	A	0.433588	0.996402	-0.002	0.0004	$1.2x10^{-8}$
rs12409187	Total dairy products §	1	4733793	0.0979031	G	A	0.953013	0.964766	0.04	0.01	$4.7x10^{-9}$
	Milk as a drink and added to cereals ‖	1	4733793	0.0979031	G	A	0.952941	0.964766	0.003	0.001	0.02
	Milk added to coffee, tea or cereals ‖	1	4733793	0.0979031	G	A	0.953013	0.964766	0.01	0.002	$7.5x10^{-7}$
	Milk as drink ‖	1	4733793	0.0979031	G	A	0.953013	0.964766	-0.001	0.002	0.55
	Yoghurt ‖	1	4733793	0.0979031	G	A	0.953013	0.964766	0.01	0.004	0.06
	Cheese ‖	1	4733793	0.0979031	G	A	0.952877	0.964766	-0.0003	0.001	0.72
rs35025768	Total dairy products §	1	43483707	0.707581	G	A	0.984486	0.857219	0.04	0.01	0.002
	Milk as a drink and added to cereals ‖	1	43483707	0.707581	G	A	0.984466	0.857219	0.002	0.002	0.42
	Milk added to coffee, tea or cereals ‖	1	43483707	0.707581	G	A	0.984486	0.857219	0.004	0.004	0.38
	Milk as drink ‖	1	43483707	0.707581	G	A	0.984486	0.857219	0.003	0.004	0.47
	Yoghurt ‖	1	43483707	0.707581	G	A	0.984486	0.857219	0.04	0.01	$1.4x10^{-8}$
	Cheese ‖	1	43483707	0.707581	G	A	0.984477	0.857219	0	0.002	0.99
rs151022760	Total dairy products §	2	136098560	1.58514	T	TATTG	0.730153	0.977032	0.03	0.004	$2.2x10^{-12}$
	Milk as a drink and added to cereals ‖	2	136098560	1.58514	T	TATTG	0.736534	0.977032	0.004	0.001	$9.8x10^{-9}$
	Milk added to coffee, tea or cereals ‖	2	136098560	1.58514	T	TATTG	0.730153	0.977032	0.01	0.001	$1.5x10^{-6}$
	Milk as drink ‖	2	136098560	1.58514	T	TATTG	0.730153	0.977032	0.01	0.001	$9.6x10^{-11}$
	Yoghurt ‖	2	136098560	1.58514	T	TATTG	0.730153	0.977032	0.01	0.002	$5.2x10^{-4}$
	Cheese ‖	2	136098560	1.58514	T	TATTG	0.736508	0.977032	-0.0003	0.0004	0.52
rs182549	Total dairy products §	2	136616754	1.58582	C	T	0.255625	0.975123	-0.03	0.004	$1.5x10^{-15}$
	Milk as a drink and added to cereals ‖	2	136616754	1.58582	C	T	0.249371	0.975123	-0.004	0.001	$3.6x10^{-12}$
	Milk added to coffee, tea or cereals ‖	2	136616754	1.58582	C	T	0.255625	0.975123	-0.01	0.001	$6.4x10^{-7}$
	Milk as drink ‖	2	136616754	1.58582	C	T	0.255625	0.975123	-0.01	0.001	$5.7x10^{-12}$
	Yoghurt ‖	2	136616754	1.58582	C	T	0.255625	0.975123	-0.01	0.002	$3.4x10^{-5}$
	Cheese ‖	2	136616754	1.58582	C	T	0.249391	0.975123	0.001	0.0004	0.23
rs28815269	Total dairy products §	2	136979530	1.58915	G	A	0.351051	0.963135	-0.02	0.003	$2x10^{-9}$
	Milk as a drink and added to cereals ‖	2	136979530	1.58915	G	A	0.347311	0.963135	-0.002	0.001	0.005
	Milk added to coffee, tea or cereals ‖	2	136979530	1.58915	G	A	0.351051	0.963135	-0.003	0.001	0.01
	Milk as drink ‖	2	136979530	1.58915	G	A	0.351051	0.963135	-0.004	0.001	$5.5x10^{-6}$
	Yoghurt ‖	2	136979530	1.58915	G	A	0.351051	0.963135	-0.005	0.002	0.007
	Cheese ‖	2	136979530	1.58915	G	A	0.347251	0.963135	0.0005	0.0004	0.2

Table A.34 (continued)

SNP	Dairy phenotype	Chromosome	Base position	Gene position	Effect allele	Minor allele	EAF	INFO †	b ‡	se	*p*
rs3940549	Total dairy products §	2	136138627	1.58515	A	G	0.723244	0.983431	0.03	0.004	$2x10^{-13}$
	Milk as a drink and added to cereals ‖	2	136138627	1.58515	A	G	0.729798	0.983431	0.004	0.001	$9.3x10^{-10}$
	Milk added to coffee, tea or cereals ‖	2	136138627	1.58515	A	G	0.723244	0.983431	0.01	0.001	$3.7x10^{-7}$
	Milk as drink ‖	2	136138627	1.58515	A	G	0.723244	0.983431	0.01	0.001	$1.1x10^{-10}$
	Yoghurt ‖	2	136138627	1.58515	A	G	0.723244	0.983431	0.01	0.002	$1.9x10^{-4}$
	Cheese ‖	2	136138627	1.58515	A	G	0.729762	0.983431	-0.0003	0.0004	0.43
rs4988235	Total dairy products §	2	136608646	1.58582	G	A	0.258622	0.972914	-0.03	0.004	$8.6x10^{-16}$
	Milk as a drink and added to cereals ‖	2	136608646	1.58582	G	A	0.252285	0.972914	-0.005	0.001	$8.9x10^{-13}$
	Milk added to coffee, tea or cereals ‖	2	136608646	1.58582	G	A	0.258622	0.972914	-0.01	0.001	$1.1x10^{-6}$
	Milk as drink ‖	2	136608646	1.58582	G	A	0.258622	0.972914	-0.01	0.001	$1.3x10^{-11}$
	Yoghurt ‖	2	136608646	1.58582	G	A	0.258622	0.972914	-0.01	0.002	$2.3x10^{-5}$
	Cheese ‖	2	136608646	1.58582	G	A	0.252309	0.972914	0.001	0.0004	0.24
rs504764	Total dairy products §	2	45154662	0.697037	C	T	0.666122	0.9923	-0.01	0.003	0.11
	Milk as a drink and added to cereals ‖	2	45154662	0.697037	C	T	0.667261	0.9923	-0.0004	0.001	0.44
	Milk added to coffee, tea or cereals ‖	2	45154662	0.697037	C	T	0.666122	0.9923	-0.002	0.001	0.08
	Milk as drink ‖	2	45154662	0.697037	C	T	0.666122	0.9923	-0.001	0.001	0.25
	Yoghurt ‖	2	45154662	0.697037	C	T	0.666122	0.9923	0.004	0.002	0.01
	Cheese ‖	2	45154662	0.697037	C	T	0.667252	0.9923	-0.002	0.0004	$4.4x10^{-8}$
rs62168795	Total dairy products §	2	136429366	1.58551	T	C	0.741356	0.949956	0.03	0.004	$3.9x10^{-12}$
	Milk as a drink and added to cereals ‖	2	136429366	1.58551	T	C	0.747538	0.949956	0.004	0.001	$6x10^{-11}$
	Milk added to coffee, tea or cereals ‖	2	136429366	1.58551	T	C	0.741356	0.949956	0.01	0.001	$2.1x10^{-6}$
	Milk as drink ‖	2	136429366	1.58551	T	C	0.741356	0.949956	0.01	0.001	$6.9x10^{-10}$
	Yoghurt ‖	2	136429366	1.58551	T	C	0.741356	0.949956	0.01	0.002	$6x10^{-4}$
	Cheese ‖	2	136429366	1.58551	T	C	0.747527	0.949956	-0.001	0.0004	0.19
rs6754311	Total dairy products §	2	136707982	1.58595	T	C	0.740847	0.968571	0.03	0.004	$3.7x10^{-14}$
	Milk as a drink and added to cereals ‖	2	136707982	1.58595	T	C	0.747096	0.968571	0.005	0.001	$3.8x10^{-13}$
	Milk added to coffee, tea or cereals ‖	2	136707982	1.58595	T	C	0.740847	0.968571	0.01	0.001	$2.5x10^{-6}$
	Milk as drink ‖	2	136707982	1.58595	T	C	0.740847	0.968571	0.01	0.001	$1.7x10^{-11}$
	Yoghurt ‖	2	136707982	1.58595	T	C	0.740847	0.968571	0.01	0.002	$1.6x10^{-5}$
	Cheese ‖	2	136707982	1.58595	T	C	0.747093	0.968571	-0.0004	0.0004	0.29
rs74775210	Total dairy products §	2	136352327	1.58521	T	TTC	0.727258	0.979213	0.03	0.004	$3x10^{-13}$
	Milk as a drink and added to cereals ‖	2	136352327	1.58521	T	TTC	0.733624	0.979213	0.004	0.001	$7.5x10^{-10}$
	Milk added to coffee, tea or cereals ‖	2	136352327	1.58521	T	TTC	0.727258	0.979213	0.01	0.001	$7.6x10^{-7}$
	Milk as drink ‖	2	136352327	1.58521	T	TTC	0.727258	0.979213	0.01	0.001	$2.8x10^{-11}$
	Yoghurt ‖	2	136352327	1.58521	T	TTC	0.727258	0.979213	0.01	0.002	$1.3x10^{-4}$
	Cheese ‖	2	136352327	1.58521	T	TTC	0.733595	0.979213	-0.0003	0.0004	0.44

Table A.34 (continued)

SNP	Dairy phenotype	Chromosome	Base position	Gene position	Effect allele	Minor allele	EAF	INFO †	b ‡	se	*p*
rs7570971	Total dairy products §	2	135837906	1.58509	C	A	0.721428	1	0.03	0.004	$9.8x10^{-14}$
	Milk as a drink and added to cereals ‖	2	135837906	1.58509	C	A	0.727838	1	0.004	0.001	$2x10^{-10}$
	Milk added to coffee, tea or cereals ‖	2	135837906	1.58509	C	A	0.721428	1	0.01	0.001	$2.3x10^{-10}$
	Milk as drink ‖	2	135837906	1.58509	C	A	0.721428	1	0.01	0.001	$2.4x10^{-7}$
	Yoghurt ‖	2	135837906	1.58509	C	A	0.721428	1	0.01	0.002	$1.9x10^{-4}$
	Cheese ‖	2	135837906	1.58509	C	A	0.727802	1	-0.0002	0.0004	0.59
rs11940694	Total dairy products §	4	39414993	0.613896	A	G	0.396908	0.982314	0.02	0.003	$5.2x10^{-11}$
	Milk as a drink and added to cereals ‖	4	39414993	0.613896	A	G	0.393814	0.982314	0.001	0.001	0.16
	Milk added to coffee, tea or cereals ‖	4	39414993	0.613896	A	G	0.396908	0.982314	0.001	0.001	0.22
	Milk as drink ‖	4	39414993	0.613896	A	G	0.396908	0.982314	0	0.001	0.9
	Yoghurt ‖	4	39414993	0.613896	A	G	0.396908	0.982314	0.01	0.002	$4.6x10^{-6}$
	Cheese ‖	4	39414993	0.613896	A	G	0.393907	0.982314	0	0.0004	0.9
rs10245608	Total dairy products §	7	41241212	0.657781	C	T	0.505258	0.989992	-0.01	0.003	0.01
	Milk as a drink and added to cereals ‖	7	41241212	0.657781	C	T	0.506704	0.989992	-0.001	0.001	0.15
	Milk added to coffee, tea or cereals ‖	7	41241212	0.657781	C	T	0.505258	0.989992	-0.001	0.001	0.13
	Milk as drink ‖	7	41241212	0.657781	C	T	0.505258	0.989992	-0.001	0.001	0.52
	Yoghurt ‖	7	41241212	0.657781	C	T	0.505258	0.989992	-0.001	0.002	0.48
	Cheese ‖	7	41241212	0.657781	C	T	0.506599	0.989992	-0.002	0.0004	$3.6x10^{-8}$
rs10264126	Total dairy products §	7	132620098	1.44451	C	T	0.742713	0.999143	0.01	0.004	0.13
	Milk as a drink and added to cereals ‖	7	132620098	1.44451	C	T	0.741229	0.999143	-0.0002	0.001	0.73
	Milk added to coffee, tea or cereals ‖	7	132620098	1.44451	C	T	0.742713	0.999143	0.001	0.001	0.42
	Milk as drink ‖	7	132620098	1.44451	C	T	0.742713	0.999143	-0.001	0.001	0.43
	Yoghurt ‖	7	132620098	1.44451	C	T	0.742713	0.999143	0.001	0.002	0.57
	Cheese ‖	7	132620098	1.44451	C	T	0.741248	0.999143	0.002	0.0004	$4.4x10^{-9}$
rs35754956	Total dairy products §	12	122511116	1.46556	A	G	0.870971	0.990334	0.03	0.005	$1x10^{-8}$
	Milk as a drink and added to cereals ‖	12	122511116	1.46556	A	G	0.869929	0.990334	0.001	0.001	0.38
	Milk added to coffee, tea or cereals ‖	12	122511116	1.46556	A	G	0.870971	0.990334	0.004	0.001	0.002
	Milk as drink ‖	12	122511116	1.46556	A	G	0.870971	0.990334	0.003	0.001	0.009
	Yoghurt ‖	12	122511116	1.46556	A	G	0.870971	0.990334	0.003	0.002	0.19
	Cheese ‖	12	122511116	1.46556	A	G	0.869881	0.990334	-0.0001	0.001	0.81

†Imputation score for the assessment of the quality of imputation done in IMPUTE2 programme. SNPs with a score <0.4 were dropped from the analyses
‡Beta coefficient and standard error as derived from BOLT-LMM software, which uses linear regression models
§A continuous phenotype was used for total dairy products in servings/day
‖Categorical phenotypes were used for milk consumed as a drink and added to cereals (total sample), milk consumed as a drink (Oxford WebQ subset), milk added to tea, coffee and cereals (Oxford WebQ subset), yoghurt (Oxford WebQ subset) and cheese (total sample) splitting the sample into consumers and non-consumers
Abbreviations: SNP: single nucleotide polymorphisms, EAF: effect allele frequency

Appendix B

Glossary of genetic terms

The following definitions of genetic terms were extracted from the website of NSW Centre for Genetics Education (http://www.genetics.edu.au/publications-and-resources/glossary) and are provided to facilitate the reader to understand key concepts of genetic epidemiology included in Chapter 8.

Allele: There are usually two copies of a gene. These two copies are called alleles. In some cases, one or both alleles will be mutated or altered in some way.

Autosomal: Of the autosome i.e. any chromosome that is not a sex chromosome (that is not an X or Y chromosome). In humans, the autosomes are the numbered chromosomes and are given the numbers 1-22. Chromosome 1 is the largest and 22 is the smallest.

Chromosome: A threadlike structure found in the nucleus of all the body cells (except red blood cells) consisting of DNA and proteins. Each chromosome can be thought of as a string of beads where every bead represents a gene.

Dominant: Every cell contains two copies of each gene. Where only one of the gene copies or allele is mutated, and the other allele is 'correct', but the person is affected by a genetic condition due to that mutation, the mutation is described as dominant. The mutated gene is said to be dominant over the other 'correct' copy of the gene. A condition or characteristic caused by a dominant gene mutation only requires one of the genes to be mutated for the person to be affected.

Gene: The basic unit of heredity; a segment of DNA which contains the information for a specific characteristic or function.

Genome: The complete set of genes carried by an individual or a cell.

Genotype: The genetic constitution of an individual

Heterozygote: An individual who has two different alleles at a particular gene locus, one on each chromosome of a pair. One allele is usually normal and the other abnormal. Such an individual may also be referred to as a carrier.

Homozygote: Refers to an individual in whom the two alleles or gene copies contain identical information. An individual can be homozygous for the correct copies of the gene or can be homozygous for the mutated copies of the gene.

Intron: The part of the genetic sequence that is not translated into the final gene product or message.

Linkage: The tendency for genes or segments of DNA which are located close together on the same chromosome to be inherited together.

Phenotype: The physical and/or biochemical characteristics of a person, an animal or other organism which are determined by their genetic make-up and/or environment.

Recessive: Every cell contains two copies of each gene. Each gene contains the information for a particular gene product, such as a protein. If a gene is mutated, the gene no longer codes for the gene product. Where an individual has one gene copy or allele mutated and the other copy 'correct', the cell will only be producing half the amount of gene product. If this does not result in any condition for the individual, the mutation is described as being hidden or 'recessive' to the correct copy of the gene. An individual with this genetic constitution is said to be a 'carrier' of a recessive gene mutation. For a recessive gene mutation to result in a particular characteristic or a condition, both copies of the genes must be mutated.

Single Nucleotide Polymorphism (SNP): A DNA sequence variation that involves a change in a single nucleotide. Variations in the genetic code at the level of one nucleotide may be useful in certain applications such as assessing the patterns of inheritance via genetic linkage studies, or forensic DNA testing or DNA fingerprinting.