Dairy Intake and Prediabetes Risk

Evidence from Multiple Cohort Studies and Meta-Analyses

Isabel Slurink

Dairy Intake and Prediabetes Risk

Evidence from Multiple Cohort Studies and Meta-Analyses

Isabel Slurink

Colofon
The work presented in this thesis was supported by a grant from the Dutch Dairy Association [Nederlandse Zuivelstichting (NZO)].

Provided by thesis specialist Ridderprint, ridderprint.nl
Printing Ridderprint
Layout and design Indah Hijmans, persoonlijkproefschrift.nl
Cover design Indah Hijmans

ISBN 978-94-6506-450-5

Dairy Intake and Prediabetes Risk

Evidence from Multiple Cohort Studies and Meta-Analyses

Proefschrift ter verkrijging van de graad van doctor aan Tilburg University op gezag van de rector magnificus, prof. dr. W.B.H.J. van de Donk, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de Aula van de Universiteit op woensdag 27 november 2024 om 14.00 uur door

Isabel Anna Laurien Slurink

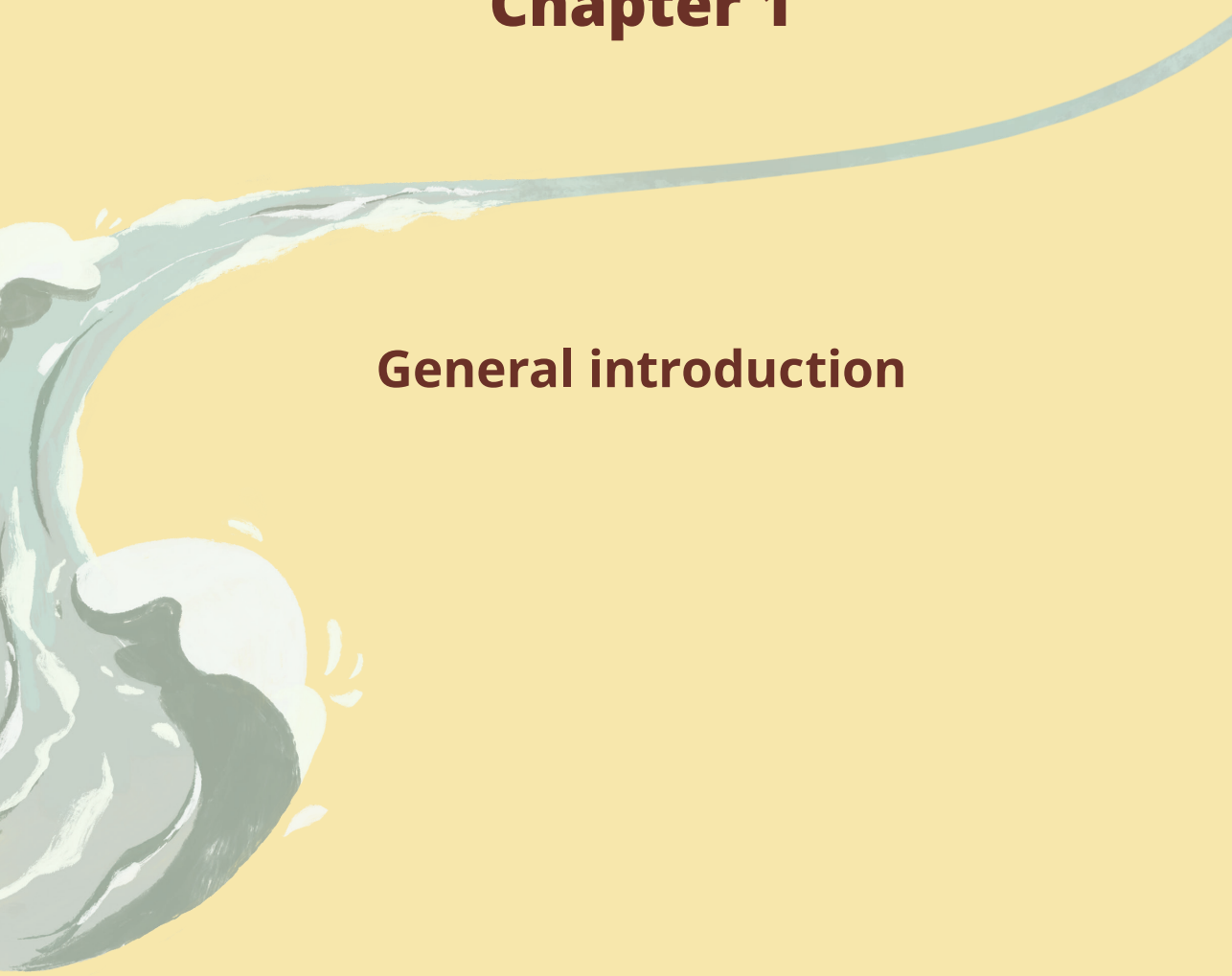
geboren te Wageningen

Promotores	prof. dr. T. Smeets, Tilburg University
	dr. Soedamah-Muthu, Tilburg University
Promotiecommissie	prof. dr. L.A.M. van de Goor, Tilburg University
	prof. dr. ir. I. Brouwer, Vrije Universiteit Amsterdam
	prof. dr. J.M. Geleijnse, Wageningen University & Research
	prof. dr. R.P. Mensink, Maastricht University
	dr. K.A.C. Berk, Erasmus MC

Als iemand echt van de natuur houdt, vindt men overal schoonheid.
Vincent van Gogh

Contents

Chapter 1



General introduction

Background

The consumption of milk and dairy foods has been a longstanding tradition in many cultures. With the genetic adaptation to lactase persistence (LP), certain populations were able to continue consuming dairy into adulthood. Thereby, these populations obtained considerable intakes of proteins, fats, and other essential nutrients from dairy consumption, which in turn contributed to their survival and reproduction. Furthermore, while early modern humans obtained calcium and other minerals from wild plant foods [1], the domestication of animals led to milk becoming a predominant source. The development of pasteurization in the 19th century has led to improved safety and shelf life of milk and has therefore enabled the upscaling of dairy farming and wider consumption of dairy products. Milk and milk products form a substantial part of traditional diets in many countries. In several countries, there have been national campaigns to promote dairy consumption by emphasizing the nutritional benefits and its place in healthy diets, particularly in the post-war period when nutritional deficits were widespread. The Netherlands has a renowned cheese sandwich culture that traces its origins back to the Middle Ages and has endured ever since. Cheese sandwiches, toasties or white buns with cheese continue to be the most popular breakfast and lunch options for many of us 'kaaskoppen' [cheese heads]. Dutch hard cheeses like Gouda and Edam have become popular worldwide. In Dutch culture, the habit of eating dairy is instilled from a young age; for me it was mandatory to drink one glass of milk per day, as well as first having a savoury 'healthy' sandwich (i.e., with cheese), only then a 'bad' one with sweet toppings was allowed.

Dairy foods are currently recommended in many guidelines worldwide as part of a healthy diet. Dairy is an interesting preventive target for maintaining cardiometabolic health, as it is a rich source of protein, odd-chained fatty acids, calcium, magnesium, potassium, vitamins A, D, B2 and B12. Nevertheless, several concerns have been raised about dairy, as it is relatively high in saturated fat (SFA), sodium, potential hormones [2] and sugar, compounded by the high prevalence of lactose intolerance. Around 65% of the global population has some degree of lactose intolerance, up to 95% in the East Asian population [3]. With aging populations, the link between diets and chronic diseases has become more apparent. The discovery of a link between SFA and coronary heart disease (CAD) has resulted in critiques on dairy consumption, and consequently many guidelines recommend low-fat dairy types. However, this claim is currently not substantiated by literature, as the current literature does not show the harmful effects of dairy that would be expected based on the SFA content [4]. Dairy contains a complex mixture of fatty acids with potentially different health effects, and because they are eaten within a matrix, the *dairy matrix*, the effect depends on the interactions between various components within this matrix, further affected by the bioactivity of several components derived from various fermentation and processing techniques. Dairy has therefore received ample attention in the literature. Nevertheless, the health effects of dairy foods remain

heterogeneous, and the debate in the scientific world as well as in the public about the role of dairy in the prevention of cardiometabolic diseases is ongoing.

Dairy consumption

Dairy types and constituents

The 'dairy' group is a heterogeneous group of foods that vary in the structure, profile and amounts of nutrients, the fat content, the addition of sugars, water content, bioactive components, and other constituents, and the processing methods including fermentation or aging. Dairy foods can be divided into several types, including milk, powdered milk, fermented products, yogurt, quark (i.e., soft cheese), cheese, custard, cream, and ice cream. Studies that investigate the health effects of dairy foods often categorize products based on their relative fat content, distinguishing between low-fat (<2% for liquid dairy types or <20% for solid dairy types) and high-fat (≥2% for liquid dairy types or ≥20% for solid dairy types). Nevertheless, the absolute fat content of cheese is much higher due to a lower moisture content compared to high-fat milk. Within each dairy type, the low-fat varieties contain similar protein, lactose, minerals (calcium, sodium, potassium, phosphorus, magnesium), and water-soluble vitamins compared to the high-fat varieties, but less fat-soluble vitamins (**Table 1**). The fat-soluble vitamins A, D, E and K and several essential fatty acids are found in milk fat. The bioavailability of calcium was found not to substantially differ among various dairy products, but differences in the bioavailability of other vitamins and minerals are understudied [5, 6]. Most European countries, the UK and Australia do not fortify dairy products with vitamin D, unlike Finland, Canada, and the United States, which have national vitamin D fortification policies [7].

Full-fat milk consists of 87.6 g water, 3.4 g milk fat and 9 g non-fat constituents including 3.3 g protein, 4.5 g lactose (carbohydrates), and 0.65 g minerals (**Table 1**). Of proteins, 80% are casein micelles and 20% are whey proteins. Raw milk normally contains ~4.4 g of fat per 100 g. In the Netherlands, and many other countries such as the UK and Australia, the fat content is standardized to 3.5% fat for full-fat milk, 1.5-1.8% for semi-skimmed milk and 0.1% for skimmed milk. There are some minor variations in international standardization. Furthermore, modifications in the composition of milk are allowed if indicated on the product packaging.

Milk is an emulsion of milk fat globules within a water-based fluid that contains dissolved carbohydrates and protein aggregates with minerals. The milk fat globule is surrounded by a membrane consisting of bioactive polar lipids (phospholipids and sphingolipids) and proteins (i.e., the milk fat globule membrane, MFGM). Homogenization of milk leads to smaller fat globules with a membrane consisting mostly of milk protein. SFA amount and composition in milk depend on region, season, physiological factors, feeding and farming practices [8-10]. SFA accounts for 67-72% of total fat in milk, including mostly

palmitic (C16:0, 30–33% total FA), myristic (C14:0, 10–11 % total FA), stearic (C18:0, 9–10 % total FA) acids and short to medium chain FA (C4:0–C12:0, < 4 % total FA) [9]. Odd-chain FA such as pentadecanoic (C15:0) and heptadecanoic (C17:0) represent 1 % and 0.5% of total FA, respectively. Furthermore, milk contains monounsaturated (MUFA) and polyunsaturated fatty acids (PUFA) (18:1n-9, 18:2n-6, and 18:3n-3), branched-chain saturated fats (BCSFA), and trace amounts of natural (ruminant) trans fats (e.g., trans-palmitoleic acid, trans-C16:1n-7) [11].

Raw milk undergoes a separation process to separate a high-fat cream layer from a low-fat milk layer. Milk is then standardized by adding cream to achieve the desired fat content. Churning the cream causes fat globules to clump together, removing the MFGM, and forming butter, which is then separated from the liquid buttermilk. In research and dietary guidelines, butter is often excluded from the definition of the dairy food category as butter contains only triglycerides and lacks the other components found in whole milk. Instead, in the Dutch dietary guidelines, butter is included in the fats and oils group [12]. For dairy desserts such as custard and ice cream, milk, cream, sugars, and other ingredients are combined.

Fermented dairy refers to dairy products that are produced through the fermentation of milk with living (lactic acid) bacteria, such as yogurt, quark, buttermilk, and kefir. Yogurt is produced by adding a bacterial starter to milk which converts the milk sugars into lactic acid, changing the protein structure to more gel-like. For cheese production, natural rennet and bacterial starter cultures are added to milk, leading to coagulation of the casein proteins and the formation of a curd, which is separated from the liquid whey. This curd undergoes further processing involving drying, salting, and ripening, during which the proteins undergo further structural changes. This curd can also serve as the base for quark production. Many fermented dairy types contain probiotics, which are used as a starter culture alone or in combination with lactic acid bacteria or are incorporated into dairy after fermentation to enhance functional properties of the product [13]. Probiotics are live microorganisms that have health benefits if administered in adequate amounts [14]. Probiotics are capable of surviving passage through the digestive tract which enables them to confer their health benefits. Lactic acid bacteria do not survive passage through the digestive tract and therefore lack the same health-promoting properties attributed to probiotics.

Table 1 Nutrients in dairy foods per 100 gram as assessed by the Dutch Food Composition Database 2023.

Dairy product		Macronutrients						Minerals					Vitamins				
		Energy (kcal)	Water (g)	Protein (g)	Fat (g)	SFA (g)	CH (g)	Na (mg)	K (mg)	Ca (mg)	P (mg)	Mg (mg)	RAE (Vit A) (µg)	Vit D (µg)	Vit K2 (µg)	Vit B2 (mg)	Vit B12 (µg)
Milk	Full-fat	61	87.6	3.3	3.4	2.2	4.5	42	163	124	104	12	36	0	0.9	0.18	0.4
	Semi skimmed	45	89.4	3.4	1.4	0.9	4.7	42	160	123	104	12	15	0	0.5	0.18	0.45
	Skimmed	35	90.3	3.7	0.1	0.1	4.9	44	169	126	106	12	1	0	0	0.18	0.44
Yogurt	Full-fat	56	88.8	3.8	2.7	1.8	3.4	39	159	145	111	12	30	0	1	0.17	0.22
	Semi skimmed	50	90	4.2	1.5	0.9	4.3	51	142	139	88	12	16	0	0.5	0.2	0.39
	Skimmed	37	91	4.1	0.2	0.1	4	43	159	152	118	13	2	0	0.1	0.17	0.28
Quark	Full-fat	118	79.5	7.5	8	5.2	3.6	35	159	134	122	11	85	0.1	5.1	0.14	0.7
	Semi skimmed	77	82.7	7.5	3.2	2.1	4	40	160	137	131	12	42	0.1	10.5	0.14	0.7
	Skimmed	51	86	8.4	0	0	3.8	37	160	139	139	12	0	0	15.9	0.14	0.7
Cheese	Full-fat 48+	370	39.1	22.9	30.6	19.9	0	670	80	819	542	33	333	0.3	64.5	0.28	2.03
	Medium fat 30+	280	45.9	30.1	17.7	11.6	0	676	77	1021	612	40	173	0.2	43.6	0.32	1.72
	Low fat 20+	246	48.1	34.5	12	7.8	0	683	100	1059	710	42	98	0.1	32.8	0.4	2.8
	Brie 60+	369	46.1	17	33	21.5	1	471	109	250	200	20	298	0.3	12.5	0.52	1.2
	Cottage cheese	92	79.9	11.2	3.9	2.5	2.3	370	85	75	141	7	47	0.1		0.25	0.76
	Ricotta	141	77.8	8	10.3	6.7	3.9	164	110	240	170	13	193			0.19	0.3
Custard	Full-fat	120	76	2.7	5.4	3.5	15.1	47	115	117	73	10	50	0	1.3	0.14	0.26
	Semi skimmed	85	80.1	2.3	1.7	1.1	15.1	53	121	89	69	9	15	0	0.5	0.17	0.26
Cream	Whipped	339	58.8	2.3	35.3	23	3.1	36	97	78	58	8	375	0.1	5.4	0.12	0.2
	Cooking	212	72.3	2.4	20.2	13.2	5	50	101	88	61	8	164	0.1		0.12	0.27
Ice cream		220	60	3	12	7.8	24.9	46	172	97	85	11	176	0		0.23	0.41
Butter		737	15.4	0.7	81.2	53.6	1.1	5	27	17	23	2	874	0.3	15	0.03	0.3

Source: (NEVO-online versie 2023/8.0, RIVM, Bilthoven; www.rivm.nl/nevo). Empty cells indicate that no information about a nutrient is available.
Abbreviations: Ca, Calcium; CH, Carbohydrate; K, Potassium; Mg, Magnesium; Na, Sodium; P, Phosphorus; RAE, Retinol Activity Equivalent; SFA, saturated fatty acids; Vit, vitamin.

Guidelines for dairy consumption

Many guidelines worldwide recommend consuming 2-3 servings daily, mostly driven by the provision of sufficient calcium into diets [15, 16]. Generally, dietary guidelines emphasize intake of plain or unsweetened, low-fat dairy to limit intake of added sugars and saturated fatty acids [15].

Current Dutch guidelines for dairy consumption

In 2015, the Dutch Health Council (HCN) formulated for the first time food-based dietary guidelines, based on complementing evidence from prospective cohort studies and RCTs on the prevention of the top 10 diseases contributing to mortality, years of potential life lost, or burden of disease in the Netherlands (i.e., CHD, stroke, heart failure, type 2 diabetes (T2D), chronic obstructive pulmonary diseases, breast cancer, colorectal cancer, lung cancer, dementia, and depression) and risk factors (i.e., blood pressure, LDL cholesterol and body weight) [17]. Based on extensive literature reviews, the Dutch Health Council recommends consuming dairy products daily, including milk or yogurt and cheese, as literature shows that yogurt is associated with lower T2D risk, and dairy and milk are associated with lower colon cancer risk. Additionally, they state that dairy foods are important sources of essential nutrients including calcium, potassium, and vitamin A, especially relevant for some groups with low intake of these nutrients. The Dutch Health Council did not recommend a specific fat content for dairy products due to insufficient evidence on whether the fat content of dairy has distinct health effects.

This advice has been translated by the 'Voedingscentrum' [The Netherlands Nutrition Centre] into the 'Schijf van Vijf' [Wheel of Five], which provides specific dietary recommendations for the general public [12] (**Figure 1**). For both men and women, the guidelines recommend consuming 2-3 servings per day, along with a maximum of 40 grams of cheese per day. This intake covers 80-90% of the Recommended Dietary Allowance (RDA) for calcium is achieved. Low-fat dairy is preferred due to the higher energy density of high-fat dairy. Furthermore, low-fat dairy options are recommended to meet the guideline of limiting total saturated fat content of the diet to ≤10% of total energy (en%). Milk and milk products included in this advice may contain ≤1.1 g of saturated fat per 100 g, ≤6 g of total sugar per 100 g, and should not have added trans fats, sodium, or sugars. Thereby, skimmed, and semi-skimmed milk and yogurt, and curd cheese are recommended, but high-fat milk and high-fat yogurt are not. Also, puddings, custard, ice creams and other dairy desserts with added sugar are excluded from this recommendation. Furthermore, dairy drinks are considered as a sugar-sweetened beverage (SSB) if the total sugar content is >6g/100g and is then excluded from recommendations. For cheese, the guidelines are ≤14g/100g SFA and ≤820mg/100g sodium, and no added trans fats or sugar. Thereby, it is recommended to consume 20+ or 30+ cheese with reduced salt, soft goat cheese, mozzarella, and dairy spread, but not 48+ and diverse foreign cheeses such as brie and blue cheeses.

The last advisory report from the HCN for the Dutch dietary guidelines was published in 2015 [17]. This advisory report was supplemented with specific recommendations for individuals with T2D in 2021 [18] and cardiovascular disease (CVD) in 2023 [19]. Both reports concluded evidence was insufficient on health effects among these patient groups to deviate from the existing dietary guidelines for the general population.

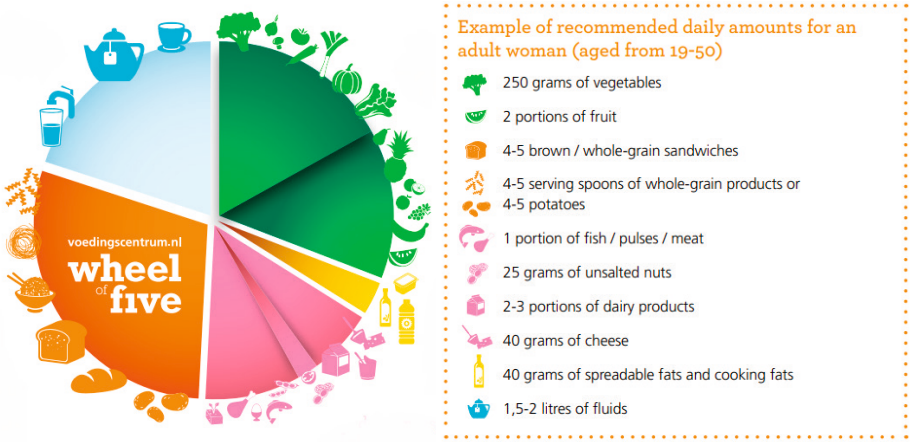


Figure 1. The 'Schijf van Vijf' [Wheel of Five] and an example of recommended daily amounts for adult women (aged from 19-50), including the recommendation to consume 2-3 portions of dairy products. Derived from: https://www.voedingscentrum.nl/

The HCN published an advisory report on a 'healthy protein transition' in 2023 [20]. This transition encompasses a change in the dietary patterns of Dutch individuals to 60% plant-based proteins and 40% animal-based proteins. Currently, 57% of proteins are from animal-based sources. Research shows that a transition in diets is needed to be able to comply with climate goals and to ensure sufficient healthy foods for the world population. The greenhouse gas emissions of milk and milk products among Dutch individuals aged 19-30 years was 12% in men and 13% in women [21]. For cheese, this was 6% and 7%, respectively. This makes dairy the second largest in greenhouse gas emissions, after 31% and 29% for red meat among men and women, respectively. To lower the environmental burden of Dutch diets, the focus should be on reducing consumption of red meat and both alcoholic and non-alcoholic beverages, increasing water and tea intake, and moderating dairy intake to recommended levels [20, 21]. The Eat-Lancet Commission on healthy diets from sustainable food systems states that a wide range of intakes may fit within an overall healthy diet, with an optimal intake of 250 g/day within a range of 0-500 g/day [22]. No specification of fat content is made, as with low-fat dairy intake, the fat proportion will remain in the human food supply for example as butter or cream, resulting in limited effects on population health.

Plant-based dairy alternatives

Achieving a sufficient intake of high-quality proteins, calcium, magnesium, and potassium in a diet without dairy products requires effort. Without dairy products in the diet, individuals need to actively seek alternative sources to meet their nutritional needs. Suitable alternative sources include legumes, nuts, and seeds, while calcium, magnesium, and potassium can be obtained from various fruits, vegetables, and fortified foods. Many plant-based alternatives for dairy have entered the market in the Netherlands, and a comparison of nutrient content is made in **Table 2**. Plant-based beverages need to adhere to criteria for dairy products including calcium ≥ 80 mg/100 g (≥ 500 mg/100g for plant-based cheese alternatives), vitamin B12 ≥ 0.24 mcg/100 g ≥ 20 en% from protein. Soy drinks provide more protein, magnesium, and vitamin K1 compared to other plant-based drinks. Except for soy drinks, other milk alternatives are not considered good sources of protein. Concerns have been raised about the lower protein quality and bioavailability of plant-based drinks compared to milk [23, 24], as well as the lower bioavailability of fortified calcium in plant-based beverages [25, 26]. Additional comparative studies on the absorption and digestibility of proteins, added vitamins and minerals, and health effects of plant-based alternatives are needed.

Dairy consumption patterns

The consumption of dairy products in different populations is highly heterogeneous due to various demographic, cultural and socioeconomic factors. Dairy consumption in the Netherlands is relatively high compared to other countries, attributed to dairy products being embedded deeply in our culture, a low percentage of adults with lactose intolerance (2%) and a highly developed dairy industry offering a diverse range of nutrient-rich products tailored to trends in consumer behaviour.

The dietary habits of Dutch citizens are monitored in the Voedselconsumptiepeiling (VCP) [Dutch National Food Consumption Survey] [27]. In the VCP of 2019-2021, the mean dairy intake was 338 grams/day of dairy products, and 13 grams/day of dairy substitutes. Milk-based drinks are consumed most frequently (51%), followed by yogurt (20%) and cheese (9%). Consumption in children (336 grams/day) and adults (348 grams/day) is comparable, although children consume more milk and milk products, and adults more yogurt, quark, and cheese. Furthermore, intake in men (338 grams/day) is higher than in women (315 grams/day) mostly due to higher milk and milk products consumption.

Table 2 Nutrients in dairy foods per 100 grams compared with plant-based alternatives as assessed by the Dutch Food Composition Database 2023.

Dairy product or alternative	Energy (kcal)	Water (g)	Protein (g)	Fat (g)	SFA (g)	CH (g)	Na (mg)	K (mg)	Ca (mg)	P (mg)	Mg (mg)	RAE (Vit A) (µg)	Vit D (µg)	K (µg)	B2 (mg)	B12 (µg)
Milk																
Animal-based																
Semi skimmed milk	45	89.4	3.4	1.4	0.9	4.7	42	160	123	104	12	15	0	0.7	0.18	0.45
Plant-based																
Soy drink no sugar fortified with calcium and vitamins	34	93.6	3.4	1.8	0.3	0.6	35	150	120	44	18	0	0.8	4.8	0.21	0.38
Oat drink no sugar fortified with calcium and vitamins	44	90.2	0.5	1.5	0.1	6.7	39	30	119	12	3	0	1.1	0.4	0.21	0.36
Yogurt																
Animal-based																
Semi skimmed yogurt	50	90	4.2	1.5	0.9	4.3	51	142	139	88	12	16	0	0.7	0.2	0.39
Plant-based																
Soy yogurt no sugar fortified with calcium and vitamins	39	92.8	4	2.3	0.4	0	96	150	120	97	19	0	0.8	4.2	0.01	0.38
Cheese																
Animal-based																
Full-fat 48+ cheese	370	39.1	22.9	30.6	19.9	0	670	80	819	542	33	333	0.3	68.4	0.28	2.03
Plant-based																
Plant-based slices based on coconut oil fortified with Ca and B12	286	54.2	0.1	21.3	18	22.7	827	9	400	15	3	52	0	0.4	0.05	1.83

Source: (NEVO-online versie 2023/8.0, RIVM, Bilthoven; www.rivm.nl/nevo). Empty cells indicate that no information about a nutrient is available.
Abbreviations: Ca, Calcium; CH, Carbohydrate; K, Potassium; Mg, Magnesium; Na, Sodium; P, Phosphorus; RAE, Retinol Activity Equivalent; SFA, saturated fatty acids; Vit, vitamin.

Dairy consumption habits change over time. Total dairy intake has decreased in the Netherlands by 14%, since the VCP of 2007-2010 among all age groups and both sexes, and both for milk and yogurt products and other dairy types. The VCP shows that since the publication of the 2015 dietary guidelines, average dairy intake has remained stable. Based on consumption data from ZuivelNL [Organization of the Dutch dairy sector] and Statistics Netherlands (CBS), the intake of milk decreased by approximately 21% (56.6 to 44.7 kg milk per person per year) from 2005 to 2021 [28]. In contrast, cheese and quark intake increased by 40% (18.4 to 25.7 kg of cheese per person per year). This shift in dairy intake patterns can be attributed to changing lifestyle trends, dietary preferences, and a deeper understanding of the health effects of various types of dairy foods. Two explanations are noteworthy here. First, over recent years it has become more apparent that increasing dietary protein is favourable for promoting muscle growth and increasing energy expenditure. Consequently, products like quark have gained greater attention in the fitness and dairy sectors due to its high protein content and favourable total fat-to-energy ratio. Second, concerns about the climate impact of dairy products might have led to consumers choosing to decrease their milk intake, especially since dairy alternatives with similar nutrient content and taste are widely available, in line with decreasing meat intake. For cheese, consumers might not be willing to decrease their intake due to its high palatability which is difficult to replace with plant-based options [29].

Worldwide, more than 6 billion people consume milk and milk products; most of these people live in developing countries. The milk and dairy production (including butter) is projected to increase globally by 0.8% in 2020-22 to 15.7 kg in 2032 (milk solids, excluding water content of milk or dairy products), driven by population and income growth mainly in India, Pakistan, and several African Countries (**Figure 2**) [30]. Of dairy consumption worldwide, 81% of milk is cow's milk, 15% buffalo and 4% other such as goat, sheep, and camel. Most dairy products are consumed in the form of fresh dairy products which underwent minimal processing (i.e., pasteurised or fermented, including milk, yogurts, quark). In low- and middle-income countries, two-thirds of the average per capita dairy production is fresh dairy. In high-income countries, consumption of processed dairy products is higher (i.e., butter, cheese, skim milk powder, whole milk powder, whey powder and, for few cases casein). In Europe, consumption trends are similar to the Netherlands. In Northern Europe, milk production is expected to decline as domestic demands stagnate due to low population growth and declining population per capita consumption of fresh dairy products, partly at the expense of increased intake of plant-based replacements. Cheese intake is increasing in many European and North American regions. Furthermore, in Southeast Asian countries, cheese intake will increase due to urbanization and income increases resulting in more fast-food intake (i.e., pizza or burgers with cheese).

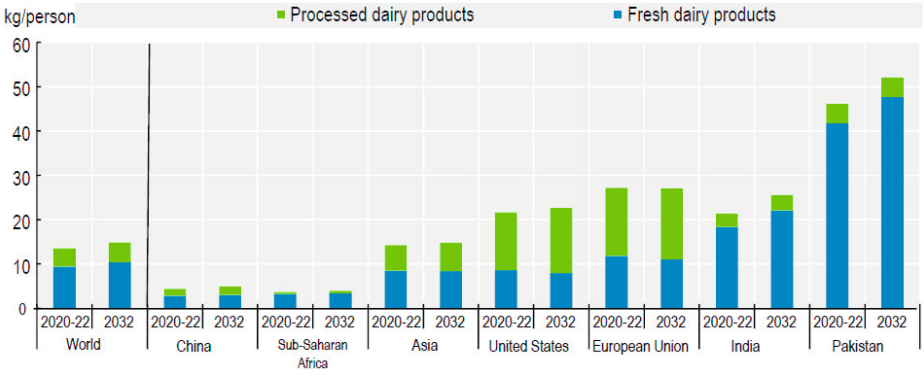


Figure 2. Per capita consumption of processed and fresh dairy products in milk solids. Milk solids are calculated by adding the amount of fat and non-fat solids for each product; Processed dairy products include butter, cheese, skim milk powder and whole milk powder. Source: OECD/FAO (2023), ''OECD-FAO Agricultural Outlook'', OECD Agriculture statistics (database), http://dx.doi.org/10.1787/agr-outl-dataen [30].

Burden of prediabetes

Prediabetes is an intermediate stage between normoglycaemia and T2D, defined by plasma glucose levels that are higher than the normal range but fall just below the diagnostic threshold for T2D [31]. Diagnostic cut-off levels for normoglycaemia and prediabetes are outlined in **Table 3**. People in this risk stage already display insulin resistance and declined pancreatic beta-cell function, resulting in impaired fasting or postprandial glycaemia. Prediabetes is asymptomatic and describes a high-risk stage for progressing to micro-and macrovascular diseases. It is important to acknowledge that not all individuals with prediabetes will eventually develop T2D, nor that individuals without prediabetes will remain free from developing T2D.

Prediabetes definition

The term 'prediabetes' was used for the first time in 1979 by the National Diabetes Data Group to describe impaired glucose tolerance (IGT), based on a two-hour postprandial glucose (2hPG) after an oral glucose tolerance test (OGTT) of 7.8–11.0 mol/L [140–199 mg/dL] [32]. This definition was adopted by the American Diabetes Organisation (ADA) and the World Health Organisation (WHO), who both later (1997 and 1998, respectively) additionally introduced a definition of impaired fasting glucose (IFG) based on fasting plasma glucose (FPG) values of 6.1–6.9 mmol/L [110–125 mg/dL] [31]. In 2003, the ADA broadened the prediabetes definition, lowering the threshold for FPG from 6.1 mmol/L to 5.4 mmol/L to optimise sensitivity and specificity in T2D risk prediction and make IFG and IGT prevalence more similar [33]. In 2009, the International Expert Committee (IEC) recommended using HbA1c for T2D diagnosis as this measure reflects long-term glucose exposure, also identifying a high-risk group at HbA1c levels of 6.0-6.5% for whom preventive interventions are recommended [34]. In 2010, the ADA defined a slightly

lower HbA1c cut-off for prediabetes diagnosis, ranging from 39-46 mmol/mol [5.7-6.4%] [35]. The WHO decided not to adopt these lower cut-offs, as the prediabetes prevalence based on the ADA cut-offs is substantially higher than based on the WHO-ICE cut-offs, creating a lower-risk group with a better cardiometabolic risk profile.

Table 3. Definitions of normoglycaemia and prediabetes according to different guidelines

Definition by	Marker	Normoglycaemia	Prediabetes
WHO [31]	FPG	<6.1 mmol/L (<110 mg/dL)	**IFG**: 6.1-6.9 mmol/L (110-125 mg/dL)
ADA [35]	FPG	<5.6 mmol/L (<100 mg/dL)	**IFG**: 5.6-6.9 mmol/L (100-125 mg/dL)
WHO/ADA [31, 35]	2hPG	<7.8 mmol/L (<140 mg/dL)	**IGT**: 7.8-11.0 mmol/L (140-199 mg/dL)
IEC [34]	HbA1c	<42 mmol/mol (<6.0%)	42-46 mmol/mol (6.0-6.4%)
ADA [35]	HbA1c	<39 mmol/mol (<5.7%)	39-46 mmol/mol (5.7-6.4%)

Abbreviations: 2hPG, two-hour postprandial glucose; ADA, American Diabetes Association; FPG, fasting plasma glucose; IEC, International Expert Committee; IFG, impaired fasting glucose; IGT, impaired glucose tolerance; HbA1c, haemoglobin A1c; WHO, World Health Organization.

Prediabetes prevalence

The prevalence of prediabetes is increasing worldwide at an alarming rate due to the aging of populations, economic developments, and unhealthier behaviours [36]. A high prevalence is especially observed in people with obesity and of older age [37]. There is limited overlap between IGT and IFG; only 20–25% of people with IGT have IFG, and 30–45% of individuals with IFG have IGT [38]. This is because IGT tends to be characterized by insulin resistance in muscle and decreased glucose uptake, while IFG is generally driven by insulin resistance in the liver and excess hepatic glucose production.

A review of over 7,000 studies showed a global age-adjusted prevalence of 9.1% (464 million) for IGT, and 5.8% (298 million) for IFG (based on WHO definitions) among adults aged 20-79 years in 2021 [36]. The prevalence was highest in high-income countries (11.2% for IGT and 6.4% IFG, respectively). By 2045, the global prevalence of IGT and IFG was projected to increase to 10.0% and 6.5% (414 million), respectively. In the Netherlands, 1.156.900 (~10%) adults have T2D [39]. Exact estimates of prediabetes are lacking, but estimates range from 10% [40] to 25% of the adult population [41].

The prevalence of prediabetes varies according to the diagnostic tests and cut-off values used. For example, in the US National Health and Nutrition Examination Survey (NHANES) of 2015-2016, the prevalence of prediabetes in adults aged 20 or older was 4.3% based on the IEC-HbA1c definition, 12.3% based on the ADA-HbA1c definition, 14.7% based on the ADA/WHO-2hPG definition, 15.4% on the WHO-FPG definition and 43.5% based on the ADA-FPG definition [42]. If any of the cut-offs were met, the prevalence was 51.3%, and the prevalence was 2.5% if all criteria were met.

Complications of prediabetes

Prediabetes increases the risk to develop T2D and CVD. A systematic review of 103 prospective studies up to 2018 from the Cochrane library assessed the overall prognosis of developing T2D [43]. The estimated cumulative T2D incidence over 5 years of follow-up was 50% (95%CI 37-63%) based on the combination of IFG and IGT, and 38% (95%CI 26-51%) based on IEC-HbA1c definition. A meta-analysis of 129 prospective studies showed that prediabetes compared to normoglycaemia was associated with an increased risk of CVD with a relative risk of 1.15 (95% confidence interval [CI] 1.11-1.18) with a median follow-up time of 9.8 years [44]. Effect estimates for coronary artery disease (CAD) and stroke were similar. Furthermore, a Mendelian randomization (MR) study of 1,326,915 participants showed that prediabetes is likely causally linked with CAD, with a 26% higher odds of CAD per mmol/L increase in FPG [45]. Additionally, early stages of typical complications of T2D, including retinopathy, nephropathy, and neuropathy, have been reported among people with prediabetes, as well as other disorders associated with T2D including periodontal disease, cognitive dysfunction, obstructive sleep apnoea, metabolic syndrome, fatty liver disease, and cancer [46].

Public health implications of prediabetes

Recognizing individuals with prediabetes can have significant public health implications. Identifying and screening for prediabetes could help preventive efforts and treatment [47]. Effective preventive strategies are essential, as interventions targeting the reversal of abnormal glucose levels and prevention of complications are most effective when initiated during the early stages. Besides screening, education, community programs, and policy interventions, lifestyle modification is one of the key preventive strategies to prevent prediabetes [42]. Landmark clinical trials have shown that intensive lifestyle modification (i.e., dietary changes and increased physical activity) among people with prediabetes can help to prevent T2D, reducing the incidence of T2D with 25-58% over 3–6-year periods compared to placebo or standard of care groups, mainly by inducing weight loss [48-51].

Moreover, considering the high prevalence of prediabetes, minor adjustments to risk factors could have a profound impact on population health. This perspective aligns with the public health approach to preventive medicine as defined by Geoffrey Rose, which emphasized shifting population risk exposure toward a lower mean, through alterations in environmental conditions that contribute to increased risk [121].

In the Netherlands, there has been limited emphasis on screening for or treating prediabetes specifically in national health initiatives or guidelines. People with elevated blood glucose levels are generally monitored by general practitioners (GPs), who may offer lifestyle advice or refer them to combined lifestyle intervention (GLI) programs aimed at managing overweight and obesity [52]. Diagnosis of individuals with prediabetes is at this moment primarily done in the context of research studies rather than routine

clinical practice, possibly due to its asymptomatic nature and varying risk outcomes, as well as limited treatment capacity.

Prediabetes in epidemiological studies

In research, distinguishing individuals with prediabetes from normoglycaemia and T2D allows for the comparisons of more homogenous groups. Risk factors of early onset of T2D among normoglycaemia can be better understood by excluding individuals with prediabetes at baseline. This approach also helps to remove a potential source of heterogeneity, as associations of dairy intake may vary depending on the level of glycaemic disturbances [53], for example due to varying levels of insulin sensitivity, metabolic disturbances, and differences in body composition. A comparison between the use of prediabetes versus glycaemic markers as outcome measures in epidemiological studies on the association between dairy intake and hyperglycaemia is provided in **Table 4**.

Table 4. Comparison of outcome measures in epidemiological studies assessing dairy intake and hyperglycaemia: prediabetes versus glycaemic markers.

	Prediabetes outcome	Glycaemic markers
Outcome definition	Definition based on a combination of glycaemic markers (i.e., FPG, 2hPG and/or HbA1c).	Many glycaemic markers have been assessed, including those of insulin sensitivity and resistance.
Distribution	Binary	Continuous
Effect estimate	Odds Ratio (OR), Relative Risk (RR) or Hazard Ratio (HR) with 95%CI	β with 95%CI
Meaning	Quantification of ratio of probability between exposed versus unexposed or lower exposure group.	Magnitude and direction of the association.
Advantages	• Provides insights into the preventive associations of certain exposures on the onset of prediabetes. • Allows for the comparisons of more homogenous groups. Helps removing heterogeneity in baseline levels. • Can be applied to populations with different disease prevalences aiding comparison of exposure effects across populations, therefore often used in prospective cohort designs. • Clear clinical relevance, directly translatable to dietary guidelines.	• Early detection of changes in glycaemic markers, even before early-disease states, providing insights on maintenance of normal glycaemic control and preventing development of metabolic disorders. • High sensitivity, small variations are captured, therefore often used as outcome in intervention studies. • Repeated measures may provide insights into glycaemic variability and long-term trends.

Table 4. Comparison of outcome measures in epidemiological studies assessing dairy intake and hyperglycaemia: prediabetes versus glycaemic markers (continued).

	Prediabetes outcome	Glycaemic markers
Disadvantages	• Effect estimates may be affected by the exact definition used. ◦ Intra-individual variability in glycaemic markers introduces noise. ◦ Cut-offs have varying sensitivity and specificity.	• Effect estimates are difficult to compare between studies as heterogeneity in baseline levels have implications for strength and direction of effect estimates. • Intra-individual variability in glycaemic markers introduces noise and make it challenging to detect meaningful changes.

Based on FPG and fasting insulin, as well as on the 2hPG, various indices for insulin sensitivity and resistance have been proposed to predict T2D development in non-diabetic populations in clinical and epidemiological studies [54]. Especially the homeostatic model of insulin resistance (HOMA-IR) is frequently used to quantify insulin resistance and β-cell function [55]. This model correlates well with the hyperinsulinemic-euglycaemic clamp technique (r = 0.88), accepted as the gold standard for assessing insulin sensitivity. Thereby, studies using indices of insulin sensitivity and resistance contribute to our understanding of T2D pathogenesis.

Dairy and hyperglycaemia

As dairy foods are widely consumed and contain many beneficial but also detrimental nutrients for (cardio)metabolic health, it may be an important factor contributing to prediabetes development. Currently, no randomized controlled trials (RCT) examine the effects of dairy intake on prediabetes. Regarding glycaemic markers, RCTs suggest inconsistent effects of high compared to low dairy intake [1-7]. Observational evidence mainly comprises cross-sectional studies with prediabetes or continuous glycaemic markers as outcome, or prospective cohort studies with T2D as outcome. In cross-sectional studies, the exposure and outcome are assessed simultaneously and thereby causality cannot be established. When the outcome precedes the exposure measurement, it introduces the possibility of reverse causality. In such cases, the outcome itself may have influenced health and dietary behaviours, as well as the reporting of these behaviours. Therefore, a major advantage of prospective studies compared to cross-sectional studies is that the exposure assessment proceeds the development of the outcome, limiting the risk for reverse causality. Nevertheless, before conduction of the studies in this dissertation, only one prospective cohort study examined the associations between dairy intake and prediabetes [53]. The observational evidence is summarized in the following paragraphs. Questions remain about the type and dosage of dairy linked to the development of prediabetes, the nature of dose-response relationships, and potential confounding factors related to health behaviours and food intake.

Dairy intake and prediabetes risk in cross-sectional studies

Four cross-sectional studies were performed recently on the association of dairy and prediabetes, with varying results (mostly neutral, inverse and some positive associations) (**Table 5)** [56-59]. A study in the Dutch Maastricht cohort found that higher consumption of skimmed dairy, fermented dairy, yogurt, and cheese was associated with lower odds of prediabetes, but high-fat dairy was not related to prediabetes [56]. In line with these results, inverse associations for skimmed, fermented dairy and prediabetes were found in the Lifelines cohort [57]. This study also found positive associations for high-fat dairy, non-fermented dairy, and custard. In contrast to these inverse associations for low-fat dairy, in the Feel4Diabetes study low intake (0-1 servings/d) compared to high intake (≥1 serving/day) of low-fat dairy was associated with lower odds of prediabetes [58]. Another study in the German KORA-FF4 found no associations between total dairy, milk, yogurt and cheese and the odds of prediabetes in the fully adjusted models [59].

Dairy intake and continuous glycaemic markers in cross-sectional studies

Twelve cross-sectional studies regarding the associations between dairy intake and glycaemic outcomes from observational studies are summarized in **Table 6**. Similar to prediabetes, neutral, inverse and some positive associations were reported. The ELSA-Brasil study including 10,010 participants found inverse associations of several dairy types with FPG, 2hPG, HbA1c, fasting insulin, and HOMA-IR, with strongest inverse associations for cheese with 2hPG, and yogurt with HbA1c [60]. Nevertheless, also desserts exhibited strong inverse associations with 2hPG in this study. Furthermore, inverse associations were found for high-fat dairy intake with fasting insulin and HOMA-IR in a Japanese cohort [61], and for yogurt intake with FPG and HOMA-IR in a UK cohort [62]. Three studies reported null associations of dairy types with FPG, specifically for total, low-fat, high-fat dairy [63], cheese [64] and yogurt and dairy desserts [65]. Three studies reported positive associations between milk intake and glycaemic outcomes, specifically of milk and HbA1c [66], milk and HOMA score [67], and high-fat milk and FPG [68]. However, no associations of milk with FPG were found in three Danish cohorts including 98,529 participants [69] and in four cohorts in the US and Spain including 7,177 participants [70].

Table 5. Cross-sectional studies reporting associations between dairy product intake and prediabetes

Author, year	Cohort, baseline, and location	N, female %	Mean age, y	Mean BMI, kg/m2	Dietary assessment	Dairy types	Prediabetes assessment	Cases	Adjustments	Main results	Conclusion
Eussen, 2015 [56]	Maastricht cohort, 2010-2015, the Netherlands	2,266, 52.9%	58.9	26.0	253-item FFQ.	Total-, high-fat-, semi skimmed-, skimmed-, fermented-, non-fermented (basic and desserts) dairy, milk, cheese (total, Dutch, foreign), yogurt (natural, fruit, drink), curd cheese (total, natural, fruits).	FPG 6.1 to 6.9mmol/l 2hPG ≥7.8 and <11.1mmol/l	470	Age, sex, BMI, PA, smoking, education, and intakes of energy, vegetables, fruits, meat, and fish.	Skimmed dairy (OR_{T3} 0.73, 95%CI 0.55-0.96), fermented dairy (0.74, 0.54-0.99), yogurt (0.67, 0.50-0.90). Cheese (OR per 20 g/d 0.89, 0.81-0.97), Dutch cheese (0.88, 0.80-0.97).	High yogurt and fermented dairy intake were associated with lower odds of prediabetes and newly diagnosed T2D. Associations of total-, full-fat-, skimmed dairy and Dutch cheese were not in similar direction and magnitude for prediabetes and newly diagnosed T2D.
Brouwer-Brolsma, 2018 [57]	Lifelines, 2006-2013, the Netherlands	112,086, 59%[1]	45	25.6	110-item FFQ (heart).	Total, high-fat, skimmed, semi-skimmed, fermented, non-fermented, milk (high-fat, semi-skimmed, skimmed), yogurt (high -fat, skimmed), buttermilk, curd cheese/quark, flavoured yogurt drinks, custard, cheese (high-fat, low-fat, Dutch)	FPG 5.6-6.9 mmol/L or HbA1c 5.7-6.4%	25,549	Age, sex, alcohol, smoking, education, PA, energy intake, energy-adjusted bread, pasta, rice, potato, fruit, vegetables, legumes, meat, fish, coffee, tea, soda/fruit juice, other dairy products, BMI, waist circumference	Inverse associations for skimmed dairy ($OR_{100\ g/d}$ 0.98, 95%CI 0.97-1.00, OR_{T3} 0.95, 0.92-0.99), fermented dairy ($OR_{100\ g/d}$ 0.98, 0.97-0.99, OR_{T3} 0.94, 0.90-0.98), and buttermilk (OR_{150g} 0.97, 0.94-1.00). Positive associations for high-fat dairy ($OR_{100\ g/d}$ 1.03, 1.01-1.06, OR_{T3} 1.10, 1.06-1.15), non-fermented dairy ($OR_{100\ g/d}$ 1.01, 1.00-1.02, OR_{T3} 1.05, 1.00-1.09), and custard (OR_{150g} 1.13, 1.03-1.24, OR_{T3} 1.05, 1.01-1.10).	Inverse associations of skimmed and fermented dairy, buttermilk and low-fat cheese and positive associations for high-fat dairy products, non-fermented dairy products, custard, high-fat milk, and high-fat yogurt with pre-diabetes were observed. The observed associations for dairy product intake and ND-T2D were less convincing, but positive associations for high-fat dairy products non-fermented dairy products, total milk and skimmed milk were found.

Table 5. Cross-sectional studies reporting associations between dairy product intake and prediabetes (continued)

Author, year	Cohort, baseline, and location	N, female %	Mean age, y	Mean BMI, kg/m2	Dietary assessment	Dairy types	Prediabetes assessment	Cases	Adjustments	Main results	Conclusion
Breuninger, 2018 [59]	KORA FF4, 2013-2014	1,332, 53.2%	57.1	26.5	Repeated 300-item 24h food list for consumption probability and a 148-item FFQ for usual portion sizes.	Total dairy, milk, yogurt, cheese	FPG 5.6-6.9 mmol/l or 2hPG 7.8-11.0 mmol/l	545	Age, sex, energy intake, BMI, waist circumference, family history of diabetes, PA, smoking, education, and hypertension.	In the minimally adjusted models, total dairy (OR_{SD} 0.86, 0.75-0.98) and yogurt intake (OR_{SD} 0.87, 0.77-0.99) were associated with prediabetes, but not in fully adjusted models.	No conclusion specifically for dairy intake.
Kontochristo-poulou, 2022 [58]	Feel4Diabetes study, 2016-2018, Bulgaria, Hungary, Belgium, Finland, Greece, and Spain	2,816, 66.8%	40.8	28.5	Self-reported standardized questionnaire	High-fat dairy, low-fat dairy	FPG 5.6-6.9 mmol/l	NR	Age, sex, region, education, occupation, marital status, weight status, weight circumference, fruits and berries, vegetables, legumes, refined grains, whole grains, red meat, white meat, fish, nuts and seeds, salty snacks, sweet snacks, other dairy intake, coffee, SSBs, NSBs, alcohol, PA, screen time, sleep hours	Low-fat dairy (OR 0.5-1 servings/d vs. >1 serving/d 0.49, 95%CI 0.24-0.96) and OR 0-0.5 servings/d 0.56, 0.33-0.94).	Low intake (0-0.5 servings or 0.5-1 servings of 240g) of low-fat dairy was associated with lower odds of prediabetes compared to high intake.

[1] Including 1,305 participants with newly diagnosed diabetes.

Abbreviations: 2hPG, 2-hour plasma glucose; BMI, body mass index; FFQ, Food Frequency Questionnaire; g/d, gram per day; HbA1c, glycated haemoglobin; IFG, impaired fasting glucose; PA, physical activity; ND, newly diagnosed; NSB, non-sugary beverages; NR, not reported; SSB, sugar sweetened beverages; T2D, type 2 diabetes.

Table 6. Cross-sectional studies reporting associations between dairy product intake and glycaemic outcomes

Author, year	Cohort, baseline, and location	N, female %	Mean age, y	Mean BMI, kg/m2	Dietary assessment	Dairy types	Glycaemic markers	Adjustments	Main results	Conclusion
Gulliford, 2001 [66]	Health Survey for England, 1994, United Kingdom	9,772, 52.1%	46.6	NR	Eating habits questionnaire	Milk	HbA1c	Age, sex, BMI, waist-hip ratio, PA, education, smoking, alcohol	Semi-skimmed vs. skimmed milk (0.06%, 95%CI 0.01-0.10). Whole vs. skimmed milk (0.09%, 95CI 0.04-0.14).	Frequent consumption of fat-containing foods was associated with higher HbA1c.
Lawlor, 2005 [67]	British Women's Heart and Health Study, 1999-2001, United Kingdom	4,286, 100%	60-79	27.6	FFQ.	Milk	HOMA score	Age	Milk drinkers vs. non milk drinkers (mean HOMA 1.72, 95%CI 1.70-1.74 vs. 1.49, 1.31-1.70).	Women who do not drink milk have reduced odds of the metabolic syndrome compared with milk consumers.
Akter, 2013 [61]	Japanese working population in Kyushu, 2009, Japan	496, 39%[1]	44.6[1]	22.4[1]	56-item brief diet history questionnaire validated against 16-d weighted dietary records (dairy intake $r = 0.70$).	Total, high-fat, low-fat dairy.	FPG, FI, HOMA-IR	Age, sex, workplace, sedentary work, non-occupational PA, smoking, alcohol, BMI, parental history of diabetes, intake of energy, fibre, PUFA/SFA and calcium.	High-fat dairy intake was associated with FI (P_{trend} 0.02), with means 4.52 μU/mL (95%CI 4.22-4.86), 4.50 (4.19-4.84), 4.40 (4.06-4.77), and 3.71 (3.34-4.12), and with HOMA-IR (p=0.02) with means 1.04 (0.96-1.12), 1.04 (0.96-1.13), 1.00 (0.91-1.08), and 0.86 (0.76-0.96) for categories never, ≤1 serving/wk, 2-6 servings/wk, and ≥1 serving/d, respectively.	The consumption of full-fat dairy products may be associated with lower IR among Japanese adults.
Wang, 2013 [62]	Framingham Heart Study Offspring, 1998-2001, and Third Generation cohorts, 2002-2005, United Kingdom	6,526, NA	49.7	27.4	126-item FFQ. Validated.	Yogurt	FPG, FI, HOMA-IR	Age, sex, diet quality, energy intake, smoking, supplement use and BMI.	Yogurt consumers had lower FPG compared to nonconsumers (97.5 mg/dL, 95%CI 96.8-98.2 vs 98.4, 97.7-99.1, P=0.02). The high yogurt intake group had lower FPG (97.1, 96.3-98.0 vs. 97.8, 96.9-98.9, P=0.01) and HOMA-IR (3.28, 3.20-3.36 vs. 3.34, 3.27-3.42, P=0.04) compared to the low-intake group.	No conclusion specifically for FPG.

Table 6. Cross-sectional studies reporting associations between dairy product intake and glycaemic outcomes (continued)

Author, year	Cohort, baseline, and location	N, female %	Mean age, y	Mean BMI, kg/m2	Dietary assessment	Dairy types	Glycaemic markers	Adjustments	Main results	Conclusion
Crichton, 2014 [63]	ORISCAV-LUX survey, 2006-2009, Luxembourg	1,352, 55.4%	44.6	26.6	134-item FFQ	Total, low-fat, high-fat dairy	FPG	Age, sex, education, and total energy intake.	No associations with FPG were found.	No conclusion specifically for FPG.
Sadeghi, 2014 [64]	Isfahan Healthy Heart Program, 1999, Iran	9,526, 50%	38.8	25.7	Validated 49-item FFQ.	Cheese	FPG	-	No difference in FPG was found between high (≥ 7 servings/d) vs. low cheese intake (<7).	No conclusion specifically for FPG.
Sun, 2014 [68]	Guangzhou Biobank Cohort Study, 2003-2006, China.	20,335, 71.2%	62.7	23.8	Validated FFQ.	High-fat milk	FPG	Age, sex, study phase, SEP, smoking, alcohol, PA, adiposity.	FPG (β >3 vs. 0 servings/d 0.08 mmol/L, 95% CI 0.01 to 0.16, P_{trend} = 0.048).	High-fat milk intake was heterogeneously associated with CVD risk factors. Higher intake was associated with lower BP and TAG, but higher HDL-c and FPG, which might suggest risk factor specific biological pathways.
Bergholdt, 2015 [69]	CCHS (1991-994), CGPS (2003), and GESUS (2010-2013), Denmark.	98,529, 55.2%	57	NA	FFQ	Milk	FPG	Age, sex, birth year, population, height, glucose-lowering therapy	No associations with FPG were found (per glass/week 0.0007 mmol/L, -0.0002, 0.0017, P = 0.12).	No evidence was found of an observational or genetic association of milk intake with risk of IHD or MI in the Danish general population.
Drehmer, 2015 [60]	ELSA-Brasil, 2008-2010, Brazil	10,010, 54.3%	51.6	26.7	114-item FFQ	Total, high-fat, low-fat, milk, cheese, yogurt, cream cheese, dessert, and fermented dairy	FPG, 2hPG, HbA1c, FI, PI, HOMA-IR	Age, sex, race, occupational status, education, family income, menopause, family history of diabetes, smoking alcohol, PA, intake of energy, fruit, vegetables, whole grains, sodas, coffee, red meat, eggs, non-dairy desserts, and sweets.	Total dairy (per serving/d) was associated with lower FPG (mean -0.46 mg/dL, SD = 0.19), 2hPG (-1.25 mg/dL, SD = 0.49), HbA1c (-0.02%, SD = 0), PI (-1.52, SD = 0.60) and HOMA-IR (-0.04, SD = 0.02). Fermented dairy was associated with lower FPG (mean -0.24 mg/dL, 95%CI -0.46; -0.02) and 2hPG (mean -0.86 mg/dL, 95%CI -1.42; -0.30). Full-fat dairy (-0.64 mg/dL, -1.16; -0.12), cheese (-0.92, -1.55; -0.29), desserts (-2.45, -4.65; -0.23) were associated with lower 2hPG and yogurt with lower HbA1c (-0.04%, -0.06; -0.01).	Total dairy intake to be inversely associated with measures of glycaemia and insulinemia, as well as with the presentation of newly diagnosed diabetes. Fermented milk products, such as cheeses and yogurts, had the strongest inverse association with glycaemia.

Table 6. Cross-sectional studies reporting associations between dairy product intake and glycaemic outcomes (continued)

Author, year	Cohort, baseline, and location	N, female %	Mean age, y	Mean BMI, kg/m2	Dietary assessment	Dairy types	Glycaemic markers	Adjustments	Main results	Conclusion
Smith, 2016 [70]	BPRHS (2004-2009), GOLDN (2002), PREDIMED (2003-2009) and WHI (1993), United States and Spain.	1,232, 70.2%; 817, 50.6%; 3,886, 57.3%; 1,242, 30.8%[2]	57.2; 49.0; 67.0; 60.8	31.8; 28.5; 29.9; 30.3	Validated FFQs (122 to 137-item).	Milk	FPG	Age, sex, field centre or ancestry (BPRHS, WHI), family (GOLDN), BMI, smoking, drinking, PA, diabetes, medication, and energy intake.	No associations with FPG were found. β 100g/d 0.13 mg/dL, SE = 0.52; β 100g/d 0.16 mg/dL, SE = 0.16; β 100g/d 0.30 mg/dL, SE = 0.20; β 100g/d 0.65 mg/dL, SE = 0.62.	Milk intake was not associated with CVD biomarkers, CVD, or mortality. Although MCM6-rs3754686 is a good milk intake proxy in these populations, attributing its associations with CVD and mortality in Mediterranean women to milk is unwarranted, as other factors limiting the assumption of causality in Mendelian randomization may exist.
Wade, 2021 [65]	Maine-Syracuse Longitudinal Study	915 (351 for FPG analysis), 60%	62.1	29.4	FFQ	Yogurt and dairy desserts	FPG	Cottage cheese, cakes/pudding/pie, ice cream, age, sex, education, PA, alcohol, total intake, smoking, food score, diabetes, homocysteine, TAG, HDL, BMI, antihypertensive and cholesterol medication	No association with FPG was found (β -0.56, 95%CI -1.62; 0.49)	No conclusion specifically for FPG.

[1] Pooled data based on lowest and highest total of dairy intake (*n* = 236).
[2] %female based on complete sample, *n* = 1,244; 817; 7,185 and 10,843, respectively.
Abbreviations: 2hPG, 2-hour plasma glucose; BMI, body mass index; CVD, cardiovascular disease; d, day; FFQ, Food Frequency Questionnaire; g/d, gram per day; FI, fasting insulin; HbA1c, glycated haemoglobin; HDL, high density lipoprotein; HOMA, homeostatic model assessment; IFG, impaired fasting glucose; IR, insulin resistance; PA, physical activity; PI, plasma insulin; PUFA, polyunsaturated fatty acid; NAFLD, non-alcoholic fatty liver disease; ND, newly diagnosed; NSB, non-sugary beverages; NR, not reported; SFA, saturated fatty acids; SD, standard deviation; T2D, type 2 diabetes; TAG, triacylglycerol; wk, week.

Dairy intake and T2D risk in prospective cohort studies

Several systematic reviews and meta-analyses have published evidence on the prospective association of dairy intake and risk of T2D [71-84]. An overview of summary estimates from systematic reviews and meta-analyses based on a review from Alvarez-Bueno *et al.* (2019) [85] updated with three additional meta-analyses [81, 83, 84] is provided in (**Table 7**).

Table 7. Summary of evidence from meta-analyses on prospective observational associations between dairy intake and T2D, derived from the overview by Alvarez-Bueno *et al.* (2019) [85] updated with the additional meta-analyses [81, 83, 84].[1]

Dairy type	High vs. low intake		Dose-response	
	N meta-analyses	RR range	N meta-analyses	RR range
Total dairy	7 (6 significant)	0.81-1.00	5 (4 significant)	0.88-0.98 per 200-400 g/d
High-fat dairy	4 (0 significant)	0.95-1.00	4 (0 significant)	0.95-0.98 per 200 g/d
Low-fat dairy	4 (4 significant)	0.81-0.83	4 (3 significant)	0.88-0.97 per 200 g/d
Fermented dairy	3 (2 significant)	0.88-0.94	2 (1 significant)[2]	0.92-0.98 per 200 g/d
Milk	6 (3 significant)	0.82-1.12	4 (2 significant)	0.83-1.27 per 200 g/d
High-fat milk	4 (1 significant)	0.87-1.12	3 (0 significant)	0.99-1.27 per 200 g/d
Low-fat milk	2 (2 significant)	0.82 (0.69-0.97)[3]	3 (1 significant)	0.83-1.01 per 200 g/d
Yogurt	7 (7 significant)	0.74-0.86	5 (3 significant)	0.78-0.94 per 50-200 g/d
Cheese	6 (3 significant)	0.82-0.94	5 (2 significant)	0.80-1.00 per 10-50 g/d

[1] Other meta-analyses identified but not included were Mohan *et al.*, (2023) [86] and Companys *et al.* (2020) [87] including fewer prospective cohort studies compared to earlier meta-analyses and Mishali *et al.* (2019) [82] pooling the estimates of total dairy and milk for studies that reported these outcomes separately.
[2] Derived from a study presenting two estimates, for low-fat or high-fat fermented dairy separately [79].
[3] Same estimates based on both meta-analyses, RR (95%CI) [85].
Abbreviations: g/d, gram per day; RR, relative risk.

Additionally, the results of two recent meta-analyses are summarized here as they were similar in search strategy and inclusion criteria, but differed in statistical analysis, included studies, cases, and results. The systematic umbrella review and meta-analysis by Giosuè *et al*. (2022) [88] including studies up to December 2021 is summarized in **Table 8**. Gijsbers *et al.* (2016) [79] updated by Soedamah-Muthu and de Goede (2018) [80] included studies up to July 2018 in **Table 9**. Gijsbers *et al.* (2016) included more studies and/or cases for some dairy types than Giosuè *et al.*, (2022) (e.g., for total milk, 8,061 vs. 17,241 cases) and explored nonlinear associations. For cheese, the most recent and complete meta-analysis was conducted by Zhang *et al.* (2023) [84].

Table 8. Summary of evidence on prospective observational associations between dairy intake and type 2 diabetes from the systematic umbrella review of meta-analyses by Giosuè *et al.* (2022) [88].

Dairy type	N studies	N cases (N total not reported)	Intake	Relative risk (95%CI)	Evidence quality based on NutriGrade[1]
Total dairy	21	42,204	200 g/day	0.95 (0.92-0.98)	Low
High-fat dairy	14	28,817	200 g/day	0.98 (0.93-1.03)	Low
Low-fat dairy	15	29,023	200 g/day	0.97 (0.93-1.00)	Low
Milk	11	8,061	200 g/day	0.90 (0.83-0.98)	Low
Yogurt	10	37,223	100 g/day	0.94 (0.90-0.98)	Moderate
Cheese	10	9,479	30 g/day	0.97 (0.91-1.04)	Moderate

NutriGrade is a scoring system for the quality of evidence, i.e., the thrust in the summary estimate, based on risk of bias/quality assessment, precision, heterogeneity, directness of evidence, publication bias, funding bias, effect size and dose-response gradient [89].

Table 9. Summary of evidence on prospective observational associations between dairy intake and type 2 diabetes from the systematic review and meta-analysis by Gijsbers *et al.* (2016) updated by Soedamah-Muthu and Goede (2018) [79, 80].

Dairy type	N studies	N cases/ N total	Intake	Relative risk (95%CI)	Hetero-geneity (I^2)	Nonlinear relative risk (95%CI)
Total dairy	21	46,905/5,741,718	200 g/day	0.97 (0.95-1.00)	63%	Linear
High-fat dairy	13	24,034/327,895	200 g/day	0.98 (0.93-1.03)	52%	Linear
Low-fat dairy	16	28,531/5,313,782	200 g/day	0.96 (0.92–1.00)	60%	Linear
Fermented dairy	5	14,311/64,227	200 g/day	0.98 (0.90-1.06)/ 0.92 (0.83-1.03)[1]	56%/ 51%	Linear
Milk	11	17,241/145,472	200 g/day	0.97 (0.93-1.02)	57%	Linear
High-fat milk	9	21,995/336,102	200 g/day	1.01 (0.97-1.05)	72%	Linear
Low-fat milk	7	20,098/267,607	200 g/day	0.99 (0.88-1.11)	84%	Linear
Yogurt	14	37,223/5,184,590	100 g/day	0.94 (0.91–0.97)	69%	At 80 g/day, 0.86 (0.83–0.90)
Cheese	12	32,936/369,697	30 g/day	1.00 (0.99-1.02)	62%	Linear
Cream	5	19,730/258,571	5 g/day	0.99 (0.97-1.01)	34%	Linear
Ice cream	5	19,730/258,571	10 g/day	0.93 (0.89-0.97)	86%	At 10 g/day, 0.81 (0.78-0.85)

[1] Summary estimates including either the estimate for low-fat or high-fat fermented dairy of one included study, respectively.
Abbreviation: CI, confidence interval.

Overall, these meta-analyses show an association between a higher intake of total dairy and a lower risk of T2D, especially for low-fat dairy and yogurt intake. High-fat dairy intake and cream were not associated with T2D. For total milk intake, the evidence is mixed, with the largest meta-analysis showing no association (RR per 200g/day 0.97, 95%CI 0.93-1.02), also not when considering high-fat or low-fat milk intake separately (**Table 9**). Yogurt intake was strongly inversely associated with T2D in all meta-analyses

(**Table 7**). This association was found to be nonlinear with the lowest risk at 80 g/day compared to 0 gram/day (RR 0.86, 95%CI 0.83-0.90), but no additional benefit with higher intakes (**Table 9**). For cheese, a moderate association was found in some studies, but not all. The most complete meta-analysis concluded that high compared to low cheese intake was associated with lower T2D risk (*n* = 25, 44,584 cases among 674,107 participants, 0.93, 0.88-0.98, I^2 = 45%) [84]. Nevertheless, no evidence for a linear or nonlinear dose-response association was found (*n* = 18, 35,449 cases among 394,508 participants, RR per 30 g/day 1.00 95%CI 0.95-1.06, I^2 = 57%), and evidence was graded as low based on the AMSTAR-2 tool [90]. Ice cream intake was non-linearly associated with lower T2D risk, with the lowest risk at 10 g/day compared to 0 g/day (RR 0.81, 0.78-0.85), with no further decrease at higher intakes (**Table 8**). Giosuè et al., (2022) graded the quality of evidence for total, high-fat and low-fat dairy, and milk as low, and for yogurt and cheese as moderate [88] (**Table 8**). Zhang et al. (2023) graded the evidence for cheese as low based on the AMSTAR-2 tool [90].

Dairy intake and prediabetes risk in prospective cohort studies

In 1,884 participants with normoglycaemia at baseline from the Framingham Offspring Cohort Study in the US, 902 cases of prediabetes (48.3%) were identified over a mean follow-up of 10.5 ± 4.1 years [53]. Higher intake of total dairy (HR highest compared with lowest intake 0.61, 95%CI 0.46-0.81, P_{trend} = 0.002), high-fat dairy (HR 0.75, 95%CI 0.47-1.17, P_{trend} = 0.03), and low-fat dairy (HR 0.68, 95%CI 0.51-0.92, P_{trend} = 0.03) were associated with lower prediabetes risk. Nonlinear associations were found for total milk, skim milk and whole milk intake with prediabetes risk, with the greatest risk reductions at moderate intakes (around 1 to <3 servings/week). This study also showed that the associations of dairy intake varied by baseline glycaemic status (i.e., normoglycaemia or prediabetes) for high-fat milk (HR ≥1 serving/week 0.84 and 1.16, respectively) and cheese intake (HR ≥4 serving/week 0.88 and 0.37, respectively). This highlights the importance of considering individual metabolic profiles when assessing the impact of dairy consumption on health outcomes.

Substitutions of dairy intake

The most common approach in nutritional epidemiology is to compare individuals with high intake of food compared to those with low intake, controlling for intake of total energy. Thereby, the estimate represents a joint effect of the investigated food and the substituted food [91]. In substitution analysis, the food that is being substituted is specified, helping to elucidate the health implications of dietary modifications and insights into the optimal composition of diets. A common approach is a leave-one-out model, involving adjustments for all dietary sources and total energy except those being substituted with the exposure.

Studies using substitution analysis in the context of associations between dairy and T2D are scarce. A study by Ibsen *et al.* (2017) in the Danish National Diabetes Register

(*n* = 54,277, 7,137 cases, mean follow-up 15.3 years) examined the relationship between substituting one dairy product subgroup for another at baseline and incidence of T2D [92]. Substituting low-fat yogurt in place of high-fat yogurt was associated with higher T2D incidence ($HR_{serving/day}$ 1.17, 95%CI 1.06-1.29). Substituting high-fat yogurt in place of low-fat milk ($HR_{serving/day}$ 0.89, 95%CI 0.83-0.96), whole-fat milk ($HR_{serving/day}$ 0.89, 95%CI 0.82-0.96) or buttermilk ($HR_{serving/day}$ 0.89; 95%CI 0.81, 0·97) was associated with a lower T2D incidence. Furthermore, in the same cohort (*n* = 39,393), Ibsen *et al.* examined the relationship between substituting one dairy product subgroup for another during 5 years of follow-up and the subsequent 10-year risk of T2D [93]. They found that replacing whole-fat yogurt for milk reduced T2D risk in those aged 56–59 years, while replacing skimmed milk for semi-skimmed milk increased risk among those aged 60–72 years. A study by the same group in the European Prospective Investigation into Cancer (EPIC)-InterAct case cohort, showed that replacing red and processed meat with cheese ($HR_{30\ g/day}$ 0.90, 95%CI 0.83-0.97) or yogurt ($HR_{70\ g/day}$ 0.90, 95%CI 0.86-0.95) was associated with a lower rate of T2D [94]. A study by Stuber *et al.* (2020) using EPIC-NL data found no associations between substitutions among milk and yogurt and T2D risk [95]. No studies have examined the associations of substituting dairy types with prediabetes risk.

Considerations on confounding factors and dairy intake associations

Confounding is a major issue in every type of research. The findings of studies are affected by various parameters related to the aims of the research, such as age, and sex, but also baseline values of the parameters under study and time-dependency. Discrepancies in the associations between different dairy types and health outcomes may be partly explained by various dietary, behavioural, and socioeconomic factors across populations [96]. Dairy intake, and especially yogurt, may function as a marker for overall health behaviours and diet quality [16, 56, 97-99]. Positive correlations have been observed between higher dairy and yogurt intake and health behaviours including non-smoking, higher physical activity and favourable cardiometabolic profile [56, 62, 63, 97-101]. Low-fat cheese intake might relate to overall higher SEP [102], female sex, higher physical activity, and better diet quality [103-106]. While studies examining dairy-hyperglycaemia associations often employ confounder adjustment models to derive independent estimates, the exact impact of these confounders on relationships remains unclear.

Potential mechanisms

It is not clear if the beneficial associations of dairy foods on the development of prediabetes and T2D are a result of indirect benefits on lower adiposity or lipidaemia, or if dairy has direct protective effects on glycaemia [107]. With decreased insulin sensitivity, the β-cells may become dysfunctional or fail to produce enough insulin to overcome insulin resistance, resulting in hyperglycaemia. Obesity is the most important

risk factor for the development of prediabetes and T2D [108], with accumulation of abdominal fat and lower lean body mass both associated with insulin resistance and heightened risk [109]. Adipose tissue dysfunction is a key factor in the onset of obesity-related insulin resistance [109]. In obese individuals, when energy intake surpasses energy expenditure, lipid overflow due to overloaded adipocytes leads to excessive fat storage in non-adipose tissues including the liver, skeletal muscle, and pancreatic islets. This, coupled with inflammation and disruptions in adipokine secretion, can contribute to the development or progression of insulin resistance. Skeletal muscle is the primary driver of glycaemic control, responsible for over 80% of insulin-stimulated postprandial glucose uptake [110]. Reduced sensitivity of skeletal muscle to insulin may appear in the earliest stages of T2D development before the onset of β-cell failure and may occur independently of obesity [111].

The Health Council Netherlands concluded that under ad libitum conditions, a daily intake of three servings of dairy, particularly when additional dairy consumption is advised, leads to a 0.5 kg increase in body weight in adults over six months [112]. This is based on two meta-analyses (based on 14 and 18 RCTs) providing compelling evidence with limited heterogeneity [113, 114]. It is unlikely that weight changes will occur when additional dairy foods are compared under isocaloric conditions [112]. In energy restricted RCTs, dairy foods have been related to weight regulation and body composition. A meta-analysis of 37 RCTs (*n* = 3,007) showed that high dairy diets increased lean mass, while decreasing body weight, body fat and waist circumference [107]. Yogurt intake was associated with a lower risk of obesity, weight gain, and elevated waist circumference based on a meta-analysis of 22 prospective studies [115]. A MR study showed an inverse association of genetically predicted milk intake with increased lean mass (0.52 kg per serving/day, SE = 0.17) but also increased waist circumference (1.33 cm per serving/day, SE = 0.62) [116]. In the context of energy-restriction, dairy foods may assist in appetite control exerting favourable effects on body composition [117]. Branched-chain amino acids (BCAA) derived from amino acids may enhance muscle protein synthesis, lean muscle mass, and skeletal muscle metabolic function [118]. Exploring non-obesity-related pathways directly relating dairy intake to disrupted glucose metabolism and insulin resistance is important for gaining a comprehensive understanding of the relationship between dairy and prediabetes development. Several relevant molecular pathways of dairy intake and prediabetes pertain to the effects of dairy proteins, calcium, the milk fat globule membrane (MFGM), dairy fat, probiotics, vitamin K and lactose (**Figure 3**) [11]. However, direct causal molecular mechanisms remain unclear and potential pathways do need further testing in experimental studies.

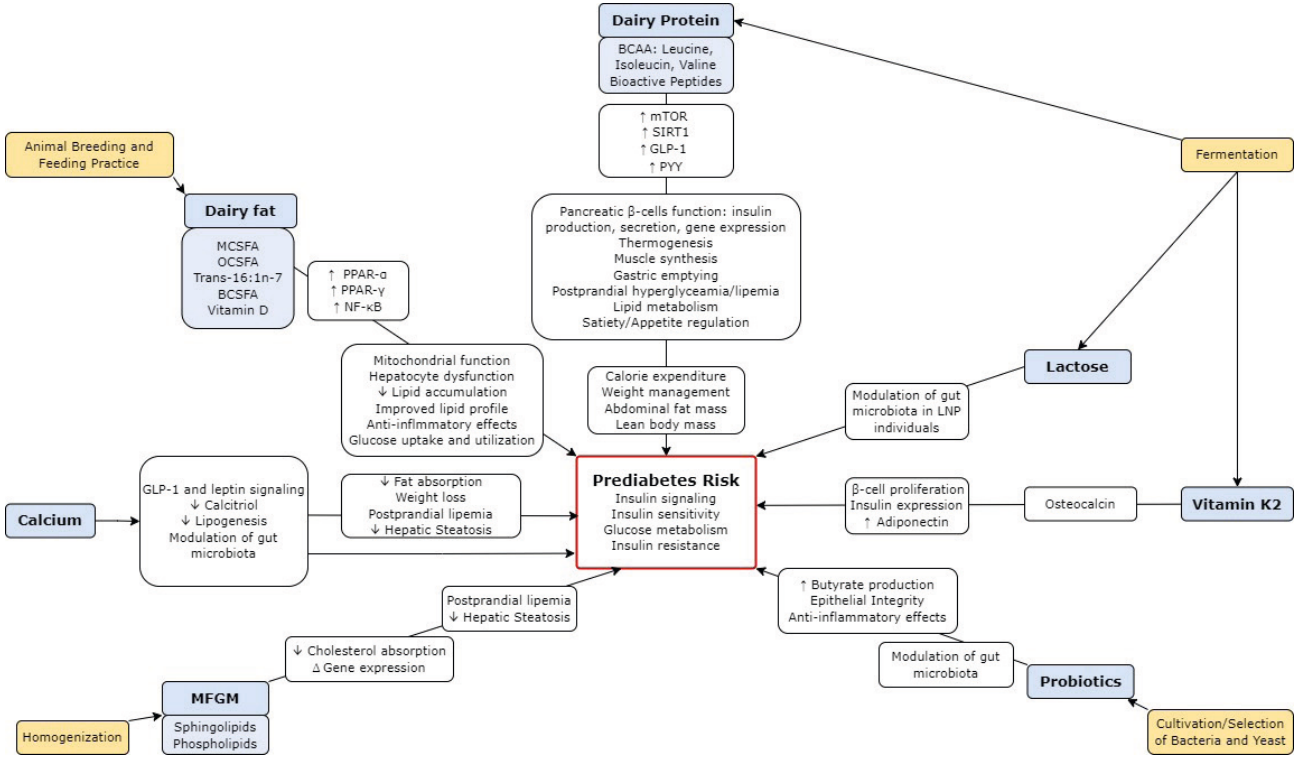


Figure 3. Relevant characteristics of dairy foods, procession methods and selected (molecular) pathways potentially associated to prediabetes risk. Drawing from the review by Mozaffarian *et al.* (2018) [11], supplemented with additional relevant literature, this figure provides an overview of how several factors within dairy products may modulate the risk of developing prediabetes. Abbreviations: BCSFA, Branched-Chain Saturated Fatty Acids; GLP-1, Glucagon-Like Peptide 1; LNP, Lactase non-persistence; mTOR, Mammalian Target of Rapamycin; MFGM, Milk Fat Globule Membrane; NF-kB, Nuclear Factor kappa B; OCSFA, Odd-Chain Saturated Fatty Acids; PPAR- α, Peroxisome Proliferator-Activated Receptor alpha; ;PPAR-γ, Peroxisome Proliferator-Activated Receptor gamma; PYY, Peptide YY; SIRT1, Silent Information Regulator Transcript 1.

Dairy protein

Dairy proteins regulate postprandial glycaemia, affect gastric emptying, regulate lipid changes induced by glucose ingestion, and promote satiety and reduce appetite which can help in weight management [119-124]. Studies indicate that higher dairy protein intake contributes to reducing fat mass while maintaining lean mass in weight loss [125]. Dairy proteins are rich in BCAA leucine, isoleucine, and valine, which have been linked to enhanced thermogenesis and insulin secretion in animal studies, by activating the signalling pathways including mTOR (mammalian target of rapamycin) and SIRT1 (silent information regulator transcript 1) [126, 127]. In vitro, BCAA have been shown to directly impact pancreatic β-cells and promote the release of incretin glucagon-like peptide 1 (GLP-1), slowing down gastric emptying and promoting insulin secretion [128]. Also, several dairy-derived peptides were linked to incretin hormones [119], however relevance might be limited due to their low bioavailability [129]. To prevent the resulting hypoglycaemia after a dairy-rich meal, insulin sensitivity might be reduced. This insulinotropic effect of dairy may explain the short-term effects of high-dairy diets on fasting insulin and insulin sensitivity in RCTs [130, 131]. However, in an RCT in healthy subjects, whey proteins increased the GLP-1 and peptide YY (PYY) secretion without altering insulin secretion [120].

Dairy fats

Dairy fats exhibit complex effects on glycaemia, depending on the exact fatty acid composition of a dairy type. Plasma phospholipid even-chain SFAs (ECSFA) [(myristic acid (C14:0), palmitic acid (C16:0), and stearic acid (C18:0)] have been associated with higher T2D risk [132, 133]. Nevertheless, these ECSFA are also sourced from exogenous non-dairy sources (such as meat, cocoa, and coconut oil) and are produced by de novo lipogenesis which is stimulated by increased intake of carbohydrates and alcohol [134, 135]. Therefore, these positive associations with T2D might only partly represent dairy intake. In contrast, the odd-chain SFAs (OCSFAs) pentadecanoic acid (C15:0), heptadecanoic acid (C17:0) and trans-palmitoleic acid (*t*C16:1n7) have been associated with lower T2D risk [133, 136]. These OCSFAs occur in ruminant milk and have no or little endogenous FA production in the body and have therefore been used as biomarkers of dairy fat intake [137, 138]. Whether they play a functional role in the aetiology of T2D, or if they merely reflect dairy fat intake is yet to be determined [139-141]. Hypothesized biological pathways relating OCSFA to insulin resistance involve facilitating the replenishment of the citric acid cycle, enhancing mitochondrial function, and protect against hepatocyte dysfunction through the promotion of MUFA synthesis by activation of PPAR-α [142]. In animal studies, dietary *t*C16:1n7 as well as BCSFA were related to inhibition of hepatic de novo lipogenesis, activation of PPAR-α and PPAR-γ, improving insulin sensitivity and reducing inflammation [11, 143, 144]. Additionally, medium chain SFAs (MCSFA) have different molecular and metabolic effects compared to long-chain SFA (C16:0 and C18:0). In rodents, MCSFA has been shown to maintain glucose homeostasis during high-fat and energy overfeeding [145]. In a RCT of 17 healthy men,

the substitution of LCSFAs with MCSFAs after overfeeding with LCSFAs resulted in a full reversal of insulin resistance in all tissues, especially in the skeletal muscle [146]. An underlying mechanism might be that MCSFA enhanced mitochondrial oxidative capacity and reduced lipid accumulation in vitro [147], while LCSFA activated NF-κB and decreased insulin sensitivity [148].

Probiotics

Probiotics have been associated with lower weight gain, lower cholesterol, and blood glucose levels in animal models, possibly by compositional and functional changes in the gut microbiome, increased butyrate production and anti-inflammatory effects [11, 149]. Probiotics in some dairy types may affect the viability and composition of the gut microbiota and influence gene expression related to improved glucose metabolism [150, 151]. These viability and compositional changes may include higher epithelial integrity, thereby reducing leakage of lipopolysaccharides into systemic circulation, and reducing low-grade inflammation [152, 153]. The effects depend on the bacterial strains and their amount [154].

Vitamin K

Two major forms of vitamin K exist in food sources: K1 (phylloquinone), which is abundant in green-leafy vegetables and certain vegetable oils, and K2 (menaquinone). Dairy, especially cheese, is a significant source of vitamin K2 due to the utilization of vitamin K2-producing bacteria species in industrial dairy fermentation. Therefore, cheese is a major contributor to vitamin K2 intake in Europe and North America [155, 156]. Both vitamin K1 and K2 have been associated with lower T2D risk [157]. In animal models of T2D, vitamin K2 supplementation showed dose-dependent reductions of HbA1c and FPG and improved insulin resistance and β-cell function [158, 159]. In 68 patients with T2D, vitamin K2 supplementation reduced FPG and HbA1c [160]. Vitamin K may improve insulin sensitivity via several pathways derived from animal models. Vitamin K2 may upregulate carboxylate osteocalcin, improving β-cell proliferation, insulin expression, and upregulation of serum adiponectin levels [161]. Adiponectin, in turn, enhances insulin sensitivity through increased fatty acid oxidation in skeletal muscles and inhibition of hepatic glucose production in the liver [162].

Dairy matrix

Potential health effects of dairy intake are not only derivable to individual nutrients but depend on the matrix of dairy products with specific nutrients, and mutual interactions, possibly further affected by bioactive components. This dairy matrix concept entails the unique structure of a dairy type, combining the single effects and interactions of nutrients and non-nutrient components (e.g. probiotics) on dairy digestion, nutrition absorption and physiological functions relevant to health [6]. This concept is derived from research showing that the health effects of dairy products differ from the effects of individual nutrients. Multiple RCTs demonstrated that isoenergetic substitution of butter

with milk or cheese has a greater impact on rising triglycerides and LDL-cholesterol levels, showing differential effects of SFAs on blood lipids when eaten in different matrices [163-166]. Of particular importance are the calcium content and the bioactive polar lipids in the MFGM. Calcium has been studied for potential anti-obesity effects. Higher calcium intake may inhibit fat absorption by the formation of insoluble "soaps" in the intestine and/or the formation of hydrophobic aggregation, thereby increasing fat exertion and affecting overall energy balance [167, 168]. This also results in a lower postprandial rise in triglycerides and LDL-cholesterol. Some animal studies showed that the effects of calcium on attenuated weight gain, hepatic steatosis and hyperglycaemia and insulin resistance were mediated by correction of leptin and GLP-1 signalling, lower calcitriol levels, suppression of lipogenesis and gut microbiome alterations, but results are mixed [11]. Also, RCTs in humans on the effects of calcium supplementation on weight outcomes are mixed [169-171], suggesting limited relevance of calcium specifically [11]. The MFGM prevents gastric lipase from digesting triacylglycerides (TAG) in the fat globule core, slowing fatty acid release resulting in lower postprandial lipidaemia. Homogenization in the production of milk and yogurt reduces the fat globule size, increases their surface area, and changes the MFGM structure, resulting in a higher susceptibility of the TAG core for gastric lipase [26]. However, gastric coagulation of caseins can enclose fat globules, hindering lipolysis. For most commercially available cheese, homogenized milk is used. The digestion of the TAG core in cheese is further hindered by the semi-solid protein matrix. In mice studies, phospholipid, and sphingolipid supplementation (as present in the MFGM) reduced serum cholesterol and hepatic lipid accumulation through lowering of cholesterol absorption and changes in hepatic gene expression [172, 173]. Stabilization of postprandial lipidaemia and reduced availability of fatty acids for gluconeogenesis can contribute to stabilizing postprandial glucose levels.

Lactose digestion

Lactose consumption results in a lower glycaemic response than would be expected based on its monosaccharide composition due to buffering salts, whey proteins and the caloric value of fat and proteins slowing down gastric emptying, limiting postprandial glucose levels [26]. A meta-analysis showed that most people with a clinical diagnosis of lactose malabsorption can tolerate up to 18 grams/day of lactose without having symptoms, as long the lactose is ingested together with other nutrients and eaten throughout the day [174]. This means that, hypothetically, most of the lactose-intolerant population could consume about 360 ml of milk per day without symptoms [175]. LP individuals can digest lactose and absorb the resulting galactose and glucose molecules in the small intestine. Individuals with lactase non-persistence (LNP) lack the expression of the lactase enzyme in the brush border of their small intestine, which hinders the digestion of lactose from milk in this specific location. In two cohorts in the US and the UK, higher milk intake was associated with lower T2D risk among lactase non-persistent (LNP) individuals (RR 0.70, 95%CI 0.52-0.94 and 0.83, 0.70-0.98, respectively) but not among lactase persistent (LP) individuals (1.19, 0.86-1.64 and 1.00, 0.94-1.05, respectively)

[176]. In LNP individuals, undigested lactose remains in the small intestine and may serve as an energy source for intestinal microbiota, affecting microbiota composition and activity [177]. Alterations in gut microbiota (e.g. *Bifidobacterium*) in LNP individuals may relate to alterations in circulating metabolites resulting in lower T2D risk [176].

Nutritional epidemiology

Nutritional epidemiology is the study of linking intake of nutrients, foods and dietary patterns, and nutritional status, to health outcomes and occurrence of disease in groups of people. Nutritional epidemiology is essential to obtain insights into these relations, aiming to inform policy and guidance, and improve the food supply, health behaviour and population health. Nevertheless, establishing clear associations between diet and specific health outcomes is extremely challenging, due to the individual variability in response to dietary exposures, the multifactorial nature of diseases, and the long latency periods of certain diseases.

A case for observational studies

Evidence-based nutrition aims to integrate the best available evidence with clinical experience [178]. The hierarchy of the evidence pyramid, derived from medicine, reflects the amount of evidence available and the strength expected from each design (**Figure 4**). The pyramid places systematic reviews and meta-analyses of all relevant RCTs, observational studies and mechanistic studies at the top [179], followed by RCTs, and cohort studies.

Figure 4. Hierarchy of evidence pyramid. *Systematic reviews and meta-analyses of RCTs, observational studies and mechanistic studies.

RCTs are often considered to provide the strongest evidence for causal relationships between exposures and health outcomes. This is because randomization reduces bias due to unknown confounding factors, the intervention can be done while minimizing external influence, blinding is possible, temporal sequences can be examined, and results can be replicated in different settings. However, RCTs studying the effect of dietary interventions on hard disease endpoints in humans have several limitations, and are often costly, impractical, and ethically challenging. Dietary exposures are difficult to blind in RCTs, the chosen comparison group and background diet may profoundly affect the outcomes, and health effects may relate to relevant contextual factors and interactions not captured in RCTs. Therefore, RCTs, even if possible, will not provide definite proof for many research questions in nutritional epidemiology [180]. Consequently, for testing of relations of dietary exposures to disease, observational studies are the only viable option.

Observational studies provide valuable insights into potential associations between diet and health outcomes with long latency periods in large groups in real-world settings. The results of observational studies are important for generating hypotheses and informing more rigorous intervention studies. Prospective cohort studies are considered as higher quality of evidence than cross-sectional or case-cohort studies, as the assessment of the exposure precedes the development of disease. This temporal sequence helps to minimize the potential for recall bias and reverse causation. For these reasons, most evidence of associations between dietary exposures and T2D is gathered from prospective cohort studies. To increase the certainty of evidence based on observational studies, careful conduct, and thorough reporting according to current standards are paramount [181]. This includes, amongst others, careful consideration of the target population and study setting with certain dietary intakes and nutritional statuses, clearly defining dietary exposures, and efforts to mitigate biases. Alongside the primary causal study hypothesis, alternative non-causal hypotheses can be considered that may explain how certain biases have led to associations between exposure and disease [180].

With sufficient consistent observational evidence, combined with coherent findings from mechanistic studies establishing biological plausibility, associations become a key point of evidence in the direction of causality, without the need for RCTs to come to any conclusions [180]. In recent years, more emphasis has been placed on observational studies in nutritional recommendations [17].

This thesis

The aim of this thesis was to study the relationship between the intake of total dairy and various dairy types and prediabetes in prospective cohort studies including general populations.

Methods

The first part of this project focussed on the assessment of associations of dairy and prediabetes in The Netherlands. Secondly, our objective shifted towards the inclusion of international cohorts, with the overarching goal of synthesizing all available evidence from prospective cohort studies in a meta-analysis. The prospective cohorts that were used in this thesis were selected based on the following criteria:

- Inclusion of adults aged 18 years or older with normoglycaemia at baseline;
- Utilization of a prospective observational cohort design;
- Collection of continuous glycaemic measures at baseline and follow-up assessments;
- Implementation of a validated dietary assessment method to quantify dairy consumption at baseline, such as an FFQ or diet history;
- Inclusion of a wide range of sociodemographic and health risk factors for a comprehensive analysis.

The prospective cohorts that met our criteria and were available for our secondary analyses included the Dutch Hoorn Studies, the Rotterdam Study, the Lifelines study, the Australian Diabetes, Obesity, and Lifestyle Study (AusDiab), and the Fenland study from the United Kingdom. Limitations of this selection is the restriction to Western high-SES populations. Therefore, we extended invitations to Principal Investigators from the CARDIA study from the United States, the Tianjin Chronic Low-Grade Systemic Inflammation and Health Cohort Study (TCLSIH) in China, and the Korea Genome and Epidemiology Study (KoGES) in South-Korea; however, we were not able to achieve collaboration. Furthermore, we explored potential collaborations with several other cohort studies and the InterConnect project [180] but did not proceed with this because of the limited availability of glycaemic measures.

Cohort	General aim	Baseline (years)	N total cohort
HOORNSTUDIES	Prevalence and risk factors for disturbances in glucose metabolism and diabetes.	HS1: 1989-1992 HS2: 2006-2007	9,733
ROTTERDAM STUDY	Aetiology and risk factors of diseases in elderly.	RS-I: 1989–1993 RS-II: 2000 RS-III: 2006	14,926
lifelines	Biobank enabling research to better prevent, predict, diagnose, and treat diseases.	2006-2013	167,729

Cohort	General aim	Baseline (years)	N total cohort
AusDiab Australian Diabetes, Obesity & Lifestyle Study	Benchmark data on prevalence and examination of the natural history of diabetes, pre-diabetes, heart disease and kidney disease.	1999–2000	11,247
FENLAND STUDY	Interaction between environmental and genetic factors in determining obesity, T2D, and related metabolic disorders.	2005-2015	12,435

Exposure assessment

All cohorts used a validated FFQ at baseline developed specifically for these cohorts to measure dietary intake. We aggregated the items from each FFQ into commonly used categories for dairy intake, including total dairy, fermented dairy, milk and milk products, yogurt, cheese, and ice cream, distinguishing between low-fat and high-fat types. Further distinction into sweetened and unsweetened dairy products was not possible.

The mean dairy intake in each of the included cohorts, and compared to the VCP [27], are shown in **Figure 5**. The FFQ used in each cohort differed considerably and therefore comparison of exact intakes is not possible. The mean dairy intake and distribution of different dairy types to the total intake is similar in Dutch cohorts, with lower milk intake in more recent cohorts. Fermented dairy intake is lower in AusDiab and the Fenland study compared to the Dutch cohorts.

We analysed the dairy types in servings/day according to serving size definitions employed by national health agencies. Thereby, relative risks (RRs) are standardized facilitating comparison across different studies and populations and are in line with national guidelines facilitating integration of research findings. An exception is the Lifelines study, where we expressed intakes in servings/day enforcing equal water content of liquid dairy types, and comparable to high-fat cheese regarding energy content. These latter serving sizes were used in the meta-analysis to allow for a better comparison of the strength and significance of relative risks for the various dairy types.

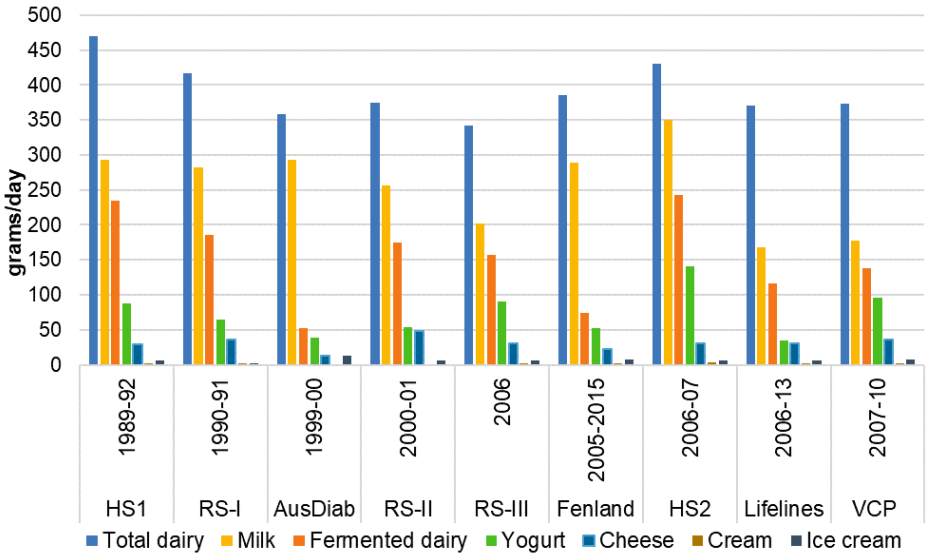


Figure 5. Mean dairy intake in the Hoorn Studies (HS), the Rotterdam Study (RS), the AusDiab study, the Fenland study, the Lifelines cohort, and the Food Consumption Survey Netherlands (VCP) [27] with year of (baseline) assessment shown.

Analysis methods

In each of the selected cohorts, we rigorously assessed our objectives based on a standardized analysis plan to establish a robust evidence base. This standardized methodology contributed to high-quality input for the meta-analysis. This also allowed for deeper insights into the potential sources of heterogeneity across cohorts with specific explorations of confounding factors, mediation effects of obesity and cardiometabolic markers, and effect modification by age, sex, and obesity. Such cross-cohort comparisons and replication analyses are imperative for drawing firmer conclusions based on observational nutritional epidemiological studies.

We employed a novel approach using network models to describe the holistic interrelationships of dietary characteristics, and sociodemographic, health and cardiometabolic risk factors and prediabetes. Network models have a unique advantage of analysing and visualizing all mutual connections simultaneously, providing a comprehensive understanding of the key variables, clustering, and pathways. With this integrative approach, we aimed to uncover insights into the found associations between dairy intake and prediabetes risk, considering the connections of dairy intake to potential confounding factors and their collective influence on prediabetes risk. Nutritional epidemiology often relies on regression analysis for the assessment of diet-disease relationships, offering insights into the independent effect of a single dietary exposure. Nevertheless, regression analysis does not fully capture the complex network

of relationships between variables. Leveraging network modelling techniques, previously employed within the field of nutritional epidemiology to identify dietary patterns [183-186] and to explore relationships between demographics, dietary behaviours, and clinical markers [187], we aimed to offer a more nuanced understanding of these associations. To our knowledge, using holistic network models to interpret the results of reductionist regression analysis represents a novel approach in the field.

We derived this network approach from the psychology field, which adapted social networks (i.e., connections to persons or entities) to networks visualizing connections between mood states, symptoms, or attitudes [188]. In network analysis, the conditional interdependencies between all these variables are estimated using machine learning regularization techniques and visualized together (**Figure 6**). We applied this methodology in two prospective cohort studies, firstly in the Lifelines study due to its large sample size, and secondly as validation in the Fenland study for its definition of prediabetes based on FPG, 2hPG and Hba1c, and its unique inclusion of objective measurements of physical activity within the cohort. Subsequently, we assessed several structural features of the network, including the predictability of variables, clustering of variables, and centrality indices.

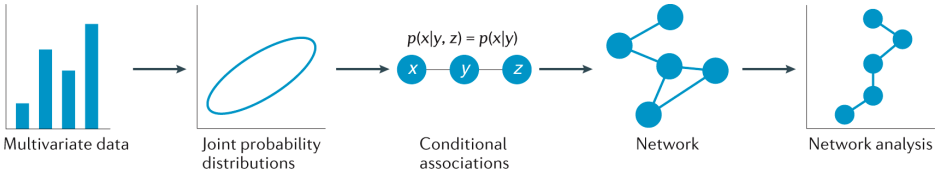


Figure 6. Graphical presentation of steps undertaken in network analysis.

Finally, we identified, summarized, and evaluated all available evidence in a systematic review and meta-analysis. Systematic reviews and meta-analyses are often considered at the top of the hierarchy of evidence and are paramount for informing evidence-based guidelines and clinical practice. We aimed to follow a rigorous and predefined methodology following the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) and Meta-analysis of Observational Studies in Epidemiology (MOOSE) guidelines to reduce bias in the selection and interpretation of studies [189]. We conducted an extensive literature search in several databases on longitudinal studies reporting associations of any dairy intake with continuous glycaemic markers or prediabetes incidence. Articles were selected and relevant information was extracted independently by two researchers. The quality of the individual studies was assessed by the Newcastle-Ottawa Scale, a well-known scoring system for observational studies [190]. Separate meta-analyses were performed for each dairy type in relation to prediabetes. Dose-response random-effects meta-analyses were used to derive incremental dose-response associations and we examined potential non-linear associations using quadratic

and restricted cubic spline models. The quality of each meta-analysis was assessed using the NutriGrade scoring system to clearly communicate the level of confidence in the findings [89]. This scoring system is tailored to characteristics, strengths and limitations, and biases of observational designs in nutritional science.

Outline of this thesis

After the General Introduction, each chapter describes the research in the five analysed prospective cohort studies. **Chapter 2** describes the prospective associations between dairy intake and prediabetes risk in the Hoorn Studies. Furthermore, we examined substituting dairy types with alternative dairy types in relation to prediabetes risk. **Chapter 3** presents the findings of the Rotterdam Study, focusing on the analysis and comparison of associations between dairy intake and incident prediabetes during the follow-up period, as well as dairy intake with repeated measures of insulin resistance. **Chapter 4** describes the association between dairy intake and prediabetes risk in the AusDiab study. In **Chapter 5**, we evaluate the associations of dairy intake and prediabetes risk in the Lifelines study, with a focus on exploring potential sources of heterogeneity of associations in the literature. We adopt a novel approach to investigate the interrelatedness of dairy intake in a network of metabolic risk factors, individual health behaviours and intake of other food groups, to contextualize and interpret the associations between dairy consumption and prediabetes risk. Furthermore, possible reverse causation is examined by relating baseline prediabetes risk and a desire to lose weight to baseline dairy intake, as potential awareness of an individual's risk might relate to certain dietary choices. **Chapter 6** presents the associations between dairy intake and prediabetes risk in the Fenland study. We analysed changes in dairy intake in relation to changes from normoglycaemia to prediabetes and T2D, as well as with changes in glycaemic markers, which has not been done before and may shed light into these complex relationships. Furthermore, we examined mediation by dairy fat biomarkers on the associations between dairy intake and prediabetes, examined the impact of the definition of prediabetes on associations between dairy types and prediabetes, and examined interrelations of dairy intake in a network with confounding factors and prediabetes. In a short ***Intermezzo***, distinct patterns of dairy type intake based on principal component analysis (PCA) in relation to prediabetes risk are examined in the Lifelines and Fenland study. In **Chapter 7**, we provide an overview of the literature and our studies on dairy intake in relation to prediabetes and glycaemic markers in healthy adult populations. In this chapter, we combined all evidence from prospective observational studies in a dose-response meta-regression analysis and graded the quality of evidence. All findings described in this thesis are discussed in the concluding chapter, **Chapter 8**. Furthermore, we discuss methodological considerations of the studies and provide implications and recommendations for future research.

References

1. Eaton, S.B. and D.A. Nelson, Calcium in evolutionary perspective. Am J Clin Nutr, 1991. 54(1 Suppl): p. 281S-287S.
2. Ganmaa, D. and A. Sato, The possible role of female sex hormones in milk from pregnant cows in the development of breast, ovarian and corpus uteri cancers. Med Hypotheses, 2005. 65(6): p. 1028-37.
3. Segurel, L. and C. Bon, On the Evolution of Lactase Persistence in Humans. Annu Rev Genomics Hum Genet, 2017. 18: p. 297-319.
4. Ludwig, D.S., et al., Low-fat diet Redux at WHO. Am J Clin Nutr, 2023. 118(5): p. 849-851.
5. Nickel, K.P., et al., Calcium bioavailability from bovine milk and dairy products in premenopausal women using intrinsic and extrinsic labeling techniques. J Nutr, 1996. 126(5): p. 1406-11.
6. Fardet, A., et al., Influence of food structure on dairy protein, lipid and calcium bioavailability: A narrative review of evidence. Crit Rev Food Sci Nutr, 2019. 59(13): p. 1987-2010.
7. Niedermaier, T., et al., Vitamin D food fortification in European countries: the underused potential to prevent cancer deaths. Eur J Epidemiol, 2022. 37(4): p. 309-320.
8. Yener, S., et al., Seasonal variation in the positional distribution of fatty acids in bovine milk fat. J Dairy Sci, 2021. 104(12): p. 12274-12285.
9. O'Donnell-Megaro, A.M., D.M. Barbano, and D.E. Bauman, Survey of the fatty acid composition of retail milk in the United States including regional and seasonal variations. J Dairy Sci, 2011. 94(1): p. 59-65.
10. Liu, N., et al., Dairy farming system markers: The correlation of forage and milk fatty acid profiles from organic, pasture and conventional systems in the Netherlands. Food Chem, 2020. 314: p. 126153.
11. Mozaffarian, D. and J.H.Y. Wu, Flavonoids, Dairy Foods, and Cardiovascular and Metabolic Health: A Review of Emerging Biologic Pathways. Circ Res, 2018. 122(2): p. 369-384.
12. Brink, E., et al., Development of healthy and sustainable food-based dietary guidelines for the Netherlands. Public Health Nutr, 2019. 22(13): p. 2419-2435.
13. Gao, J., et al., Probiotics in the dairy industry-Advances and opportunities. Compr Rev Food Sci Food Saf, 2021. 20(4): p. 3937-3982.
14. Hill, C., et al., Expert consensus document. The International Scientific Association for Probiotics and Prebiotics consensus statement on the scope and appropriate use of the term probiotic. Nat Rev Gastroenterol Hepatol, 2014. 11(8): p. 506-14.
15. Comerford, K.B., et al., Global Review of Dairy Recommendations in Food-Based Dietary Guidelines. Front Nutr, 2021. 8: p. 671999.
16. Weaver, C.M., How sound is the science behind the dietary recommendations for dairy? Am J Clin Nutr, 2014. 99(5 Suppl): p. 1217S-22S.
17. Dutch Health Council (Gezondheidsraad), Dutch dietary guidelines 2015 (Richtlijnen goede voeding 2015). Publication nr. 2015/24. ISBN 978-94-6281-089-1. The Hague. 2015.
18. Gezondheidsraad. Richtlijnen goede voeding voor mensen met diabetes type 2. Den Haag: Gezondheidsraad 2021; publicatienr. 2021/41.
19. Gezondheidsraad. Richtlijnen goede voeding voor mensen met hart- en vaatziekten door atherosclerose. Den Haag: Gezondheidsraad 2023; publicatienr. 2023/02.
20. Gezondheidsraad. Gezonde eiwittransitie. Den Haag: Gezondheidsraad 2023; publicatienr. 2023/19. .
21. Temme, E.H., et al., Greenhouse gas emission of diets in the Netherlands and associations with food, energy and macronutrient intakes. Public Health Nutr, 2015. 18(13): p. 2433-45.
22. Willett, W., et al., Food in the Anthropocene: the EAT-Lancet Commission on healthy diets from sustainable food systems. Lancet, 2019. 393(10170): p. 447-492.
23. Walther, B., et al., Comparison of nutritional composition between plant-based drinks and cow's milk. Frontiers in nutrition, 2022. 9: p. 2645.
24. Hertzler, S.R., et al., Plant Proteins: Assessing Their Nutritional Quality and Effects on Health and Physical Function. Nutrients, 2020. 12(12).
25. Muleya, M., F.B. E, and H.B. E, A comparison of the bioaccessible calcium supplies of various plant-based products relative to bovine milk. Food Res Int, 2024. 175: p. 113795.

26. Huppertz, T., et al., Dairy Matrix Effects: Physicochemical Properties Underlying a Multifaceted Paradigm. Nutrients, 2024. 16(7): p. 943.
27. van Rossum, C., et al., The diet of the Dutch. Results of the Dutch National Food Consumption Survey 2019-2021 on food consumption and evaluation with dietary guidelines. 2023.
28. Hans, D. and V. David, Melk (producten), boter, kaas en eieren: Consumptie per hoofd van de bevolking in Nederland, 2005-2022. 2023.
29. Autio, M., et al., Towards de-dairyfication of the diet?—Consumers downshifting milk, yet justifying their dairy pleasures. Frontiers in Sustainability, 2023. 4: p. 975679.
30. OECD/FAO (2020), OECD-FAO Agricultural Outlook 2020-2029, FAO, Rome/OECD Publishing, Paris, https://doi.org/10.1787/1112c23b-en.
31. World Health Organization (WHO), Definition and diagnosis of diabetes mellitus and intermediate hyperglycaemia: report of a WHO/IDF consultation. 2006.
32. Group, N.D.D., Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. diabetes, 1979. 28(12): p. 1039-1057.
33. Kahn, R., Follow-up report on the diagnosis of diabetes mellitus: the expert committee on the diagnosis and classifications of diabetes mellitus. Diabetes care, 2003. 26(11): p. 3160.
34. International Expert, C., International Expert Committee report on the role of the A1C assay in the diagnosis of diabetes. Diabetes Care, 2009. 32(7): p. 1327-34.
35. American Diabetes Association Professional Practice, C., 2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes-2022. Diabetes Care, 2022. 45(Suppl 1): p. S17-S38.
36. Rooney, M.R., et al., Global Prevalence of Prediabetes. Diabetes Care, 2023: p. dc222376.
37. Liu, C., et al., Trends in Self-reported Prediabetes and Metformin Use in the USA: NHANES 2005-2014. J Gen Intern Med, 2020. 35(1): p. 95-101.
38. Abdul-Ghani, M.A., D. Tripathy, and R.A. DeFronzo, Contributions of beta-cell dysfunction and insulin resistance to the pathogenesis of impaired glucose tolerance and impaired fasting glucose. Diabetes Care, 2006. 29(5): p. 1130-9.
39. VZinfo.nl (2023): https://www.vzinfo.nl/diabetes-mellitus/leeftijd-en-geslacht, RIVM: Bilthoven. Accessed at 3-2-2024.
40. Diabetes Fonds. Diabetes in cijfers. [cited 2024 15-01]; Available from: https://www.diabetesfonds.nl/over-diabetes/diabetes-in-het-algemeen/diabetes-in-cijfers#:~:text=Hoeveel%20mensen%20hebben%20diabetes%3F,voorfase%20van%20diabetes%20type%202.
41. Stehouwer, C. Onze verhalen: Het grijze gebied genaamd prediabetes. Maastricht UMC+. 2022 [cited 2024 15-01]; Available from: https://www.mumc.nl/actueel/onze-verhalen/het-grijze-gebied-genaamd-prediabetes.
42. Echouffo-Tcheugui, J.B. and E. Selvin, Prediabetes and what it means: the epidemiological evidence. Annual review of public health, 2021. 42: p. 59-77.
43. Richter, B., et al., Development of type 2 diabetes mellitus in people with intermediate hyperglycaemia. Cochrane Database Syst Rev, 2018. 10(10): p. CD012661.
44. Cai, X., et al., Association between prediabetes and risk of all cause mortality and cardiovascular disease: updated meta-analysis. BMJ, 2020. 370: p. m2297.
45. Mutie, P.M., et al., An investigation of causal relationships between prediabetes and vascular complications. Nat Commun, 2020. 11(1): p. 4592.
46. Buysschaert, M., et al., Prediabetes and associated disorders. Endocrine, 2015. 48(2): p. 371-93.
47. Force, U.S.P.S.T., et al., Screening for Prediabetes and Type 2 Diabetes: US Preventive Services Task Force Recommendation Statement. JAMA, 2021. 326(8): p. 736-743.
48. Tuomilehto, J., et al., Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. New England Journal of Medicine, 2001. 344(18): p. 1343-1350.
49. Knowler, W.C., Diabetes Prevention Program Research Group: Reduction in the incidence of type 2 diabetes with life-style intervention or metformin. N. Engl. J. Med., 2002. 346: p. 393-403.
50. Pan, X.R., et al., Effects of diet and exercise in preventing NIDDM in people with impaired glucose tolerance. The Da Qing IGT and Diabetes Study. Diabetes Care, 1997. 20(4): p. 537-44.
51. Ramachandran, A., et al., The Indian Diabetes Prevention Programme shows that lifestyle modification and metformin prevent type 2 diabetes in Asian Indian subjects with impaired glucose tolerance (IDPP-1). Diabetologia, 2006. 49: p. 289-297.

52. Oosterhoff, M., et al., Jaarrapportage monitor gecombineerde leefstijl interventie (GLI) 2023. 2023.
53. Hruby, A., et al., Associations of Dairy Intake with Incident Prediabetes or Diabetes in Middle-Aged Adults Vary by Both Dairy Type and Glycemic Status. J Nutr, 2017. 147(9): p. 1764-1775.
54. Gutch, M., et al., Assessment of insulin sensitivity/resistance. Indian journal of endocrinology and metabolism, 2015. 19(1): p. 160-164.
55. Matthews, D.R., et al., Homeostasis model assessment: insulin resistance and β-cell function from fasting plasma glucose and insulin concentrations in man. diabetologia, 1985. 28: p. 412-419.
56. Eussen, S.J., et al., Consumption of dairy foods in relation to impaired glucose metabolism and type 2 diabetes mellitus: the Maastricht Study. Br J Nutr, 2016. 115(8): p. 1453-61.
57. Brouwer-Brolsma, E.M., et al., Dairy product consumption is associated with pre-diabetes and newly diagnosed type 2 diabetes in the Lifelines Cohort Study. Br J Nutr, 2018. 119(4): p. 442-455.
58. Kontochristopoulou, A.M., et al., Sociodemographic, anthropometric, and lifestyle correlates of prediabetes and type 2 diabetes in europe: The Feel4Diabetes study. Nutr Metab Cardiovasc Dis, 2022. 32(8): p. 1851-1862.
59. Breuninger, T.A., et al., Differential associations between diet and prediabetes or diabetes in the KORA FF4 study. J Nutr Sci, 2018. 7: p. e34.
60. Drehmer, M., 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): p. 775-82.
61. Akter, S., et al., Dairy consumption is associated with decreased insulin resistance among the Japanese. Nutr Res, 2013. 33(4): p. 286-92.
62. Wang, H., et al., Yogurt consumption is associated with better diet quality and metabolic profile in American men and women. Nutr Res, 2013. 33(1): p. 18-26.
63. Crichton, G.E. and A.a. Alkerwi, Dairy food intake is positively associated with cardiovascular health: findings from Observation of Cardiovascular Risk Factors in Luxembourg study. Nutrition research, 2014. 34(12): p. 1036-1044.
64. Sadeghi, M., et al., Cheese consumption in relation to cardiovascular risk factors among Iranian adults- IHHP Study. Nutr Res Pract, 2014. 8(3): p. 336-41.
65. Wade, A.T., et al., Higher yogurt intake is associated with lower blood pressure in hypertensive individuals: Cross-sectional findings from the Maine-Syracuse Longitudinal study. Int Dairy J, 2021. 122.
66. Gulliford, M.C. and O.C. Ukoumunne, Determinants of glycated haemoglobin in the general population: associations with diet, alcohol and cigarette smoking. Eur J Clin Nutr, 2001. 55(7): p. 615-23.
67. Lawlor, D., et al., Avoiding milk is associated with a reduced risk of insulin resistance and the metabolic syndrome: findings from the British Women's Heart and Health Study. Diabetic medicine, 2005. 22(6): p. 808-811.
68. Sun, Y., et al., Milk consumption and cardiovascular risk factors in older Chinese: the Guangzhou Biobank Cohort Study. PLoS One, 2014. 9(1): p. e84813.
69. Bergholdt, H.K., et al., Milk intake is not associated with ischaemic heart disease in observational or Mendelian randomization analyses in 98 529 Danish adults. International journal of epidemiology, 2015. 44(2): p. 587-603.
70. Smith, C.E., et al., Associations of the MCM6-rs3754686 proxy for milk intake in Mediterranean and American populations with cardiovascular biomarkers, disease and mortality: Mendelian randomization. Sci Rep, 2016. 6(1): p. 33188.
71. Elwood, P.C., et al., The survival advantage of milk and dairy consumption: an overview of evidence from cohort studies of vascular diseases, diabetes and cancer. J Am Coll Nutr, 2008. 27(6): p. 723S-34S.
72. Elwood, P.C., et al., The consumption of milk and dairy foods and the incidence of vascular disease and diabetes: an overview of the evidence. Lipids, 2010. 45(10): p. 925-39.
73. Tong, X., et al., Dairy consumption and risk of type 2 diabetes mellitus: a meta-analysis of cohort studies. Eur J Clin Nutr, 2011. 65(9): p. 1027-31.
74. Aune, D., et al., 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): p. 1066-83.

75. Gao, D., et al., Dairy products consumption and risk of type 2 diabetes: systematic review and dose-response meta-analysis. PLoS One, 2013. 8(9): p. e73965.
76. Chen, M., et al., Dairy consumption and risk of type 2 diabetes: 3 cohorts of US adults and an updated meta-analysis. BMC medicine, 2014. 12(1): p. 1-14.
77. Schwingshackl, L., et al., Food groups and risk of type 2 diabetes mellitus: a systematic review and meta-analysis of prospective studies. Eur J Epidemiol, 2017. 32(5): p. 363-375.
78. Tian, S., et al., Dietary Protein Consumption and the Risk of Type 2 Diabetes: A Systematic Review and Meta-Analysis of Cohort Studies. Nutrients, 2017. 9(9): p. 982.
79. Gijsbers, L., et al., Consumption of dairy foods and diabetes incidence: a dose-response meta-analysis of observational studies. Am J Clin Nutr, 2016. 103(4): p. 1111-24.
80. Soedamah-Muthu, S.S. and J. de Goede, Dairy Consumption and Cardiometabolic Diseases: Systematic Review and Updated Meta-Analyses of Prospective Cohort Studies. Curr Nutr Rep, 2018. 7(4): p. 171-182.
81. Fan, M., et al., Dietary Protein Consumption and the Risk of Type 2 Diabetes: ADose-Response Meta-Analysis of Prospective Studies. Nutrients, 2019. 11(11): p. 2783.
82. Mishali, M., et al., Association between dairy intake and the risk of contracting type 2 diabetes and cardiovascular diseases: a systematic review and meta-analysis with subgroup analysis of men versus women. Nutr Rev, 2019. 77(6): p. 417-429.
83. Zhang, K., P. Bai, and Z. Deng, Dose-Dependent Effect of Intake of Fermented Dairy Foods on the Risk of Diabetes: Results From a Meta-analysis. Can J Diabetes, 2022. 46(3): p. 307-312.
84. Zhang, M., et al., Cheese consumption and multiple health outcomes: an umbrella review and updated meta-analysis of prospective studies. Adv Nutr, 2023. 14(5): p. 1170-1186.
85. Alvarez-Bueno, C., et al., Effects of Milk and Dairy Product Consumption on Type 2 Diabetes: Overview of Systematic Reviews and Meta-Analyses. Adv Nutr, 2019. 10(suppl_2): p. S154-S163.
86. Mohan, V., et al., Effect of Milk and Cultured Milk Products on Type 2 Diabetes: A Global Systematic Review and Meta-analysis of Prospective Cohort Studies. Journal of the Indian Institute of Science, 2023. 103(1): p. 167-190.
87. Companys, J., et al., Fermented Dairy Products, Probiotic Supplementation, and Cardiometabolic Diseases: A Systematic Review and Meta-analysis. Adv Nutr, 2020. 11(4): p. 834-863.
88. Giosue, A., et al., Consumption of different animal-based foods and risk of type 2 diabetes: An umbrella review of meta-analyses of prospective studies. Diabetes Res Clin Pract, 2022. 191: p. 110071.
89. Schwingshackl, L., et al., Perspective: NutriGrade: A Scoring System to Assess and Judge the Meta-Evidence of Randomized Controlled Trials and Cohort Studies in Nutrition Research. Adv Nutr, 2016. 7(6): p. 994-1004.
90. Shea, B.J., et al., AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. bmj, 2017. 358.
91. Laursen, A.S.D., et al., Substitutions of dairy product intake and risk of stroke: a Danish cohort study. Eur J Epidemiol, 2018. 33(2): p. 201-212.
92. Ibsen, D.B., et al., Substitutions between dairy product subgroups and risk of type 2 diabetes: the Danish Diet, Cancer and Health cohort. British journal of nutrition, 2017. 118(11): p. 989-997.
93. Ibsen, D.B., et al., Changes in intake of dairy product subgroups and risk of type 2 diabetes: modelling specified food substitutions in the Danish Diet, Cancer and Health cohort. Eur J Nutr, 2021. 60(6): p. 3449-3459.
94. Ibsen, D.B., et al., Replacement of Red and Processed Meat With Other Food Sources of Protein and the Risk of Type 2 Diabetes in European Populations: The EPIC-InterAct Study. Diabetes Care, 2020. 43(11): p. 2660-2667.
95. Stuber, J.M., et al., Substitution among milk and yogurt products and the risk of incident type 2 diabetes in the EPIC-NL cohort. J Hum Nutr Diet, 2021. 34(1): p. 54-63.
96. Mitri, J., et al., Dairy intake and type 2 diabetes risk factors: A narrative review. Diabetes Metab Syndr, 2019. 13(5): p. 2879-2887.
97. Panahi, S., et al., Yogurt consumption, body composition, and metabolic health in the Quebec Family Study. Eur J Nutr, 2018. 57(4): p. 1591-1603.
98. Mena-Sanchez, G., et al., Fermented dairy products, diet quality, and cardio-metabolic profile of a Mediterranean cohort at high cardiovascular risk. Nutr Metab Cardiovasc Dis, 2018. 28(10): p. 1002-1011.

99. Cormier, H., et al., Association between yogurt consumption, dietary patterns, and cardio-metabolic risk factors. European journal of nutrition, 2016. 55: p. 577-587.
100. Possa, G., et al., 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. Nutr Res, 2015. 35(8): p. 700-6.
101. Feeney, E.L., et al., Patterns of dairy food intake, body composition and markers of metabolic health in Ireland: results from the National Adult Nutrition Survey. Nutr Diabetes, 2017. 7(2): p. e243.
102. Sanchez-Villegas, A., et al., A systematic review of socioeconomic differences in food habits in Europe: consumption of cheese and milk. Eur J Clin Nutr, 2003. 57(8): p. 917-29.
103. Slurink, I. and F. Imamura, Dairy prediabetes Fenland.
104. Slurink, I.A., et al., Dairy Product Consumption and Incident Prediabetes in the Australian Diabetes, Obesity, and Lifestyle Study With 12 Years of Follow-Up. J Nutr, 2023. 153(6): p. 1742-1752.
105. Slurink, I.A., et al., Dairy consumption and incident prediabetes: prospective associations and network models in the large population-based Lifelines study. Am J Clin Nutr, 2023. 118(6): p. 1077-1090.
106. Slurink, I.A.L., et al., Dairy Product Consumption in Relation to Incident Prediabetes and Longitudinal Insulin Resistance in the Rotterdam Study. Nutrients, 2022. 14(3).
107. Geng, T., L. Qi, and T. Huang, Effects of dairy products consumption on body weight and body composition among adults: an updated meta-analysis of 37 randomized control trials. Molecular nutrition & food research, 2018. 62(1): p. 1700410.
108. Chan, J.M., et al., Obesity, fat distribution, and weight gain as risk factors for clinical diabetes in men. Diabetes care, 1994. 17(9): p. 961-969.
109. Goossens, G.H., The role of adipose tissue dysfunction in the pathogenesis of obesity-related insulin resistance. Physiol Behav, 2008. 94(2): p. 206-18.
110. DeFronzo, R.A., et al., The effect of insulin on the disposal of intravenous glucose. Results from indirect calorimetry and hepatic and femoral venous catheterization. Diabetes, 1981. 30(12): p. 1000-7.
111. Merz, K.E. and D.C. Thurmond, Role of skeletal muscle in insulin resistance and glucose uptake. Comprehensive Physiology, 2011. 10(3): p. 785-809.
112. Gezondheidsraad. Zuivel - Achtergronddocument bij Richtlijnen goede voeding 2015. Den Haag: Gezondheidsraad, 2015; publicatienr. A15/32.
113. Benatar, J.R., K. Sidhu, and R.A. Stewart, Effects of high and low fat dairy food on cardio-metabolic risk factors: a meta-analysis of randomized studies. PLoS One, 2013. 8(10): p. e76480.
114. Chen, M., et al., Effects of dairy intake on body weight and fat: a meta-analysis of randomized controlled trials. Am J Clin Nutr, 2012. 96(4): p. 735-47.
115. Schwingshackl, L., et al., Consumption of Dairy Products in Relation to Changes in Anthropometric Variables in Adult Populations: A Systematic Review and Meta-Analysis of Cohort Studies. PLoS One, 2016. 11(6): p. e0157461.
116. Mendelian Randomization of Dairy Consumption Working, G. and C. consortium, Dairy Intake and Body Composition and Cardiometabolic Traits among Adults: Mendelian Randomization Analysis of 182041 Individuals from 18 Studies. Clin Chem, 2019. 65(6): p. 751-760.
117. Gilbert, J.A., et al., Milk supplementation facilitates appetite control in obese women during weight loss: a randomised, single-blind, placebo-controlled trial. Br J Nutr, 2011. 105(1): p. 133-43.
118. Zhang, S., et al., Novel metabolic and physiological functions of branched chain amino acids: a review. J Anim Sci Biotechnol, 2017. 8(1): p. 10.
119. Jakubowicz, D. and O. Froy, Biochemical and metabolic mechanisms by which dietary whey protein may combat obesity and Type 2 diabetes. J Nutr Biochem, 2013. 24(1): p. 1-5.
120. Akhavan, T., et al., Mechanism of action of pre-meal consumption of whey protein on glycemic control in young adults. J Nutr Biochem, 2014. 25(1): p. 36-43.
121. Gunnerud, U.J., E.M. Ostman, and I.M. Bjorck, Effects of whey proteins on glycaemia and insulinaemia to an oral glucose load in healthy adults; a dose-response study. Eur J Clin Nutr, 2013. 67(7): p. 749-53.
122. Onvani, S., et al., Dairy products, satiety and food intake: A meta-analysis of clinical trials. Clin Nutr, 2017. 36(2): p. 389-398.

123. Vogtschmidt, Y.D., et al., Is protein the forgotten ingredient: Effects of higher compared to lower protein diets on cardiometabolic risk factors. A systematic review and meta-analysis of randomised controlled trials. Atherosclerosis, 2021. 328: p. 124-135.
124. Chen, L., et al., Dairy Milk Casein and Whey Proteins Differentially Alter the Postprandial Lipidome in Persons with Prediabetes: A Comparative Lipidomics Study. J Agric Food Chem, 2022. 70(33): p. 10209-10220.
125. Josse, A.R., et al., 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. J Nutr, 2011. 141(9): p. 1626-34.
126. Li, H., et al., Leucine supplementation increases SIRT1 expression and prevents mitochondrial dysfunction and metabolic disorders in high-fat diet-induced obese mice. Am J Physiol Endocrinol Metab, 2012. 303(10): p. E1234-44.
127. Ren, M., et al., Different Lipopolysaccharide Branched-Chain Amino Acids Modulate Porcine Intestinal Endogenous beta-Defensin Expression through the Sirt1/ERK/90RSK Pathway. J Agric Food Chem, 2016. 64(17): p. 3371-9.
128. Chen, Q. and R.A. Reimer, Dairy protein and leucine alter GLP-1 release and mRNA of genes involved in intestinal lipid metabolism in vitro. Nutrition, 2009. 25(3): p. 340-9.
129. Foltz, M., P.C. van der Pijl, and G.S. Duchateau, Current in vitro testing of bioactive peptides is not valuable. J Nutr, 2010. 140(1): p. 117-8.
130. Eelderink, C., et al., The effect of high compared with low dairy consumption on glucose metabolism, insulin sensitivity, and metabolic flexibility in overweight adults: a randomized crossover trial. Am J Clin Nutr, 2019. 109(6): p. 1555-1568.
131. Schmidt, K.A., et al., The impact of diets rich in low-fat or full-fat dairy on glucose tolerance and its determinants: a randomized controlled trial. Am J Clin Nutr, 2021. 113(3): p. 534-547.
132. Chen, G., et al., Biomarkers of fatty acids and risk of type 2 diabetes: a systematic review and meta-analysis of prospective cohort studies. Crit Rev Food Sci Nutr, 2021. 61(16): p. 2705-2718.
133. Huang, L., et al., Circulating Saturated Fatty Acids and Incident Type 2 Diabetes: A Systematic Review and Meta-Analysis. Nutrients, 2019. 11(5): p. 998.
134. Hodson, L., C.M. Skeaff, and B.A. Fielding, 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): p. 348-380.
135. Imamura, F., et al., Fatty acids in the de novo lipogenesis pathway and incidence of type 2 diabetes: A pooled analysis of prospective cohort studies. PLoS Med, 2020. 17(6): p. e1003102.
136. Imamura, F., 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. 15(10): p. e1002670.
137. Jenkins, B., J.A. West, and A. Koulman, 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): p. 2425-44.
138. Jaudszus, A., et al., trans Palmitoleic acid arises endogenously from dietary vaccenic acid. Am J Clin Nutr, 2014. 99(3): p. 431-5.
139. Wolk, A., M. Furuheim, and B. Vessby, Fatty acid composition of adipose tissue and serum lipids are valid biological markers of dairy fat intake in men. J Nutr, 2001. 131(3): p. 828-33.
140. Brevik, A., et al., 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): p. 1417-22.
141. Wolk, A., et al., Evaluation of a biological marker of dairy fat intake. Am J Clin Nutr, 1998. 68(2): p. 291-5.
142. Ikwuobe, O.J., et al., Odd-Chain Fatty Acids Predict Insulin Sensitivity in People with T2DM and Protect HepG2 Cells from Palmitate-Induced Insulin Resistance Via PPARα. Free Radical Biology and Medicine, 2016. 100: p. S174.
143. Tremblay, B.L. and I. Rudkowska, Nutrigenomic point of view on effects and mechanisms of action of ruminant trans fatty acids on insulin resistance and type 2 diabetes. Nutr Rev, 2017. 75(3): p. 214-223.
144. Kraft, J., et al., Dairy-derived bioactive fatty acids improve pancreatic ß-cell function. The FASEB Journal, 2015. 29: p. 608.25.
145. Wein, S., et al., Medium-chain fatty acids ameliorate insulin resistance caused by high-fat diets in rats. Diabetes Metab Res Rev, 2009. 25(2): p. 185-94.

146. Lundsgaard, A.M., et al., Small Amounts of Dietary Medium-Chain Fatty Acids Protect Against Insulin Resistance During Caloric Excess in Humans. Diabetes, 2021. 70(1): p. 91-98.
147. Montgomery, M.K., et al., Contrasting metabolic effects of medium- versus long-chain fatty acids in skeletal muscle. J Lipid Res, 2013. 54(12): p. 3322-33.
148. Hommelberg, P.P., et al., Fatty acid-induced NF-kappaB activation and insulin resistance in skeletal muscle are chain length dependent. Am J Physiol Endocrinol Metab, 2009. 296(1): p. E114-20.
149. Kim, D.H., et al., Kefir alleviates obesity and hepatic steatosis in high-fat diet-fed mice by modulation of gut microbiota and mycobiota: targeted and untargeted community analysis with correlation of biomarkers. J Nutr Biochem, 2017. 44: p. 35-43.
150. Naito, E., et al., Beneficial effect of oral administration of Lactobacillus casei strain Shirota on insulin resistance in diet-induced obesity mice. J Appl Microbiol, 2011. 110(3): p. 650-7.
151. Tabuchi, M., et al., Antidiabetic effect of Lactobacillus GG in streptozotocin-induced diabetic rats. Biosci Biotechnol Biochem, 2003. 67(6): p. 1421-4.
152. Bron, P.A., et al., Can probiotics modulate human disease by impacting intestinal barrier function? British Journal of Nutrition, 2017. 117(1): p. 93-107.
153. Caricilli, A.M. and M.J. Saad, The role of gut microbiota on insulin resistance. Nutrients, 2013. 5(3): p. 829-51.
154. Nikbakht, E., et al., Effect of probiotics and synbiotics on blood glucose: a systematic review and meta-analysis of controlled trials. Eur J Nutr, 2018. 57(1): p. 95-106.
155. Schurgers, L.J. and C. Vermeer, Determination of phylloquinone and menaquinones in food. Effect of food matrix on circulating vitamin K concentrations. Haemostasis, 2000. 30(6): p. 298-307.
156. Vermeer, C., et al., Menaquinone Content of Cheese. Nutrients, 2018. 10(4): p. 446.
157. Beulens, J.W., et al., Dietary phylloquinone and menaquinones intakes and risk of type 2 diabetes. Diabetes Care, 2010. 33(8): p. 1699-705.
158. Hussein, A.G., et al., Vitamin K(2) alleviates type 2 diabetes in rats by induction of osteocalcin gene expression. Nutrition, 2018. 47: p. 33-38.
159. Iwamoto, J., et al., Vitamin K(2) prevents hyperglycemia and cancellous osteopenia in rats with streptozotocin-induced type 1 diabetes. Calcif Tissue Int, 2011. 88(2): p. 162-8.
160. Rahimi Sakak, F., et al., Glycemic control improvement in individuals with type 2 diabetes with vitamin K(2) supplementation: a randomized controlled trial. Eur J Nutr, 2021. 60(5): p. 2495-2506.
161. Kanazawa, I., Osteocalcin as a hormone regulating glucose metabolism. World journal of diabetes, 2015. 6(18): p. 1345.
162. Li, Y., et al., Effect of vitamin K2 on type 2 diabetes mellitus: A review. Diabetes Res Clin Pract, 2018. 136: p. 39-51.
163. Feeney, E.L., et al., Dairy matrix effects: response to consumption of dairy fat differs when eaten within the cheese matrix-a randomized controlled trial. Am J Clin Nutr, 2018. 108(4): p. 667-674.
164. de Goede, J., et al., Effect of cheese consumption on blood lipids: a systematic review and meta-analysis of randomized controlled trials. Nutr Rev, 2015. 73(5): p. 259-75.
165. Pradeilles, R., et al., Effect of Isoenergetic Substitution of Cheese with Other Dairy Products on Blood Lipid Markers in the Fasted and Postprandial State: An Updated and Extended Systematic Review and Meta-Analysis of Randomized Controlled Trials in Adults. Adv Nutr, 2023. 14(6): p. 1579-1595.
166. Rosqvist, F., et al., Potential role of milk fat globule membrane in modulating plasma lipoproteins, gene expression, and cholesterol metabolism in humans: a randomized study. Am J Clin Nutr, 2015. 102(1): p. 20-30.
167. Christensen, R., et al., Effect of calcium from dairy and dietary supplements on faecal fat excretion: a meta-analysis of randomized controlled trials. Obes Rev, 2009. 10(4): p. 475-86.
168. Jacobsen, R., et al., Effect of short-term high dietary calcium intake on 24-h energy expenditure, fat oxidation, and fecal fat excretion. Int J Obes (Lond), 2005. 29(3): p. 292-301.
169. Booth, A.O., et al., 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): p. 1013-25.
170. Onakpoya, I.J., et al., Efficacy of calcium supplementation for management of overweight and obesity: systematic review of randomized clinical trials. Nutr Rev, 2011. 69(6): p. 335-43.

171. Booth, A.O., et al., Effect of increasing dietary calcium through supplements and dairy food on body weight and body composition: a meta-analysis of randomised controlled trials. British Journal of Nutrition, 2015. 114(7): p. 1013-1025.
172. Chung, R.W., et al., Dietary sphingomyelin lowers hepatic lipid levels and inhibits intestinal cholesterol absorption in high-fat-fed mice. PLoS One, 2013. 8(2): p. e55949.
173. Norris, G.H., et al., Dietary sphingomyelin attenuates hepatic steatosis and adipose tissue inflammation in high-fat-diet-induced obese mice. J Nutr Biochem, 2017. 40: p. 36-43.
174. Suarez, F.L., et al., Tolerance to the daily ingestion of two cups of milk by individuals claiming lactose intolerance. Am J Clin Nutr, 1997. 65(5): p. 1502-6.
175. Gerbault, P., et al., Evolution of lactase persistence: an example of human niche construction. Philos Trans R Soc Lond B Biol Sci, 2011. 366(1566): p. 863-77.
176. Luo, K., et al., Variant of the lactase LCT gene explains association between milk intake and incident type 2 diabetes. Nature Metabolism, 2024: p. 1-18.
177. JanssenDuijghuijsen, L., et al., Changes in gut microbiota and lactose intolerance symptoms before and after daily lactose supplementation in individuals with the lactase nonpersistent genotype. Am J Clin Nutr, 2024. 119(3): p. 702-710.
178. Mann, J.I., Evidence-based nutrition: does it differ from evidence-based medicine? Annals of medicine, 2010. 42(7): p. 475-486.
179. Higgins, J.P. and S. Green, Cochrane handbook for systematic reviews of interventions. The Cochrane Collaboration. London, UK, 2011.
180. Rothman, K.J. and S. Greenland, Causation and causal inference in epidemiology. American journal of public health, 2005. 95(S1): p. S144-S150.
181. Lachat, C., et al., Strengthening the Reporting of Observational Studies in Epidemiology-Nutritional Epidemiology (STROBE-nut): An Extension of the STROBE Statement. PLoS Med, 2016. 13(6): p. e1002036.
182. InterConnect, Gobal data for diabetes and obesity research. Available online: http://www.interconnect-diabetes.eu/ (accessed on 17 Januari 2024).
183. Iqbal, K., et al., Gaussian graphical models identified food intake networks and risk of type 2 diabetes, CVD, and cancer in the EPIC-Potsdam study. Eur J Nutr, 2019. 58(4): p. 1673-1686.
184. Jahanmiri, R., et al., Saturated fats network identified using Gaussian graphical models is associated with metabolic syndrome in a sample of Iranian adults. Diabetology & Metabolic Syndrome, 2022. 14(1): p. 1-11.
185. Hoang, T., J. Lee, and J. Kim, Differences in Dietary Patterns Identified by the Gaussian Graphical Model in Korean Adults With and Without a Self-Reported Cancer Diagnosis. J Acad Nutr Diet, 2021. 121(8): p. 1484-1496 e3.
186. Schwedhelm, C., et al., Meal and habitual dietary networks identified through semiparametric Gaussian copula graphical models in a German adult population. PLoS One, 2018. 13(8): p. e0202936.
187. Hoang, T., J. Lee, and J. Kim, Network Analysis of Demographics, Dietary Intake, and Comorbidity Interactions. Nutrients, 2021. 13(10): p. 3563.
188. Epskamp, S. and E.I. Fried, A tutorial on regularized partial correlation networks. Psychol Methods, 2018. 23(4): p. 617-634.
189. Page, M.J., et al., The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. International journal of surgery, 2021. 88: p. 105906.
190. Wells, G.A., et al., The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. 2000, Oxford.

Chapter 2

Dairy intake and prediabetes risk in the Hoorn Studies

Manuscript based on this chapter

Isabel A.L. Slurink, Nicole R. den Braver, Femke Rutters, Nina Kupper, Tom Smeets, Petra J. Elders, Joline W. Beulens & Sabita S. Soedamah-Muthu (2021). *Dairy product consumption and incident prediabetes in Dutch middle-aged adults: the Hoorn Studies prospective cohort.* European Journal of Nutrition, 1-14.

Abstract

Objective

Our aim was to investigate prospective associations of consumption of total dairy and dairy types with risk of prediabetes in a Dutch population-based study.

Methods

Two enrolment waves of the Hoorn Studies were harmonized, resulting in an analytic sample of 2,262 participants without (pre-)diabetes at enrolment (mean age 56 ± 7.3 years; 50% male). Baseline dietary intake was assessed by validated food frequency questionnaires. Relative risks (RRs) were calculated between dairy, fermented dairy, milk, yogurt (all total/ high/low-fat), cream and ice cream and prediabetes. Additionally, substituting one serving/day of dairy types associated with prediabetes with alternative dairy types was analysed.

Results

During a mean 6.4 ± 0.7 years of follow-up, 810 participants (35.9%) developed prediabetes. High-fat fermented dairy, cheese and high-fat cheese were associated with a 17% (RR 0.83, 95%CI 0.69-0.99, P_{trend} = 0.04), 14% (RR 0.86, 95%CI 0.73-1.02, P_{trend} = 0.04) and 21% (RR 0.79, 95%CI 0.66-0.94, P_{trend} = 0.01) lower risk of prediabetes, respectively, in top compared to bottom quartiles, after adjustment for confounders. High-fat cheese consumption was continuously associated with lower prediabetes risk ($RR_{serving/day}$ 0.94, 95%CI 0.88-1.00). Total dairy and other dairy types were not associated with prediabetes risk in adjusted models, irrespective of fat content (RR ~ 1). Replacing high-fat cheese with alternative dairy types was not associated with prediabetes risk.

Conclusion

The highest intake of high-fat fermented dairy, cheese and high-fat cheese were associated with a lower risk of prediabetes, whereas other dairy types were not associated. Cheese seems to be inversely associated with type 2 diabetes risk, despite high levels of saturated fatty acids and sodium.

Introduction

Prediabetes is a condition characterized by blood glucose levels that are above the normal range, but still fall below the diagnostic threshold for type 2 diabetes (T2D) [1]. The prevalence of prediabetes is rapidly rising worldwide from 374 million in 2019 to an expected 548 million in 2045 [2]. People with prediabetes are at increased risk of developing T2D and cardiovascular diseases (CVD) [3, 4], but may reverse to normoglycaemia with lifestyle adaptation [5]. This emphasizes the need to identify modifiable risk factors that could prevent or reverse this condition. Suboptimal diet is causally linked to incidence of prediabetes and T2D, and majority of cases can be prevented by dietary modification [6, 7].

Dairy products are widely consumed and may provide considerable quantities of beneficial nutrients for metabolic health, including protein, minerals (calcium, magnesium, potassium) and vitamins (A, D, B2, B12), but also contain saturated fatty acids (SFAs) and sodium. A recent summary of meta-analyses reported dose-response relations of low-fat dairy (RRs ranging from 0.88-0.98), yogurt (0.78-0.94) and cheese (0.80-1.00) with T2D, inconsistent results for milk (RRs 0.83-1.27), with considerable heterogeneity present between studies [8]. One possible explanation for heterogeneity, proposed by Hruby *et al.*, could be differences in participant's baseline glycaemic status [9], and the precise moment along the physiological progress of T2D at which specific dairy products modify risk is largely unknown. Therefore, studies aiming to elucidate associations between dairy and early-risk stages are warranted.

Only one prospective cohort study investigated the associations of dairy products with incident prediabetes, based on fasting plasma glucose (FPG)[9]. In the US FHS Offspring Cohort (*n* = 1867, 12-year follow-up), highest consumption of total, low-fat and high-fat dairy was associated with 39%, 32% and 25% lower prediabetes incidence, respectively, compared to lowest consumption, with nonlinear protective associations for milk and yogurt. In the French DESIR study (9-year follow-up), prediabetes and T2D were combined as one outcome, inhibiting interpretation of associations with prediabetes alone. The DESIR study observed an association between higher total dairy (except cheese) intake with lower hyperglycaemia incidence [10]. Furthermore, two studies investigating continuous outcomes of glucose metabolism showed no associations of any dairy types [11, 12], except for an association of higher fermented dairy with lower FPG and HbA1c in the Danish Inter99 study [11]. Evidence from cross-sectional studies indicated that mainly higher fermented and skimmed dairy intake were associated with lower prediabetes risk [13-15], with one study also reporting associations for higher non-fermented and high-fat dairy intake and higher prediabetes risk [14]. Thus, although there are some indications for beneficial associations of dairy consumption on prediabetes risk, associations are highly heterogeneous, partly underlined by different definitions of prediabetes outcomes and large variations in dairy consumption habits, advocating the need for country and region-specific prospective data.

Therefore, this study aimed to investigate prospective associations between consumption of total dairy and dairy types with incident prediabetes.

Methods

Study design and population

This study used data of the Hoorn Studies, a prospective population-based cohort study with the first enrolment wave in 1989-1992 (Hoorn Study 1, HS1) and a second wave in 2006-2007 (HS2). The aim was to study prevalence and risk factors for disturbances in glucose metabolism and T2D. Both enrolment waves were similar in design, population characteristics and questionnaires [16], and could, therefore, be harmonized to increase sample size and study power. Furthermore, this harmonization resulted in increased variation in dairy product intake and inclusion of more up-to-date information. People from the general population were recruited, aged 50-75 years in the HS1 and 40-65 years in the HS2 at time of inclusion. Follow-up measurements were performed between the years 1996-1998 in the HS1 and 2013-2015 in the HS2. Visits took place at the Diabetes Care Center in the city of Hoorn, the Netherlands. Written informed consent was obtained from all participants. The study was approved by the Ethics Committee of the Amsterdam University Medical Centers, location VUMC.

From participants with follow-up data available (*n* = 3,245), we excluded participants with prediabetes (*n* = 557) or T2D at baseline (*n* = 229) or follow-up (*n* = 528) (**Figure 1**). Exclusion of prediabetes was based on FPG between 6.1-7.0 mmol/L, 2 h plasma glucose (2hPG) between 7.8 and 11.0 mmol/L and/or HbA1c levels between 6.0 and 6.5% [1, 17]. Exclusion of T2D was based on diagnosis by a general practitioner, diabetes medication user and/or an FPG ≥ 7.0 mmol/L, 2hPG ≥ 11.1 mmol/L, or HbA1c ≥ 6.5%. Other exclusion criteria were self-reported history of T2D prior to baseline (*n* = 59), extreme energy intake (top and bottom 0.5%) (*n* = 34) or missing information on dairy intake (*n* = 42), and/or missing data on prediabetes at follow-up (*n* = 9). After exclusion, the analytic sample consisted of 2,262 participants.

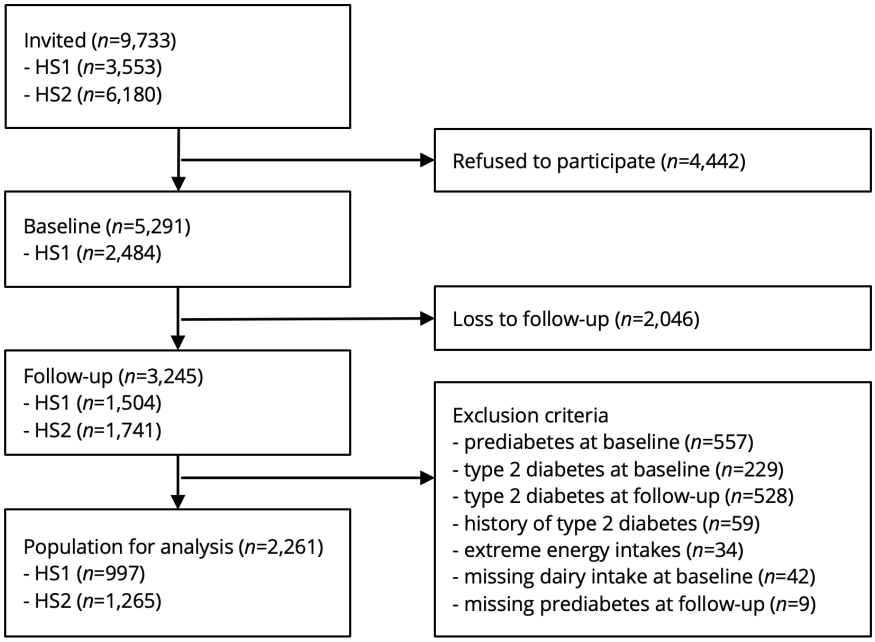


Figure 1. Flow-chart for inclusion of participants for the present analysis of the Hoorn Studies (HS).

Ascertainment of prediabetes

At all study visits, bloods samples were drawn to determine FPG, 2hPG levels after a 75-g oral glucose tolerance test (OGTT) and HbA1c levels, except at the HS2 follow-up visit, where no OGTT was conducted and HbA1c was measured in fasting capillary blood samples obtained using a blood spot card. FPG and 2hPG levels were determined using the glucose dehydrogenase method (Merck, Darmstadt, Germany). In the HS1, HbA1c was determined by ion-exchange high-performance liquid chromatography with a Modular Diabetes Monitoring System (Bio-Rad, Veenendaal, The Netherlands). In the HS2 at baseline, HbA1c levels were assessed using standardized reverse-phase cation-exchange chromatography (HA 8160 analyzer; Menarini, Florence, Italy). In the HS2 at follow-up, HbA1c levels were derived from blood spot cards, using thermo immunoturbidimetry according to a validated protocol [18]. Prediabetes at follow-up was defined according to the diagnostic criteria of the World Health Organization of 2006 [1], complemented with the HbA1c cut-offs values proposed by the 2009 international expert committee for prediabetes [17], with FPG between 110 and 125 mg/dL (6.1 and 6.9 mmol/L), 2hPG between 140 and 199 mg/dL (7.8 and 11.0 mmol/L) and/or HbA1c levels between 42 and 46 mmol/mol (6.0 and 6.4%).

Dietary assessment

Baseline dietary intake was assessed with a 92-item food frequency questionnaire (FFQ) in the HS1 and a 104-item FFQ in the HS2. The HS1 FFQ was validated against a dietary history in 74 males and females and was valid for ranking individuals according to energy intake (r = 0.72), and main nutrients in dairy products; animal protein (r = 0.68), SFAs (r = 0.73) and calcium (r = 0.75) [19]. The HS2 FFQ was validated against actual energy intake in controlled feeding trials for energy intake (r = 0.82) [20] and validated against three 24-h recalls for animal protein (r = 0.49), SFA (r = 0.44) and calcium (r = 0.56) [21].

Participants were asked to report their usual frequency of consumption, serving size and preparation in the past year. Seasonal variations in milk consumption were assessed with separate questions for winter and summer intakes. Participants completed the questionnaire at home and checked for completeness by a trained dietician. Intake (gram/day) per FFQ item was calculated using the Dutch food composition table (NEVO) 1989/1990 for HS1 and the NEVO 2006 for HS2. FFQ items were combined and categorized as total dairy, fermented dairy, and by subtypes milk (all types and plain milk), yogurt, cheese, cream, and ice cream (**Table 1**). Each dairy category was further divided into low-fat (liquid products, ≤ 2%; cheese ≤ 20%) and high-fat (liquid products, > 2%; cheese > 20%). Intakes were converted to servings per day according to Dutch standard serving sizes: milk, 200 mL; yogurt, 150 mL; cheese, 20 g; cream, 3 g; ice cream, 100 g (https://portie-online.rivm.nl/). In the total dairy category, a serving of liquid dairy products was defined as 200 mL and a serving of cheese as 20 g. Because two different FFQs were used, dietary intakes of food groups and dairy types stratified by enrolment wave are reported in **Supplementary Table 1**.

Table 1. Food items and their fat content included in total dairy and dairy types and consumption in the Hoorn Studies (*n* = 2,262).

Dairy product	Included dairy types	Consumers[1]	Intake (servings/d)[2]	
		%	Mean ± SD	Median [IQR]
Total dairy	All dairy products	99.8%	3.0 ± 1.7	2.7 [1.8; 3.9]
High-fat	All high-fat dairy products	98.1%	1.5 ± 1.3	1.3 [0.5; 2.2]
Low-fat	All low-fat dairy products	92.7%	1.5 ± 1.2	1.3 [0.6; 2.1]
Total fermented dairy	Full fat yogurt, full fat fruit yogurt, full fat curd, high-fat cheese, full fat luxury cheese	98.5%	2.2 ± 1.4	1.9 [1.1; 2.8]
High-fat[3]	Full fat yogurt, full fat fruit yogurt, full fat curd, high-fat cheese, full fat luxury cheese	87.7%	1.3 ± 1.2	1.0 [0.4; 1.9]
Low-fat[4]	Semi-skimmed yogurt, skimmed yogurt, skimmed fruit yogurt, semi-skimmed curd, skimmed curd, semi-skimmed fruit curd, skimmed fruit curd, low-fat cheese, low-fat luxury cheese	83.8%	0.9 ± 0.9	0.7 [0.2; 1.3]

Table 1. Food items and their fat content included in total dairy and dairy types and consumption in the Hoorn Studies (*n* = 2,262). (continued)

Dairy product	Included dairy types	Consumers[1]	Intake (servings/d)[2]	
		%	Mean ± SD	Median [IQR]
Total milk	All milks	91.6%	1.1 ± 1.0	0.8 [0.3; 1.5]
High-fat	Full fat milk, full fat chocolate milk, milk powder, full fat milk added to the coffee, drinking yogurt[5], fruit flavoured milk[5]	40.8%	0.2 ± 0.5	0.0 [0.0; 0.1]
Low-fat	Semi-skimmed milk, skimmed milk, buttermilk, semi-skimmed chocolate milk, skimmed chocolate milk, semi-skimmed milk added to the coffee, skimmed milk added to the coffee, semi-skimmed fruit milk	82.8%	0.9 ± 0.9	0.7 [0.1; 1.4]
Plain milk	All plain milks	81.1%	0.9 ± 1.0	0.7 [0.1; 1.4]
High-fat	Full fat milk	12.1%	0.1 ± 0.4	0.0 [0.0; 0.0]
Low-fat	Semi-skimmed milk, skimmed milk, buttermilk	76.3%	0.8 ± 0.9	0.5 [0.0; 1.4]
Yogurt	Yogurt	83.3%	0.5 ± 0.5	0.5 [0.1; 0.9]
High-fat[3]	Full fat yogurt, full fat fruit yogurt	30.4%	0.1 ± 0.3	0.0 [0.0; 0.1]
Low-fat[6]	Semi-skimmed yogurt, skimmed yogurt, skimmed fruit yogurt	68.0%	0.4 ± 0.5	0.3 [0.0; 0.6]
Cheese	All cheeses	95.2%	1.4 ± 1.1	1.1 [0.6; 1.9]
High-fat	HS1: regular cheese, cheese cubes; HS2: 40 + cheese (e.g. Edam), 48 + cheese (e.g. Gouda, cheddar, cheese spread, goat cheese), full fat luxury cheese (e.g. cream brie, cream cheese, mon chou), cheese cubes, grated cheese, feta, cheese fondue	85.6%	1.1 ± 1.1	0.9 [0.3; 1.8]
Low-fat	HS1: skimmed cheese; HS2: 20 + and 30 + cheese (e.g. cheese spread, cottage cheese), low-fat luxury cheese (e.g. brie, goat cheese)	31.9%	0.2 ± 0.5	0.0 [0.0; 0.3]
Cream	Whipped cream, coffee cream, semi-skimmed coffee cream, sour cream, crème fraiche, cooking cream	58.5%	0.8 ± 2.5	0.3 [0.0; 0.5]
Ice cream	Ice cream	68.3%	0.1 ± 0.1	0.0 [0.0; 0.1]

[1] Consumers were defined as consuming >0 servings/day of a specific dairy type.
[2] Serving sizes were milk, 200 mL; yogurt, 150 mL; cheese, 20 g; cream, 3 g and ice cream, 100 g.
[3] Includes oatmeal porridge, rice porridge and full fat custard in HS2.
[4] Includes buttermilk in HS1, buttermilk porridge and skimmed custard in HS2.
[5] Only assessed in HS2.
[6] Includes semi-skimmed curd, skimmed curd, semi-skimmed fruit curd, skimmed fruit curd, buttermilk porridge and skimmed custard in HS2.
Abbreviations: HS1 Hoorn Study 1 (first enrolment wave), HS2 Hoorn Study 2 (second enrolment wave).

Covariates

The self-administered baseline questionnaire for both enrolment waves included questions on socio-demographic, life- style and clinical factors. Responses were verified in a per- sonal interview. Smoking status was categorized as current, former, or never. Highest educational level was obtained in eight levels, which were subsequently categorized to low (no education or primary school), middle (secondary education) and high (tertiary education). Moderate physical activity in hours/week was assessed using the SQUASH questionnaire, for which the Spearman correlation for overall reproducibility was 0.58 in 50 participants compared to an activity monitor [22]. The activities included sports, bicycling, gardening, walking, doing chores and housekeeping. Alcohol intake was categorized as non-drinker, ≤ 10, 10-30 and ≥ 30 g/day. Family history of diabetes was defined as having at least a grandparent, parent, sibling, or child with diabetes.

Physical measurements were performed at baseline. BMI was calculated as weight divided by height squared (kg/m2), and categorized as < 25 kg/m2, 25-30 kg/m2 and ≥ 30 kg/ m2. Blood pressure was measured on the right arm with a random-zero sphygmomanometer (Hawksley-Gelman Ltd, Lancing, United Kingdom) while participants were sitting. Plasma levels of total cholesterol, triglycerides, high-density lipoprotein (HDL) were measured in fasting blood samples by enzymatic techniques (Boehringer-Mannheim, Mannheim, Germany) and low-density lipoprotein (LDL) was calculated using Friedewald's formula (except for participants with triglycerides > 4.55 mmol/L)[23].

Statistical analysis

Statistical analyses were performed using R version 3.6.2 (R Foundation for Statistical Computing, Vienna, Austria). Baseline characteristics are displayed as means ± SD, medians (IQR) or percentages for the total study population and in quartiles of total dairy intake. Missing values in con- founding variables (for 6% of participants, highest 2% for physical activity) were imputed using multiple imputation (n = 10) (**Supplemental table 2**).

Poisson regression with robust variance was used to examine associations between dairy product intakes and prediabetes, because of the high incidence of prediabetes (35.9%), in which case the odds ratio overestimates the strength of the association [24]. RRs with 95%CIs were calculated for quartiles of dairy intake (reference lowest) and on a continuous scale (serving/day). Dairy products for which many participants reported no intake were divided in a non-consumer category (reference) and consumers in tertiles. Linear trend across intake range categories were assessed by including median values of each category as a continuous variable in the model. Linearity was assumed in all models, as no indications for non-linearity presented assessed by adding a quadratic term to model 3. Regression coefficients for each of the imputed datasets were pooled.

Confounder models were constructed based on literature [25] and on distributions of baseline characteristics across quartiles. Model 1 included age, sex, follow-up duration and enrolment wave. Model 2 additionally adjusted for energy intake, education, smoking, physical activity, alcohol consumption and family history of diabetes. Model 3 additionally adjusted for food groups associated with T2D including intakes of fruit, vegetables, tea, coffee, grains (whole and refined), meat (processed and red) and sugar-sweetened beverages [26, 27]. BMI, blood pressure (systolic and diastolic) and LDL cholesterol were added separately in model 4 because of their potential mediating or otherwise confounding effect. We checked for effect modification by enrolment wave by including an interaction term in model 3 and stratified associations by enrolment wave to assess if associations differed for each wave of the Hoorn Studies. Furthermore, effect modification by age, sex and BMI was examined, and associations were stratified in case of significance.

We provided a supplementary baseline table stratified by the dairy types that were significantly associated with prediabetes in the main analyses, to examine confounding of associations by healthy lifestyle. Furthermore, we examined substituting one serving/day of significant dairy types with alternative dairy types in model 3. Models included a total dairy intake variable (servings/day), all individual dairy types (servings/day) except for the dairy type to be substituted, and energy intake. The estimated RR for each alternative dairy type can be interpreted as the RR for substitution of a daily serving of the alternative dairy type for a daily serving of the excluded dairy type [28].

A series of sensitivity analyses were conducted using model 3. First, the independence of the associations of specific dairy types was evaluated by mutually adjusting for intake of other dairy types. Second, we repeated the analysis excluding participants with self-reported CVD (*n* = 261). Third, we repeated the analysis in 'normal energy reporters' only, identified using the Goldberg method [29]. For this method, the basal metabolic rate (BMR) was calculated for each participant using Schofield equations specifically for age and sex categories based on weight [30]. Following, the ratio of energy intake (EI) and BMR was calculated. Using the Goldberg cut-offs described by Black *et al.*, participants with EI:BMR < 1.08 were classified as under-reporters, participants with 1.08 ≤ EI:BMR ≤ 2.22 were classified as 'normal reporters' and those with EI:BMR > 2.22 were defined as 'over-reporters' [29]. Lastly, to address possible misclassification in prediabetes defined at baseline, we repeated analyses including participants with prediabetes at baseline (*n* = 557, final sample *n* = 2,661).

Results

Participant characteristics

The mean age of the study population was 55.9 ± 7.3 years, 50% were male and 22% were current smokers (**Table 2**, by enrolment wave **Supplemental Table 1**). The mean BMI was 25.7 ± 3.4 kg/m2 and 10% of participants were obese (BMI ≥ 30 kg/m2). The

mean dairy intake was 3.0 ± 1.7 servings/day (357 ± 237 g/day) (**Table 1**). Participants in the top (3.9-15.4 servings/day) compared to the bottom quartile (0-1.8 servings/day) of dairy intake were more often male (54% vs 43%), with low education (16% vs 8%), more physically active (median (IQR): 9.0 (5.1-13.7) vs 6.5 (3.5-10.5) hours/week) and had higher LDL cholesterol levels (4.2 ± 1.1 vs 3.5 ± 1.0 mmol/L). With increasing dairy intake over the quartiles, energy, calcium, fruit, and processed meat intakes were higher and vegetable and alcohol intakes were lower. Participants with complete follow-up data (*n* = 3,245) were similar to participants lost-to-follow-up (*n* = 2046) with regard to age (56.9 ± 7.6 vs 58.0 ± 9.0 years), sex (51% vs 49% male), physical activity (7.5 IQR 4.0-12.3 vs 7.0 IQR 3.5-12.7 h/week), BMI (26.3 ± 3.6 vs 26.6 ± 4.1 kg/m2) and fasting glucose (5.6 ± 1.0 versus 5.8 ± 1.6 mmol/L) (**Supplemental Table 3**). Participants lost-to-follow-up were slightly more often lower educated (22% vs 15%), current smoker (31% vs 23%) and had a lower alcohol intake (median 5.0 IQR 0.0-15.3 vs 7.2 IQR 2.0-17.2).

Dairy intake and prediabetes risk

During a mean follow-up duration of 6.4 ± 0.7 years, 811 out of 2,262 participants developed prediabetes (35.9%). High-fat fermented dairy intake was significantly associated with lower prediabetes risk in model 3 (RR_{Q4vsQ1} 0.83, 95%CI 0.69-0.99, P_{trend} = 0.04) (**Table 3**). High-fat fermented dairy intake mainly consisted of cheese intake (63%), which was marginally significantly associated with lower risk of prediabetes in the top versus bottom quartile (RR_{Q4vsQ1} 0.86, 95%CI 0.73-1.02, P_{trend} = 0.04). Specifically higher intake of high-fat cheese (52% of high-fat fermented dairy, 83% of total cheese intake) was significantly associated with lower prediabetes risk (RR_{Q4vsQ1} 0.79, 95%CI 0.66-0.94, P_{trend} = 0.006). High-fat cheese was the only dairy type continuously associated with prediabetes ($RR_{serving/day}$ 0.94, 95%CI 0.88-1.00). Total dairy, fermented dairy, milk, plain milk, yogurt, cream, and ice cream intake both in quartiles and continuously were not associated with risk of prediabetes in multivariate adjusted models (RR ~ 1). Further adjustment for BMI, LDL cholesterol and blood pressure in model 4 did not change the associations.

Table 2. Baseline characteristics of participants in the Hoorn Studies in the total population and across quartiles (Q) of total dairy intake (*n* = 2,262).

	Total	**Q1**	**Q2**	**Q3**	**Q4**
	(*n* = 2,262)	(*n* = 562)	(*n* = 573)	(*n* = 561)	(*n* = 566)
Dairy intake (servings/day)	3.0 ± 1.7	1.2 ± 0.5	2.3 ± 0.3	3.3 ± 0.3	5.3 ± 1.3
Range	0-15.4	0-1.8	1.8-2.7	2.7-3.9	3.9-15.4
Median	2.7	1.3	2.3	3.3	5.0
Follow-up time (year)	6.4 ± 0.7	6.6 ± 0.7	6.4 ± 0.7	6.4 ± 0.7	6.2 ± 0.6
Sex (men)	50% (1132)	43% (243)	52% (296)	52% (289)	54% (304)
Age (year)	55.9 ± 7.3	54.1 ± 7.1	55.6 ± 7.3	56.6 ± 7.5	57.3 ± 7.1

Table 2. Baseline characteristics of participants in the Hoorn Studies in the total population and across quartiles (Q) of total dairy intake (*n* = 2,262). (continued)

	Total	**Q1**	**Q2**	**Q3**	**Q4**
	(*n* = 2,262)	(*n* = 562)	(*n* = 573)	(*n* = 561)	(*n* = 566)
Education level					
Low	13% (299)	8% (46)	15% (84)	14% (79)	16% (90)
Middle	58% (1310)	60% (336)	54% (309)	57% (317)	61% (348)
High	28% (631)	31% (172)	30% (174)	28% (159)	22% (126)
Smoking					
Current	22% (493)	20% (115)	23% (131)	22% (124)	22% (123)
Previous (> 2 months ago)	38% (851)	40% (226)	37% (211)	37% (208)	36% (206)
Never	40% (908)	38% (216)	40% (228)	41% (228)	42% (236)
Cigarette years	210 [9-480]	290 [100-560]	220 [30-470]	180 [0-450]	160 [0-460]
Alcohol intake					
0 g/day	18% (401)	15% (87)	18% (103)	17% (97)	20% (114)
≤ 10 g/day	42% (959)	42% (237)	43% (246)	40% (226)	44% (250)
10-30 g/day	30% (678)	30% (170)	31% (178)	32% (179)	27% (151)
≥ 30 g/day	10% (223)	12% (68)	8% (45)	11% (59)	9% (51)
Family history of diabetes	24% (553)	25% (141)	27% (152)	21% (119)	25% (141)
PA, moderate intensity, hours/week	7.5 [4.2-12.0]	6.5 [3.5-10.5]	7.0 [3.8-12.0]	7.8 [4.5-12.6]	9.0 [5.1-13.7]
BMI (kg/m2)	25.7 ± 3.4	25.9 ± 3.5	25.5 ± 3.2	25.5 ± 3.3	26.1 ± 3.4
Fasting glucose (mmol/L)	5.3 ± 0.4	5.3 ± 0.4	5.3 ± 0.4	5.2 ± 0.4	5.2 ± 0.4
Systolic blood pressure (mmHg)	130 ± 17	130 ± 16	130 ± 18	130 ± 17	130 ± 18
Diastolic blood pressure (mmHg)	78 ± 11	78 ± 11	78 ± 11	79 ± 10	80 ± 10
Antihypertensive medication use	13% (304)	16% (92)	12% (70)	13% (73)	12% (69)
LDL cholesterol (mmol/L)	3.8 ± 1.1	3.5 ± 1.0	3.8 ± 1.1	3.9 ± 1.1	4.2 ± 1.1
Lipid lowering medication	5% (108)	9% (50)	4% (22)	3% (18)	3% (18)
Dietary intake					
Energy intake (kcal/day)	2100 ± 600	1900 ± 540	2100 ± 530	2200 ± 570	2400 ± 620
DHD15-index score	70 ± 14	66 ± 14	71 ± 14	72 ± 14	70 ± 13
Fruit (g/day)	200 ± 140	160 ± 130	190 ± 130	200 ± 130	240 ± 150
Vegetables (g/day)	150 ± 85	170 ± 91	150 ± 83	150 ± 86	140 ± 75
Grain (g/day)	200 ± 95	200 ± 95	200 ± 92	200 ± 96	200 ± 96
Red meat (g/day)	34 ± 23	37 ± 23	35 ± 22	33 ± 23	31 ± 23
Processed meat (g/day)	46 ± 33	37 ± 30	44 ± 32	47 ± 33	54 ± 35
Lean fish (g/day)	11 ± 13	11 ± 12	11 ± 13	11 ± 12	12 ± 14
Fatty fish (g/day)	5.0 ± 8.4	4.9 ± 7.2	5.4 ± 10.6	4.7 ± 7.4	4.9 ± 8.1

Table 2. Baseline characteristics of participants in the Hoorn Studies in the total population and across quartiles (Q) of total dairy intake (*n* = 2,262). (continued)

	Total	**Q1**	**Q2**	**Q3**	**Q4**
	(*n* = 2,262)	(*n* = 562)	(*n* = 573)	(*n* = 561)	(*n* = 566)
Coffee (g/day)	500 ± 270	450 ± 290	490 ± 270	520 ± 250	540 ± 270
Tea (g/day)	280 ± 260	270 ± 290	280 ± 250	290 ± 250	300 ± 260
Fruit juice (g/day)	60 ± 95	69 ± 110	60 ± 89	54 ± 83	55 ± 98
SSBs (g/day)	110 ± 140	130 ± 160	110 ± 140	100 ± 130	110 ± 150
Saturated fat (en%)	14.6 ± 3.7	12.6 ± 3.3	14.2 ± 3.3	15.2 ± 3.4	16.6 ± 3.5
Protein (en%)	14.5 ± 2.4	13.8 ± 2.3	14.3 ± 2.3	14.6 ± 2.2	15.4 ± 2.5
Calcium (g/day)	1000 ± 370	650 ± 180	880 ± 160	1100 ± 160	1500 ± 300

Values are mean ± SD for continuous variables with a normal distribution, or median [IQR] for continuous variables with a skewed distribution, percentages for categorical variables, based on unimputed data. Abbreviations: BMI, body mass index; DHD15-index, Dutch Healthy Diet 2015 index score [56]; en%, percentage of total energy intake; LDL, low-density lipoprotein; PA, physical activity; SSBs, sugar-sweetened beverages.

Interactions were present between the exposures low-fat dairy (p = 0.03), low-fat fermented dairy (p = 0.01) and low-fat cheese (p < 0.0001), and enrolment wave. In stratified analysis, low-fat fermented dairy and low-fat cheese were associated with prediabetes in HS1 ($RR_{serving/day}$, respectively, 1.10, 95%CI 1.03-1.19 and 1.33, 95%CI 1.20-1.47) but not in HS2 ($RR_{serving/day}$ 0.92, 95%CI 0.81-1.05 and 0.92, 95%CI 0.81-1.05) (**Supplemental Table 4**). Low-fat cheese intake was much lower in HS1 as compared to HS2 (10.3% vs 46.2% of low-fat fermented dairy). Furthermore, interactions with age were present for low-fat dairy (p = 0.004), low-fat fermented dairy (p < 0.001), yogurt (p = 0.01), low-fat yogurt (p = 0.003), low-fat cheese (p = 0.002) and interactions were present with BMI for low-fat fermented dairy (p = 0.01) (**Supplemental Table 5**). Associations between these dairy exposures and prediabetes in participants aged 56 years and over, and in participants with a BMI ≥ 30 were similar as in HS1. Other stratified analyses were not significant.

None of the adjusted associations for substitution of high-fat cheese for alternative dairy products were significant (**Table 4**). We further examined potential confounding of inverse associations between high-fat cheese and prediabetes by assessing lifestyle and risk factors according to high-fat cheese intake (**Supplemental Table 6**). In the highest compared to lowest quartile of high-fat cheese intake, participants were more often male (53% vs 44%), current smoker (24% vs 17%), used less medication (antihypertensive 10% vs 17%; lipid lowering 2% vs 9%) and had higher LDL cholesterol levels (4.1 ± 1.1 vs 3.5 ± 1.0 mmol/L). Low-fat cheese intake was lower and intakes of all other dairy types, energy and processed meat were higher.

Table 3. The associations of dairy intake and prediabetes risk in the Hoorn Studies (n = 2,262).

	Relative risk (95% CI) across intake range categories[1]					Continuous[2]
	Q1	Q2	Q3	Q4	Ptrend	RR (95%CI)
Total dairy						
n cases/n total	216/562	216/573	183/561	196/566		
Median, servings/d	1.3	2.3	3.3	5.0		
Model 1	1	1.00 (0.86-1.16)	0.88 (0.75-1.03)	0.96 (0.82-1.14)	0.49	0.98 (0.95-1.02)
Model 2	1	1.00 (0.86-1.17)	0.87 (0.73-1.02)	0.95 (0.80-1.14)	0.41	0.98 (0.94-1.02)
Model 3	1	1.00 (0.86-1.16)	0.87 (0.74-1.03)	0.95 (0.79-1.13)	0.38	0.98 (0.94-1.02)
Model 4	1	1.01 (0.87-1.18)	0.88 (0.74-1.04)	0.93 (0.78-1.12)	0.30	0.97 (0.93-1.01)
High-fat dairy						
n cases/n total	217/565	216/567	198/563	180/567		
Median, servings/d	0.1	0.9	1.7	3.1		
Model 1	1	1.01 (0.87-1.18)	0.95 (0.81-1.12)	0.88 (0.74-1.05)	0.11	0.97 (0.92-1.01)
Model 2	1	1.00 (0.86-1.16)	0.92 (0.78-1.08)	0.84 (0.70-1.01)	0.04*	0.95 (0.91-1.00)
Model 3	1	1.01 (0.87-1.17)	0.94 (0.79-1.11)	0.85 (0.71-1.03)	0.06	0.96 (0.91-1.01)
Model 4	1	1.00 (0.86-1.16)	0.94 (0.80-1.11)	0.85 (0.71-1.03)	0.07	0.95 (0.91-1.00)
Low-fat dairy						
n cases/n total	203/566	210/561	208/576	190/559		
Median, servings/d	0.2	0.9	1.6	2.8		
Model 1	1	1.03 (0.89-1.21)	1.00 (0.86-1.16)	0.95 (0.81-1.12)	0.46	1.01 (0.96-1.06)
Model 2	1	1.06 (0.91-1.24)	1.02 (0.87-1.19)	0.98 (0.83-1.15)	0.67	1.01 (0.97-1.06)
Model 3	1	1.06 (0.91-1.23)	1.01 (0.87-1.18)	0.96 (0.82-1.13)	0.48	1.01 (0.96-1.06)
Model 4	1	1.06 (0.91-1.23)	1.00 (0.86-1.17)	0.94 (0.80-1.11)	0.35	1.00 (0.96-1.05)
Total fermented dairy						
n cases/n total	224/570	188/562	205/566	194/564		
Median, servings/d	0.7	1.5	2.4	3.7		
Model 1	1	0.86 (0.73-1.00)	0.94 (0.81-1.10)	0.94 (0.80-1.10)	0.73	0.98 (0.94-1.02)
Model 2	1	0.87 (0.75-1.02)	0.95 (0.81-1.10)	0.94 (0.79-1.11)	0.68	0.98 (0.94-1.02)

Table 3. The associations of dairy intake and prediabetes risk in the Hoorn Studies (n = 2,262). (continued)

	Relative risk (95% CI) across intake range categories[1]					Continuous[2]
	Q1	**Q2**	**Q3**	**Q4**	**Ptrend**	**RR (95%CI)**
Model 3	1	0.88 (0.75-1.03)	0.95 (0.81-1.11)	0.95 (0.80-1.13)	0.81	0.98 (0.94-1.03)
Model 4	1	0.89 (0.76-1.04)	0.95 (0.81-1.11)	0.94 (0.79-1.11)	0.65	0.98 (0.93-1.02)
High-fat fermented dairy						
n cases/n total	226/571	188/513	219/613	178/565		
Median, servings/d	0.0	0.7	1.4	2.6		
Model 1	1	0.94 (0.81-1.09)	0.93 (0.80-1.09)	0.85 (0.72-1.00)	0.05	0.96 (0.91-1.01)
Model 2	1	0.93 (0.79-1.08)	0.91 (0.78-1.07)	0.81 (0.68-0.97)	0.02*	0.94 (0.89-1.00)
Model 3	1	0.94 (0.80-1.09)	0.93 (0.79-1.08)	0.83 (0.69-0.99)	0.04*	0.95 (0.90-1.01)
Model 4	1	0.94 (0.80-1.10)	0.93 (0.80-1.09)	0.82 (0.68-0.98)	0.03*	0.95 (0.89-1.00)
Low-fat fermented dairy						
n cases/n total	207/565	204/566	203/557	197/574		
Median, servings/d	0.0	0.4	0.9	1.9		
Model 1	1	0.97 (0.83-1.13)	0.97 (0.83-1.13)	0.95 (0.81-1.11)	0.55	1.02 (0.97-1.09)
Model 2	1	0.99 (0.84-1.15)	0.99 (0.84-1.15)	0.97 (0.83-1.14)	0.73	1.03 (0.97-1.10)
Model 3	1	0.99 (0.85-1.16)	0.98 (0.84-1.15)	0.96 (0.82-1.13)	0.64	1.03 (0.97-1.10)
Model 4	1	0.99 (0.85-1.16)	0.99 (0.84-1.15)	0.96 (0.82-1.12)	0.57	1.03 (0.97-1.09)
Total milk						
n cases/n total	Q1	Q2	Q3	Q4		
Median, servings/d	197/562	206/570	222/566	186/564		
Median	0.1	0.6	1.2	2.1		
Model 1	1	1.04 (0.89-1.21)	1.15 (0.99-1.34)	1.00 (0.85-1.19)	0.83	1.01 (0.95-1.07)
Model 2	1	1.05 (0.90-1.22)	1.16 (1.00-1.35)	1.01 (0.85-1.20)	0.78	1.01 (0.95-1.08)
Model 3	1	1.02 (0.87-1.19)	1.13 (0.97-1.31)	0.96 (0.81-1.15)	0.79	1.00 (0.94-1.06)
Model 4	1	1.02 (0.88-1.19)	1.14 (0.98-1.32)	0.96 (0.81-1.14)	0.81	0.99 (0.94-1.06)

Table 3. The associations of dairy intake and prediabetes risk in the Hoorn Studies (n = 2,262). (continued)

	Relative risk (95% CI) across intake range categories[1]					Continuous[2]
	Q1	Q2	Q3	Q4	Ptrend	RR (95%CI)
High-fat milk						
n cases/n total	497/1,339	97/307	82/248	135/368		
Median, servings/d	0.0	0.0	0.2	0.5		
Model 1	1	0.89 (0.74-1.06)	0.99 (0.80-1.22)	1.08 (0.91-1.28)	0.27	1.05 (0.93-1.18)
Model 2	1	0.90 (0.76-1.08)	1.00 (0.81-1.23)	1.06 (0.89-1.25)	0.43	1.03 (0.92-1.16)
Model 3	1	0.91 (0.76-1.09)	1.01 (0.82-1.24)	1.04 (0.87-1.24)	0.55	1.01 (0.89-1.14)
Model 4	1	0.91 (0.76-1.09)	1.01 (0.82-1.24)	1.06 (0.89-1.26)	0.42	1.02 (0.91-1.15)
Low-fat milk						
n cases/n total	203/569	185/520	198/499	225/674		
Median, servings/d	0.0	0.4	0.8	1.7		
Model 1	1	1.01 (0.86-1.18)	1.06 (0.91-1.22)	1.00 (0.84-1.18)	0.96	1.00 (0.94-1.07)
Model 2	1	1.03 (0.88-1.21)	1.07 (0.93-1.24)	1.02 (0.86-1.21)	0.79	1.01 (0.94-1.07)
Model 3	1	1.01 (0.86-1.19)	1.04 (0.90-1.21)	0.98 (0.83-1.17)	0.88	0.99 (0.93-1.06)
Model 4	1	1.01 (0.86-1.18)	1.03 (0.89-1.20)	0.98 (0.82-1.15)	0.80	0.99 (0.93-1.05)
Total plain milk						
n cases/n total	183/534	187/521	290/755	151/452		
Median, servings/d	0.0	0.3	1.0	2.1		
Model 1	1	1.08 (0.92-1.28)	1.14 (0.98-1.32)	1.07 (0.89-1.29)	0.41	1.02 (0.96-1.08)
Model 2	1	1.09 (0.93-1.29)	1.15 (0.99-1.33)	1.07 (0.89-1.30)	0.42	1.02 (0.96-1.09)
Model 3	1	1.07 (0.91-1.26)	1.11 (0.95-1.28)	1.02 (0.85-1.24)	0.79	1.00 (0.94-1.07)
Model 4	1	1.07 (0.91-1.25)	1.11 (0.96-1.29)	1.01 (0.84-1.23)	0.83	1.00 (0.94-1.07)
High-fat plain milk						
n cases/n total	711/1,989	30/91	34/82	36/100		
Median, servings/d	0.0	0.2	0.7	1.8		
Model 1	1	0.98 (0.72-1.32)	1.19 (0.91-1.55)	1.07 (0.82-1.41)	0.41	1.04 (0.92-1.17)
Model 2	1	0.95 (0.70-1.28)	1.17 (0.90-1.53)	1.04 (0.79-1.37)	0.56	1.02 (0.90-1.16)

Table 3. The associations of dairy intake and prediabetes risk in the Hoorn Studies (n = 2,262). (continued)

	Relative risk (95% CI) across intake range categories[1]					Continuous[2]
	Q1	Q2	Q3	Q4	Ptrend	RR (95%CI)
Model 3	1	0.94 (0.69-1.27)	1.18 (0.90-1.54)	0.99 (0.75-1.31)	0.81	1.00 (0.88-1.13)
Model 4	1	0.94 (0.70-1.27)	1.19 (0.90-1.57)	1.03 (0.78-1.36)	0.61	1.01 (0.89-1.15)
Low-fat plain milk						
n cases/n total	212/586	170/512	236/598	193/566		
Median, servings/d	0.0	0.3	0.7	1.9		
Model 1	1	0.94 (0.80-1.11)	1.12 (0.97-1.30)	0.99 (0.85-1.16)	0.87	1.01 (0.95-1.08)
Model 2	1	0.95 (0.80-1.12)	1.13 (0.98-1.31)	1.00 (0.85-1.17)	0.81	1.02 (0.96-1.09)
Model 3	1	0.93 (0.79-1.09)	1.10 (0.95-1.28)	0.97 (0.82-1.13)	0.87	1.01 (0.94-1.07)
Model 4	1	0.91 (0.78-1.07)	1.10 (0.95-1.27)	0.95 (0.81-1.12)	0.79	1.00 (0.94-1.07)
Total yogurt						
n cases/n total	201/564	196/538	205/583	209/577		
Median, servings/d	0.0	0.4	0.6	1.0		
Model 1	1	1.00 (0.86-1.18)	0.99 (0.85-1.15)	1.01 (0.86-1.18)	0.98	1.02 (0.91-1.14)
Model 2	1	1.03 (0.88-1.21)	1.01 (0.86-1.18)	1.04 (0.89-1.22)	0.69	1.05 (0.93-1.17)
Model 3	1	1.05 (0.89-1.23)	1.02 (0.87-1.20)	1.04 (0.88-1.21)	0.73	1.04 (0.93-1.17)
Model 4	1	1.04 (0.89-1.22)	1.03 (0.88-1.21)	1.05 (0.90-1.23)	0.56	1.06 (0.94-1.18)
High-fat yogurt						
n cases/n total	550/1,575	86/229	88/219	87/239		
Median, servings/d	0.0	0.1	0.3	0.8		
Model 1	1	1.08 (0.90-1.29)	1.14 (0.95-1.36)	1.08 (0.90-1.29)	0.25	1.07 (0.91-1.27)
Model 2	1	1.09 (0.91-1.31)	1.16 (0.97-1.39)	1.08 (0.90-1.30)	0.24	1.08 (0.91-1.29)
Model 3	1	1.10 (0.92-1.32)	1.19 (0.99-1.43)	1.09 (0.91-1.32)	0.19	1.09 (0.92-1.29)
Model 4	1	1.08 (0.90-1.29)	1.20 (1.01-1.44)	1.11 (0.92-1.34)	0.13	1.12 (0.94-1.34)
Low-fat yogurt						
n cases/n total	256/724	223/604	152/416	180/518		
Median, servings/d	0.0	0.2	0.6	0.9		

Table 3. The associations of dairy intake and prediabetes risk in the Hoorn Studies (n = 2,262). (continued)

	Relative risk (95% CI) across intake range categories[1]					Continuous[2]
	Q1	Q2	Q3	Q4	Ptrend	RR (95%CI)
Model 1	1	1.01 (0.87-1.18)	1.00 (0.85-1.18)	0.97 (0.84-1.13)	0.70	0.99 (0.88-1.12)
Model 2	1	1.04 (0.89-1.21)	1.02 (0.86-1.20)	1.00 (0.86-1.17)	0.94	1.01 (0.90-1.15)
Model 3	1	1.04 (0.90-1.21)	1.03 (0.87-1.21)	0.99 (0.85-1.16)	0.82	1.01 (0.89-1.14)
Model 4	1	1.02 (0.88-1.19)	1.02 (0.87-1.20)	0.99 (0.85-1.15)	0.89	1.01 (0.89-1.14)
Total cheese						
n cases/n total	215/563	216/561	202/581	178/557		
Median, servings/d	0.3	0.9	1.5	2.6		
Model 1	1	1.01 (0.87-1.17)	0.91 (0.78-1.07)	0.87 (0.74-1.02)	0.04*	0.96 (0.91-1.01)
Model 2	1	1.01 (0.87-1.17)	0.90 (0.77-1.04)	0.84 (0.71-0.99)	0.02*	0.94 (0.89-1.00)
Model 3	1	1.02 (0.88-1.19)	0.90 (0.77-1.05)	0.86 (0.73-1.02)	0.04*	0.95 (0.90-1.01)
Model 4	1	1.02 (0.88-1.18)	0.89 (0.76-1.03)	0.84 (0.71-1.00)	0.02*	0.94 (0.89-1.00)
High-fat cheese						
n cases/n total	223/558	225/600	188/532	175/572		
Median, servings/d	0.0	0.6	1.2	2.3		
Model 1	1	0.96 (0.83-1.11)	0.92 (0.79-1.08)	0.81 (0.69-0.95)	0.009**	0.94 (0.89-1.00)
Model 2	1	0.94 (0.82-1.09)	0.90 (0.77-1.06)	0.77 (0.65-0.92)	0.002**	0.93 (0.88-0.99)*
Model 3	1	0.95 (0.82-1.10)	0.91 (0.77-1.07)	0.79 (0.66-0.94)	0.006**	0.94 (0.88-1.00)*
Model 4	1	0.95 (0.82-1.11)	0.91 (0.77-1.07)	0.78 (0.65-0.93)	0.003**	0.93 (0.88-0.99)*
Low-fat cheese						
n cases/n total	533/1,540	85/223	94/258	99/241		
Median, servings/d	0.0	0.2	0.6	1.2		
Model 1	1	1.03 (0.85-1.24)	0.98 (0.81-1.18)	1.10 (0.92-1.31)	0.39	1.03 (0.93-1.13)
Model 2	1	1.03 (0.85-1.26)	0.99 (0.82-1.19)	1.09 (0.91-1.30)	0.40	1.03 (0.93-1.13)
Model 3	1	1.03 (0.85-1.25)	0.99 (0.82-1.19)	1.12 (0.93-1.34)	0.30	1.04 (0.94-1.16)
Model 4	1	1.01 (0.83-1.22)	0.97 (0.80-1.16)	1.09 (0.91-1.30)	0.45	1.03 (0.93-1.14)

Table 3. The associations of dairy intake and prediabetes risk in the Hoorn Studies (n = 2,262). (continued)

	Relative risk (95% CI) across intake range categories[1]					Continuous[2]
	Q1	Q2	Q3	Q4	Ptrend	RR (95%CI)
Cream						
n cases/n total	341/938	172/421	151/466	147/437		
Median, servings/d	0.0	0.3	0.4	1.4		
Model 1	1	1.01 (0.86-1.20)	0.88 (0.75-1.03)	0.88 (0.74-1.04)	0.09	0.98 (0.96-1.01)
Model 2	1	1.03 (0.87-1.22)	0.89 (0.76-1.05)	0.87 (0.73-1.03)	0.06	0.98 (0.96-1.01)
Model 3	1	1.02 (0.86-1.21)	0.89 (0.76-1.05)	0.87 (0.74-1.03)	0.08	0.98 (0.96-1.01)
Model 4	1	1.03 (0.87-1.21)	0.90 (0.77-1.05)	0.87 (0.73-1.03)	0.06	0.98 (0.96-1.01)
Ice cream						
n cases/n total	263/718	98/280	287/783	163/481		
Median, servings/d	0.00	0.02	0.05	0.13		
Model 1	1	1.04 (0.86-1.27)	0.99 (0.87-1.13)	0.96 (0.82-1.12)	0.49	0.82 (0.44-1.52)
Model 2	1	1.05 (0.86-1.27)	0.99 (0.87-1.14)	0.95 (0.81-1.12)	0.47	0.79 (0.43-1.48)
Model 3	1	1.06 (0.87-1.28)	0.99 (0.86-1.13)	0.95 (0.81-1.12)	0.44	0.79 (0.42-1.48)
Model 4	1	1.04 (0.86-1.26)	0.98 (0.86-1.12)	0.91 (0.78-1.08)	0.22	0.72 (0.38-1.38)

[1] Relative risks (95CIs) were estimated across four categories split by quartile values (Q1 to Q4) or non-consumers + tertile categories with the lowest category as the reference, adjusted for covariates as follows: Model 1 included age (continuous), sex, follow-up duration (continuous) and cohort (2 categories). Model 2 was additionally adjusted for energy intake (continuous), education (3 categories), smoking (3 categories), physical activity (3 categories), alcohol consumption (continuous), and a family history of diabetes (yes/no). Model 3 was additionally adjusted for food groups associated with type 2 diabetes, including intakes of fruit, vegetables, tea, coffee, grains (whole and refined), meat (processed and red) and sugar-sweetened beverages (SSB) (continuous). Model 4 included BMI (continuous), LDL cholesterol (continuous), systolic and diastolic blood pressure (continuous). Linear trend across intake range categories was assessed by including median values of each category as a continuous variable in the model.

[2] Relative risks per 1 serving/day (see definition in Table 1) were estimated.

P-value significance level: *0.05, **0.01, ***0.001.

Abbreviations: CI, Confidence Interval; Q, Quartile.

Table 4. The associations of substitution of high-fat cheese with alternative dairy products and prediabetes risk in the Hoorn Studies (*n* = 2,262).

High-fat cheese	Continuous relative risk (95%CI)[1]
High-fat milk	1.06 (0.93-1.21)
Low-fat milk	1.06 (0.96-1.16)
High-fat yogurt	1.20 (0.99-1.46)
Low-fat yogurt	1.09 (0.95-1.26)
Low-fat cheese	1.09 (0.98-1.22)
Cream	1.05 (0.98-1.12)
Ice cream	0.82 (0.44-1.54)

[1] Continuous relative risks per 1 serving/day (see definition in Table 1) were estimated, adjusted for covariates as follows: age (continuous), sex, follow-up duration, cohort, education (3 categories), smoking (3 categories), physical activity (continuous), alcohol consumption (4 categories), family history of diabetes (yes/no), intakes of fruit, vegetables, tea, coffee, grains (whole and refined), meat (processed and red) and sugar-sweetened beverages (continuous). The substitution model included total servings/day of dairy intake and energy intake (kcal).
Abbreviations: CI, Confidence Interval.

Sensitivity analyses

Mutual adjustment for intake of all other dairy types did not result in different associations (**Supplemental Table 7**). In the sample without CVD at baseline, associations between high-fat fermented, total cheese and high-fat cheese and prediabetes had similar effect estimates compared to associations in all participants, but these were no longer significant. Using Goldberg cut-offs, we identified 481 'under reporters' (21%) and 62 'over reporters' (3%). Repeating analyses in 1,716 'normal reporters' (76%) resulted in stronger associations for high-fat fermented dairy (RR_{Q4vsQ1} 0.79, 95%CI 0.65-0.97, P_{trend} = 0.04), total cheese (RR_{Q4vsQ1} 0.78, 95%CI 0.64-0.96, P_{trend} = 0.01, $RR_{serving/day}$ 0.92, 95%CI 0.86-0.99) and high-fat cheese (RR_{Q4vsQ1} 0.73, 95%CI 0.60-0.90, P_{trend} = 0.003, $RR_{serving/day}$ 0.92, 95%CI 0.86-0.99). In 'normal reporters', continuous associations were significant for total dairy ($RR_{serving/day}$ 0.95, 95%CI 0.91-0.99) and total fermented dairy ($RR_{serving/day}$ 0.95, 95%CI 0.90-1.00) and prediabetes, although associations in quartiles were not. In analysis including participants with prediabetes at baseline, only high-fat cheese remained significantly associated with prediabetes (RR_{Q4vsQ1} 0.85, 95%CI 0.73-0.98). Associations between high-fat fermented and total cheese and prediabetes were attenuated but remained in the same direction. Sensitivity analyses of substitution of one serving/day of high-fat cheese with alternative dairy types resulted in similar associations as in main analysis (**Supplemental Table 8**).

Discussion

A high intake of high-fat fermented dairy, total cheese and high-fat cheese were associated with lower risk of prediabetes in this population-based cohort. Associations were driven by high-fat cheese intake, as 52% and 83% of, respectively, high-fat fermented

dairy and total cheese intake consisted of high-fat cheese. We found no associations for substitutions of high-fat cheese with other dairy products and risk of prediabetes. Total dairy and milk, yogurt, cream, and ice cream intake were not associated with prediabetes, irrespective of fat content.

Results in context

Our observed associations of high-fat cheese with prediabetes were consistent with the prospective FHS Offspring Cohort [9]. They reported a non-significant association between the highest and lowest cheese intake (HR 0.86, 95%CI 0.69-1.07), although the median cheese intake in their highest category was substantially lower than ours (median < 1 vs 2.6 servings/day). The cross-sectional Dutch Maastricht study reported an association for higher cheese intake with lower prediabetes risk ($OR_{20g/day}$ 0.88, 95%CI 0.80-0.97) [13]. Similar inverse associations of cheese and 2hPG were also found in the longitudinal Inter99 study ($\beta_{20g/day}$ -0.05, 95%CI -0.01; -0.001 mmol/L) [11] and in the cross-sectional ELSA-Brasil ($\beta_{30/day}$ -0.05, 95%CI -0.09; -0.02 mmol/L) [15]. Our findings are in line with a review of three meta-analyses reporting moderate evidence of a prospective association between higher cheese intake and lower T2D risk [8], and the prospective Urban Rural Epidemiology (PURE) study including 131,481 individuals from 21 countries showing a 24% lower risk at > 1 servings/day compared to 0 servings/day (HR 0.76, 95%CI 0.64-0.91, P_{trend} = 0.001) [31]. Overall, current evidence indicates that higher cheese consumption is associated with lower prediabetes and T2D risk.

Our results of no associations between total, low-fat and high-fat dairy and prediabetes are not in accordance with results from the FHS Offspring Cohort, which reported a lower risk of, respectively, 39%, 32% and 25% in the highest compared to lowest intake category [9]. These associations with prediabetes risk were driven by moderate consumption of low-fat, skim milk and whole milk consumption, none of which were significant in our study. Several explanations may underscore these different findings, including a longer follow-up in their study (12 years vs 6.4 years), a higher prediabetes incidence (48.3% vs 35.9%), and a different prediabetes definition (no use of 2hPG and HbA1c, lower cut-off of 5.6 vs 6.0 mmol/L for FPG). Furthermore, they used repeated measurements of dairy intake during follow- up to account for within-person variability, which may have strengthened the associations. Also, in US populations, high dairy consumption is associated with an overall healthier dietary pattern [32], whereas in Europe, dairy consumption is more widespread across a range of populations.

Our results pointed towards effect modification of associations of low-fat dairy types and yogurt with prediabetes by enrolment wave, age, and BMI. Positive associations for low-fat dairy types and prediabetes were shown in HS1, but not in HS2, which could be explained by differences in low-fat dairy consumption patterns, such as negligible low-fat cheese consumption in HS1. Changes in dairy intake observed between the two enrolment waves correspond to the changes observed from 1987/88 to 2007/10

as assessed in the Dutch National Food Consumption Survey [33]. They showed an increase in intake of dairy, especially of low-fat types, skimmed and semi-skimmed yogurt and cheese. Furthermore, differences in characteristics between the enrolment waves could explain effect modification, especially as the HS1 population was somewhat older compared to HS2. Effect modification by enrolment wave may explain positive associations of low-fat dairy types and prediabetes present in those with higher age and BMI, as these are not in line with previous studies [9, 34].

We found no associations of yogurt consumption with prediabetes, yet especially higher yogurt consumption has been associated to lower T2D risk in previous high-quality research [8]. Significant nonlinear associations of yogurt with prediabetes were found in the FHS Offspring Cohort, with 25% risk reduction observed for 1 to ≤ 3 servings/ week compared to no consumption, yet risk increased with higher intakes [9]. Inverse associations between yogurt and pre- diabetes were also reported by two cross-sectional studies [13, 15], but no associations were found in the much larger Lifelines study [14]. Our neutral observations for yogurt could be due to the inclusion of porridge and custard in this category for HS2, as the HS2 FFQ combined these in a single question. Furthermore, yogurt consumption has been related to healthier diet and lifestyle [13, 35], and residual confounding could explain discrepancies in results across studies.

There are several mechanisms that may explain the observed associations of cheese with prediabetes. Despite an average fat content of 24-35 g/100 g [36] of which 70% SFAs, various RCTs demonstrated less adverse effects of SFAs contained within the cheese matrix compared to SFAs in different matrixes [36-38]. Beneficial associations have been found between ruminant trans fatty acids and insulin resistance and T2D [39, 40], with potential mechanisms suggested in animal studies being inhibition of hepatic de novo lipogenesis, activation of PPAR-α and PPAR-γ, improving insulin sensitivity and reducing inflammation [40, 41]. As shown by meta-analysis of 15 RCTs [42], calcium may affect energy balance by increasing faecal fat excretion, due to formation of insoluble Ca-fatty acid soaps and/or formation of hydrophobic aggregations. However, in a meta-analysis of 20 RCTs, calcium supplementation did not reduce body weight or body fat [43]. Fermented foods contain lactic acid bacteria and bioactive molecules, which are beneficial for viability and composition of the gut microbiota and influence gene expression related to glucose and insulin metabolism [41, 44]. Furthermore, vitamin K2 (menaquinones) in dairy is synthesized during fermentation, and cheese is the richest source of vitamin K2 in Western diets (12.7 μg/20 g) [45]. Higher vitamin K2 intake has been related to lower T2D risk in the Dutch EPIC cohort [46]. In animal models of T2D, vitamin K2 supplementation showed dose-dependent reductions of HbA1c and FPG and improved insulin resistance and β-cell function [47, 48]. Vitamin K2 may upregulate carboxylated osteocalcin, resulting in increased serum adiponectin levels, which enhances insulin sensitivity through increased fatty acid oxidation in skeletal muscles

and inhibition of hepatic glucose production in the liver [49]. Whether the vitamin K2 induced pathways underline long-term effects of cheese warrants further investigation.

Despite multiple studies pointing to a role of cheese in diabetes prevention, the exact place of cheese in healthy diets is unclear. Current American and European dietary guidelines only advise low-fat cheese to limit intake of SFAs and sodium [50], although evidence of more favourable associations of low-fat cheese with cardiometabolic outcomes is lacking [51]. The limited evidence available is not sufficient to justify changes to dietary guidelines, and additional well- designed controlled trials are needed.

Strengths and limitations

This study has several strengths, including the assessment of a wide range of dairy subtypes and possibility to disseminate low and high-fat types, also for yogurt and cheese. Other strengths include the longitudinal design with 6 years of follow-up and extensive adjustments for confounders. Despite these notable strengths, there are certain limitations that should be mentioned. First, although both FFQs were validated [19, 20], measurement errors in reported dietary intake due to recall bias are unavoidable. We corrected for energy misreporting and found slightly stronger associations in 'normal reporters', indicating attenuated effect sizes in main analyses due to energy misreporting. The FFQs used in the two waves were overall comparable, and slight differences in brands and dairy products included largely reflect the changes in dietary patterns between two time periods [33], for example the inclusion of more low-fat cheese types in the HS2 cohorts. No compositional changes in dairy products were observed between both NEVO-tables used for calculating nutrient intakes by the FFQ. The main findings and conclusions were similar when stratifying associations by enrolment wave. Second, no repeated measurements of diet were available, and although dietary patterns have shown to be somewhat stable [52, 53], for dairy product intake specifically in comparison to other food groups [53], errors in single dietary measurements may result in bias of associations towards the null. We addressed reverse causality due to dietary changes related to diagnoses of disease by excluding participants with history of diabetes, and excluding participants with prevalent CVD, resulting in similar associations although no longer significant, likely because of less power. Lastly, identification of prediabetes cases was less sensitive in HS2 compared to HS1, because no OGTT was done at follow-up and capillary sample HbA1c levels were used, which tend to be higher than venous sample HbA1c levels [54]. Furthermore, reproducibility of FPG and 2hPG is only moderate [55], and participants may revert to normoglycaemia during follow-up [5]. We addressed this possible misclassification by including participants with prediabetes at baseline in sensitivity analysis, showing similar significant associations for high-fat cheese in quartiles, and other associations remained in similar direction.

Conclusions

In conclusion, higher consumption of high-fat cheese was continuously associated with lower risk of prediabetes, suggesting a potential preventive role of cheese in T2D development. Observed associations between higher high-fat fermented dairy and total cheese consumption and lower prediabetes risk were likely driven by high-fat cheese consumption. We found no associations between total dairy, fermented dairy, milk, yogurt, cream and ice cream, regard- less of fat content, with prediabetes development. Further prospective and intervention research is needed to elucidate health effects of cheese considering high SFAs and sodium content, and its place in healthy diets.

List of supplementary materials chapter 2

Supplemental Table 1. Baseline characteristics and dairy intakes of participants in the Hoorn Studies across enrolment waves (*n* = 2,262).

Supplemental table 2. Missing values of covariates in participants before imputation.

Supplementary Table 3. Baseline characteristics and dairy intakes of participants in the Hoorn Studies according to whether participants had complete follow-up or were lost-to-follow-up.

Supplemental table 4. The associations of dairy intake and prediabetes risk in the Hoorn Studies, stratified for the first enrolment wave in 1989-1992 (Hoorn Study 1, HS1, *n* = 997) and a second wave in 2006-2007 (HS2, *n* = 1,265).

Supplemental Table 5. Associations of dairy product types and prediabetes risk in the Hoorn Studies, stratified by age and BMI (*n* = 2,262).

Supplemental Table 6. Baseline characteristics of participants in the Hoorn Studies in the total population and across quartiles) of high-fat cheese intake (*n* = 2,262).

Supplemental Table 7. Sensitivity analyses for the association between dairy intake and prediabetes risk the Hoorn Studies (*n* = 2,262).

Supplemental table 8. Sensitivity analyses for associations of substitution of high-fat cheese with alternative dairy products and prediabetes risk in the Hoorn Studies.

Scan this QR code to download the supplementary materials.

References

1. World Health Organization (WHO), *Definition and diagnosis of diabetes mellitus and intermediate hyperglycaemia: report of a WHO/IDF consultation.* 2006.
2. Saeedi, P., et al., *Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9(th) edition.* Diabetes Res Clin Pract, 2019. **157**: p. 107843.
3. Gerstein, H.C., et al., *Annual incidence and relative risk of diabetes in people with various categories of dysglycemia: a systematic overview and meta-analysis of prospective studies.* Diabetes Res Clin Pract, 2007. **78**(3): p. 305-12.
4. Yeboah, J., et al., *Impaired fasting glucose and the risk of incident diabetes mellitus and cardiovascular events in an adult population: MESA (Multi-Ethnic Study of Atherosclerosis).* Journal of the American College of Cardiology, 2011. **58**(2): p. 140-146.
5. Perreault, L., et al., *Effect of regression from prediabetes to normal glucose regulation on long-term reduction in diabetes risk: results from the Diabetes Prevention Program Outcomes Study.* The Lancet, 2012. **379**(9833): p. 2243-2251.
6. Long, G.H., et al., *Healthy behaviours and 10-year incidence of diabetes: a population cohort study.* Prev Med, 2015. **71**: p. 121-7.
7. Hu, F.B., et al., *Diet, lifestyle, and the risk of type 2 diabetes mellitus in women.* N Engl J Med, 2001. **345**(11): p. 790-7.
8. Alvarez-Bueno, C., et al., *Effects of Milk and Dairy Product Consumption on Type 2 Diabetes: Overview of Systematic Reviews and Meta-Analyses.* Adv Nutr, 2019. **10**(suppl_2): p. S154-S163.
9. Hruby, A., et al., *Associations of Dairy Intake with Incident Prediabetes or Diabetes in Middle-Aged Adults Vary by Both Dairy Type and Glycemic Status.* J Nutr, 2017. **147**(9): p. 1764-1775.
10. Fumeron, F., 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): p. 813-7.
11. Struijk, E.A., et al., *Dairy product intake in relation to glucose regulation indices and risk of type 2 diabetes.* Nutr Metab Cardiovasc Dis, 2013. **23**(9): p. 822-8.
12. Snijder, M.B., et al., *A prospective study of dairy consumption in relation to changes in metabolic risk factors: the Hoorn Study.* Obesity (Silver Spring), 2008. **16**(3): p. 706-9.
13. Eussen, S.J., et al., *Consumption of dairy foods in relation to impaired glucose metabolism and type 2 diabetes mellitus: the Maastricht Study.* Br J Nutr, 2016. **115**(8): p. 1453-61.
14. Brouwer-Brolsma, E.M., et al., *Dairy product consumption is associated with pre-diabetes and newly diagnosed type 2 diabetes in the Lifelines Cohort Study.* Br J Nutr, 2018. **119**(4): p. 442-455.
15. Drehmer, M., 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): p. 775-82.
16. Rutters, F., et al., *Cohort Profile: The Hoorn Studies.* Int J Epidemiol, 2018. **47**(2): p. 396-396j.
17. *International Expert Committee report on the role of the A1C assay in the diagnosis of diabetes.* Diabetes Care, 2009. **32**(7): p. 1327-34.
18. Lakshmy, R. and R. Gupta, *Measurement of glycated hemoglobin A1c from dried blood by turbidimetric immunoassay.* Journal of diabetes science and technology, 2009. **3**(5): p. 1203-1206.
19. Grootenhuis, P.A., et al., *A semiquantitative food frequency questionnaire for use in epidemiologic research among the elderly: Validation by comparison with dietary history.* Journal of Clinical Epidemiology, 1995. **48**(7): p. 859-868.
20. Siebelink, E., A. Geelen, and J.H. de Vries, *Self-reported energy intake by FFQ compared with actual energy intake to maintain body weight in 516 adults.* Br J Nutr, 2011. **106**(2): p. 274-81.
21. Streppel, M.T., et al., *Relative validity of the food frequency questionnaire used to assess dietary intake in the Leiden Longevity Study.* Nutr J, 2013. **12**: p. 75.
22. Wendel-Vos, G.C., et al., *Reproducibility and relative validity of the short questionnaire to assess health-enhancing physical activity.* J Clin Epidemiol, 2003. **56**(12): p. 1163-9.

23. Friedewald, W.T., R.I. Levy, and D.S. Fredrickson, *Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge.* Clinical chemistry, 1972. **18**(6): p. 499-502.
24. Knol, M.J., et al., *Overestimation of risk ratios by odds ratios in trials and cohort studies: alternatives to logistic regression.* CMAJ, 2012. **184**(8): p. 895-9.
25. Soedamah-Muthu, S.S. and J. de Goede, *Dairy Consumption and Cardiometabolic Diseases: Systematic Review and Updated Meta-Analyses of Prospective Cohort Studies.* Curr Nutr Rep, 2018. **7**(4): p. 171-182.
26. *Health Council of the Netherlands. Dutch dietary guidelines 2015. The Hague: Health Council of the Netherlands, 2015; publication no. 2015/24E.*
27. Schwingshackl, L., et al., *Food groups and risk of type 2 diabetes mellitus: a systematic review and meta-analysis of prospective studies.* Eur J Epidemiol, 2017. **32**(5): p. 363-375.
28. Laursen, A.S.D., et al., *Substitutions of dairy product intake and risk of stroke: a Danish cohort study.* Eur J Epidemiol, 2018. **33**(2): p. 201-212.
29. Black, A.E., *Critical evaluation of energy intake using the Goldberg cut-off for energy intake:basal metabolic rate. A practical guide to its calculation, use and limitations.* Int J Obes Relat Metab Disord, 2000. **24**(9): p. 1119-30.
30. Schofield, W.N., *Predicting basal metabolic rate, new standards and review of previous work.* Human nutrition. Clinical nutrition, 1985. **39**: p. 5-41.
31. Bhavadharini, B., et al., *Association of dairy consumption with metabolic syndrome, hypertension and diabetes in 147 812 individuals from 21 countries.* BMJ Open Diabetes Research & Care, 2020. **8**(1): p. e000826.
32. *USDA, HHS (2016) 2015-2020 Dietary Guidelines - health.gov.* . [cited 2020 3 July]; Available from: http://health.gov/dietaryguidelines/2015/guidelines/.
33. Rijksinstituut voor Volksgezondheid en Milieu (RIVM), *Veranderingen in het aanbod van voedingsmiddelen en de voedselconsumptie.* 2013.
34. Gijsbers, L., et al., *Consumption of dairy foods and diabetes incidence: a dose-response meta-analysis of observational studies.* Am J Clin Nutr, 2016. **103**(4): p. 1111-24.
35. Tremblay, A. and S. Panahi, *Yogurt consumption as a signature of a healthy diet and lifestyle.* The Journal of Nutrition, 2017. **147**(7): p. 1476S-1480S.
36. Thorning, T.K., et al., *Whole dairy matrix or single nutrients in assessment of health effects: current evidence and knowledge gaps.* Am J Clin Nutr, 2017. **105**(5): p. 1033-1045.
37. Feeney, E.L., et al., *Dairy matrix effects: response to consumption of dairy fat differs when eaten within the cheese matrix-a randomized controlled trial.* Am J Clin Nutr, 2018. **108**(4): p. 667-674.
38. de Goede, J., et al., *Effect of cheese consumption on blood lipids: a systematic review and meta-analysis of randomized controlled trials.* Nutr Rev, 2015. **73**(5): p. 259-75.
39. Imamura, F., 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. **15**(10): p. e1002670.
40. Tremblay, B.L. and I. Rudkowska, *Nutrigenomic point of view on effects and mechanisms of action of ruminant trans fatty acids on insulin resistance and type 2 diabetes.* Nutr Rev, 2017. **75**(3): p. 214-223.
41. Mozaffarian, D. and J.H.Y. Wu, *Flavonoids, Dairy Foods, and Cardiovascular and Metabolic Health: A Review of Emerging Biologic Pathways.* Circ Res, 2018. **122**(2): p. 369-384.
42. Christensen, R., et al., *Effect of calcium from dairy and dietary supplements on faecal fat excretion: a meta-analysis of randomized controlled trials.* Obes Rev, 2009. **10**(4): p. 475-86.
43. Booth, A.O., et al., *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): p. 1013-25.
44. Fernandez, M.A. and A. Marette, *Novel perspectives on fermented milks and cardiometabolic health with a focus on type 2 diabetes.* Nutr Rev, 2018. **76**(Suppl 1): p. 16-28.
45. Schurgers, L.J. and C. Vermeer, *Determination of phylloquinone and menaquinones in food. Effect of food matrix on circulating vitamin K concentrations.* Haemostasis, 2000. **30**(6): p. 298-307.
46. Beulens, J.W., et al., *Dietary phylloquinone and menaquinones intakes and risk of type 2 diabetes.* Diabetes Care, 2010. **33**(8): p. 1699-705.

47. Hussein, A.G., et al., *Vitamin K2 alleviates type 2 diabetes in rats by induction of osteocalcin gene expression.* Nutrition, 2018. **47**: p. 33-38.
48. Iwamoto, J., et al., *Vitamin K 2 prevents hyperglycemia and cancellous osteopenia in rats with streptozotocin-induced type 1 diabetes.* Calcified tissue international, 2011. **88**: p. 162-168.
49. Li, Y., et al., *Effect of vitamin K2 on type 2 diabetes mellitus: A review.* Diabetes Res Clin Pract, 2018. **136**: p. 39-51.
50. Soedamah-Muthu, S. and J. Guo, *Dairy consumption and cardiometabolic diseases: Evidence from prospective studies. In D. I. Givens (Ed.), Milk and dairy foods: Their functionality in human health and disease (pp. 1-28). Academic Press.* 2020.
51. Drouin-Chartier, J.P., et al., *Systematic Review of the Association between Dairy Product Consumption and Risk of Cardiovascular-Related Clinical Outcomes.* Adv Nutr, 2016. **7**(6): p. 1026-1040.
52. Jankovic, N., et al., *Stability of dietary patterns assessed with reduced rank regression; the Zutphen Elderly Study.* Nutrition journal, 2014. **13**(1): p. 1-9.
53. Nagel, G., et al., *Long-term reproducibility of a food-frequency questionnaire and dietary changes in the European Prospective Investigation into Cancer and Nutrition (EPIC)-Heidelberg cohort.* Br J Nutr, 2007. **98**(1): p. 194-200.
54. Geethanjali, F., R.S. Kumar, and M. Seshadri, *Accuracy of filter paper method for measuring glycated hemoglobin.* The Journal of the Association of Physicians of India, 2007. **55**: p. 115-119.
55. Kuzuya, T., et al., *Report of the Committee on the classification and diagnostic criteria of diabetes mellitus.* Diabetes Res Clin Pract, 2002. **55**(1): p. 65-85.
56. Looman, M., et al., *Development and evaluation of the Dutch Healthy Diet index 2015.* Public Health Nutr, 2017. **20**(13): p. 2289-2299.

Chapter 3

Dairy intake in relation to incident prediabetes and longitudinal insulin resistance in the Rotterdam Study

Manuscript based on this chapter:

Isabel A.L. Slurink, Trudy Voortman, Carolina Ochoa-Rosales, Fariba Ahmadizar, Maryam Kavousi, Nina Kupper, Tom Smeets, Sabita S. Soedamah-Muthu. *Dairy Product Consumption in Relation to Incident Prediabetes and Longitudinal Insulin Resistance in the Rotterdam Study.* Nutrients 2022, 14, 415.

Abstract

Background

Evidence suggests neutral or moderately beneficial effects of dairy intake on type 2 diabetes mellitus risk. Nevertheless, evidence on associations with early phases of type 2 diabetes remains inconsistent.

Aims

We aimed to examine associations between dairy-type intake with prediabetes risk and longitudinal insulin resistance.

Methods

The analytic sample consisted of 6,770 participants (aged 62 ± 4 years, 59% female) free of (pre-)diabetes at baseline from the prospective population-based Rotterdam Study. Dairy intake was measured at baseline using food frequency questionnaires. Data on prediabetes (fasting blood glucose 6.1-6.9 mmol/L or non-fasting 7.7-11.1 mmol/L) and the longitudinal homeostatic model assessment of insulin resistance (HOMA-IR) were available from 1993-2015. Associations with these outcomes were analysed with dairy intake in quartiles (Q4 versus Q1) and continuous using multivariable Cox proportional hazard models and linear mixed models.

Results

During a mean follow-up of 11.3 ± 4.8 years, 1139 incident prediabetes cases were documented (18.8%). In models adjusting for sociodemographic, lifestyle and dietary factors, a higher intake of high-fat yogurt was associated with lower prediabetes risk (HR_{Q4vsQ1} 0.70, 95%CI 0.54-0.91 and $HR_{serving/day}$ 0.67, 0.51-0.89). In addition, a higher intake of high-fat milk was associated with lower prediabetes risk (HR_{Q4vsQ1} 0.81, 0.67-0.97, HRserving/day 0.88, 0.79-0.99). Associations were found for low-fat dairy, low-fat milk, and total cheese with a higher prediabetes risk ($HR_{serving/day}$ ranging from 1.05-1.07, not significant in quartiles). Associations with longitudinal HOMA-IR were similar to prediabetes for high-fat yogurt, low-fat dairy and low-fat milk. Fermented dairy, low-fat yogurt, high-fat cheese, cream, and ice cream were not associated with the outcomes.

Conclusions

A higher intake of high-fat yogurt was associated with a lower prediabetes risk and lower longitudinal insulin resistance. Additionally, high-fat milk was associated with a lower prediabetes risk. Some low-fat dairy types were inconsistently associated with these outcomes. Studies are needed to confirm associations and to examine the influence of confounding by population characteristics.

Introduction

The diabetes mellitus epidemic is a global public health problem, making it crucial to identify preventive strategies effective in the early stages of the disease. In the early stages, progressive loss of β-cell capacity and mass result in the development of insulin resistance and elevation of fasting blood glucose [1]. Prediabetes is a widely prevalent condition, characterized by elevated blood glucose levels above the normal range, but below the diagnostic threshold for type 2 diabetes (T2D) [2]. Prediabetes is associated with cardiovascular disease and mortality [3, 4] and likely causally with coronary artery disease, as concluded by a Mendelian randomization study [3]. Thus, it is essential to identify modifiable risk factors that could prevent or revert insulin resistance and prediabetes or delay its transitions to T2D.

An unhealthy diet, physical inactivity and excess weight are of major importance in the development of T2D, and many cases can be prevented with lifestyle modifications [5-8]. Dairy products can be an interesting preventive target in maintaining cardiometabolic health, as they are a rich source of calcium, potassium, and vitamins. Furthermore, dairy proteins have been associated with favourable body composition and improved insulin sensitivity [9]. Vitamin K2 (menaquinones) may improve insulin sensitivity via several pathways [10] and has been associated with reduced T2D risk [11]. Nevertheless, dairy also contains sodium, saturated fatty acids (SFAs) and may contain added sugars. Dietary guidelines of many countries worldwide recommend consuming 1 to 4 servings of dairy foods daily, focusing on selecting low-fat options to lower SFA intake [12]. However, this recommendation is insufficiently substantiated as evidence on the harmful effects of high-fat dairy on cardiometabolic health is lacking [13, 14]. The health effects differ by individual SFA type, further modulated by the dairy food matrix [9]. Thus, there has been controversy concerning the place of dairy and its subtypes in healthy diets.

The overall evidence indicates a neutral or moderately beneficial association between dairy intake and T2D, especially yogurt [15-17]. Not much is known about dairy in relation to earlier phases of T2D, while this may provide further insights into its role in the aetiology of the disease. Only two prospective cohort studies have studied the associations of dairy and prediabetes risk. Data from The US Framingham Offspring Cohort with 12 years of follow-up (n = 1867) showed a 39%, 32% and 25% lower prediabetes risk for, respectively, total, low-fat, and high-fat dairy for the top versus bottom quartile and nonlinear inverse associations for milk and yogurt [18]. On the contrary, in our recent analysis in the Dutch Hoorn Studies (n = 2262; [19]), we did not observe associations for total dairy or most of the studied dairy types and prediabetes risk. However, high-fat fermented dairy, cheese and high-fat cheese were associated with a lower risk of prediabetes during a mean follow-up of 6.4 years. In the cross-sectional ELSA-Brasil study, inverse associations of dairy and insulin resistance were found [20]. A meta-analysis from Randomized Controlled Trials (RCTs) showed that beneficial effects of dairy on insulin resistance were more likely

to be observed in studies longer than 12 weeks [21]. Nevertheless, studies with long-term follow-up to confirm these results are lacking. Considerable heterogeneity between results, possibly explained by variation in dairy type intake or outcome definition, underline the need for extensive longitudinal studies.

Therefore, we examined associations between the consumption of total dairy and dairy subtypes with incident prediabetes and longitudinal insulin resistance in the prospective Rotterdam Study populations.

Materials and Methods

Study Population

This study was embedded in three sub-cohorts of the Rotterdam Study (RS), a prospective cohort study ongoing since 1990. It is comprised of middle-aged and elderly persons living in the district Ommoord in Rotterdam, the Netherlands. Details of the study design are described elsewhere [22]. The first sub-cohort (RS-I) was established in 1989-1993 among inhabitants aged 55 and over (*n* = 7,983). The second sub-cohort (RS-II) was recruited in January 2000 among people who had become 55 years of age or moved into the study district (*n* = 3,011). The third sub-cohort (RS-III) was initiated in 2006 for which subjects aged 45 years and older were recruited (*n* = 3,932). These three sub-cohorts of the Rotterdam Study comprised of 14,926 subjects at baseline with an overall response of 72%. Examinations were repeated every 3-5 years. The study was approved by the Medical Ethics Committee of Erasmus University Medical Centre, and all participants provided written informed consent.

For the current analysis, we excluded participants with prevalent T2D and without dietary data at baseline, resulting in 6,770 participants (RS-I-1: *n* = 2,971, RS-II- 1: *n* = 1,413, and RS-III-1: *n* = 2,386) (**Figure 1**). For the prediabetes incidence analysis, we additionally excluded participants with prediabetes at baseline or without follow-up data on prediabetes, resulting in 6,053 participants (RS-I: *n* = 2,617, RS-II: *n* = 1,250, RS-III: *n* = 2,186). For analyses on insulin resistance, we excluded participants without data on the homeostatic model assessment of insulin resistance (HOMA-IR) at baseline and follow-up, resulting in 6,593 participants (RS-I: *n* = 2,892, RS-II: *n* = 1,391, RS-III: *n* = 2,310). Data on outcome measures were available until 2015.

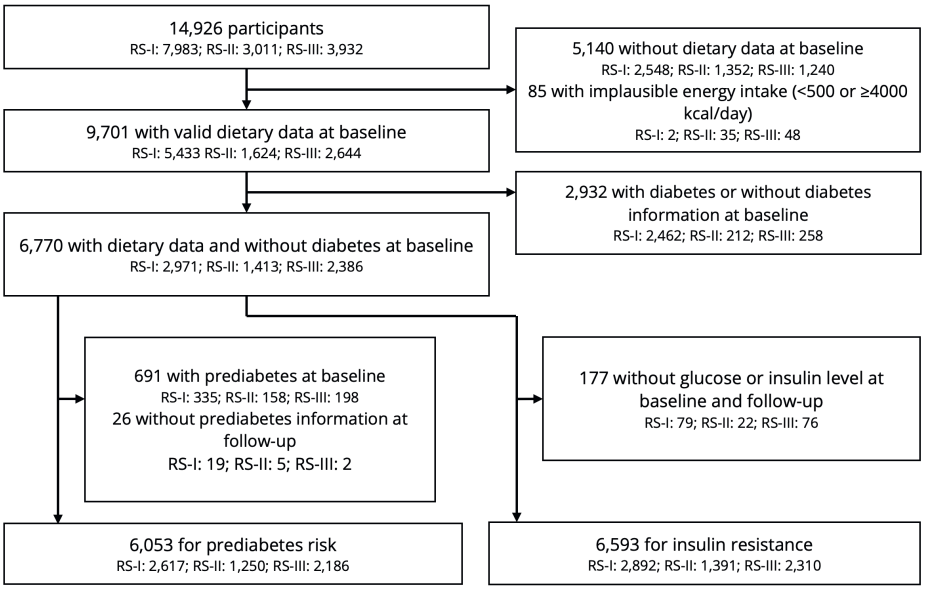


Figure 1. Flow-chart for inclusion of participants in the present analyses for prediabetes and longitudinal insulin resistance in the Rotterdam Study (RS).

Assessment of Outcomes

Fasting blood was drawn at two time points in each sub-cohort; at RS-I-3 (1997-1999) and I-5 (2009-2011), at RS-II-1 (2000-2001) and II-3 (2011-2012), and RS-III-1 (2006-2008) and III-2 (2012-2014). Glucose levels were measured using the glucose hexokinase method [23]. We set the third visit of RS-I (RS-I-3; 1997-1999) as a baseline, as fasting blood samples were not collected at the first two visits of RS-I. Information from general practitioners, structured home interviews, pharmacy dispensing records and follow-up examinations in the research facility was used to identify prediabetes and T2D cases at baseline and during follow-up. Prediabetes was defined as having a fasting blood glucose between 110 and 125 mg/dL (6.1 and 6.9 mmol/L) or non-fasting blood glucose between 138.6 and 199.8 mg/dL (7.7 and 11.1 mmol/L), according to WHO guidelines [2]. T2D was defined as a fasting plasma glucose level ≥ 126 mg/dL (7 mmol/L), a non-fasting plasma glucose level ≥199.8 mg/dL (11.1 mmol/L) or the use of blood glucose-lowering medication [2]. Two study physicians independently identified all cases. In case of a disagreement, a consensus was sought by consulting endocrinologists. Serum insulin levels were measured using the Roche Modular Analytics E170 analyzer. The HOMA-IR was calculated by multiplying fasting insulin (mU/L) by fasting glucose (mmol/L) divided by 22.5.

Methods

Assessment of Dairy Intake

Food intake data at baseline and follow-up were obtained with a 170-item (RS-I-1 and RS-II-1) and with a 389-item (RS-III-1, RS-I-5, and RS-II-3) food frequency questionnaire (FFQ) as described in detail elsewhere [24]. The FFQs were checked during an interview by a trained dietician at the study centre. The 170-item FFQ was validated in 80 RS participants against fifteen 24 h food records with adjusted Pearson's correlations of 0.66 for protein intake, 0.52 for saturated fat intake, 0.58 for sodium intake and against 24 h urine collections showing a Spearman correlation coefficient of 0.67 for protein [25]. The 389-item FFQ was validated in two Dutch populations against a 9-day dietary record and a 4-week dietary history with Pearson's correlation coefficients of 0.61 for total protein intake, 0.73-0.75 for saturated fat and 0.60 for calcium, and 0.60 for milk and milk products and 0.61 for cheese [26, 27]. Nutrient and energy intake were calculated using the Dutch Food Composition Tables 1993, 2001, 2006 and 2011 (NEVO) depending on the year of data collection in the sub-cohorts. Participants with an unreliable dietary intake according to the trained dietician or extreme energy intakes (<500 or >5000 kcal/day) were excluded [24]. Dairy categories included total dairy, fermented dairy, milk, yogurt, cheese, cream, and ice cream (**Table 1**). Each dairy category was further divided into low-fat (liquid products ≤ 2%, cheese ≤ 20%) and high-fat (liquid products > 2%, cheese > 20%). Intakes were expressed in servings/day according to Dutch serving sizes: milk, 200 mL; yogurt, 150 mL; cheese, 20 g; cream 3 g; ice cream, 50 g (https://portie-online.rivm.nl/, accessed on 1 September 2021). In the total dairy category, a serving of liquid dairy products was defined as 200 mL and cheese as 20 g.

Table 1. Food items included in total dairy and dairy types and consumption in the Rotterdam Studies (*n* = 6770). Grouping of dairy products included for the present analysis of the Rotterdam Studies.

Dairy type	Definition
Total dairy	All dairy products
High-fat[1]	All high-fat dairy products
Low-fat[1]	All low-fat dairy products
Fermented dairy	All fermented products
High-fat[2]	High-fat yogurt, high-fat cheese, full fat curd, full fat fruit curd
Low-fat	Low-fat yogurt, low-fat cheese, semi-skimmed curd, skimmed curd, semi-skimmed fruit curd, skimmed fruit curd
Milk, all types	All milk and milk products
High-fat	Full fat milk, full fat chocolate milk, milk powder, full fat milk added to the coffee
Low-fat	Semi-skimmed milk, skimmed milk, buttermilk, semi-skimmed chocolate milk, skimmed chocolate milk, semi-skimmed milk added to the coffee, skimmed milk added to the coffee

Table 1. Food items included in total dairy and dairy types and consumption in the Rotterdam Studies (*n* = 6770). Grouping of dairy products included for the present analysis of the Rotterdam Studies. (continued)

Dairy type	Definition
Yogurt	All yogurts
High-fat	Full fat yogurt, full fat fruit yogurt
Low-fat[3]	Semi-skimmed yogurt, skimmed yogurt, semi-skimmed fruit yogurt skimmed fruit yogurt
Cheese	All cheeses
High-fat	40+ cheese (e.g., Edam), 48+ cheese (e.g. Gouda, cheddar, cheese spread, goat cheese), full fat luxury cheese (e.g. cream brie, cream cheese, mon chou), cheese cubes, grated cheese, feta, cheese fondue
Low-fat	20+ and 30+ cheese (e.g., cheese spread, cottage cheese), low-fat luxury cheese (e.g. brie, goat cheese)
Cream	Whipped cream, coffee cream, semi-skimmed coffee cream, sour cream, crème fraiche, cooking cream
Ice cream	Ice cream

[1] High-fat dairy includes full fat custard for RS-I and RS-III. Low-fat dairy includes skimmed custard for RS-I.
[2] Includes mousse and chipolata pudding in RS-III.
[3] Includes skimmed custard, semi-skimmed curd, skimmed curd, semi-skimmed fruit curd and skimmed fruit curd in RS-III.
Abbreviations: RS, Rotterdam Study.

Assessment of Covariates

Information on demographic factors, education, health status, medical history and smoking behaviour was obtained during home interviews at baseline. Education attainment was defined as primary (primary education), low (lower/intermediate general education or lower vocational education), intermediate (intermediate vocational education or higher general education) or high (higher vocational education or university). Participants were classified as never, former, or current smokers. Height (cm) and weight (kg) were assessed during a physical examination at the research centre, and body mass index (BMI) was calculated as kg/m2. Waist circumference was assessed at baseline and during follow-up (RS-I-3 (1997-1999), RS-I-4 (2002-2004), RS-I-5 (2009-2011) and RS-I-6 (2014-2015); RS-II-2 (2004-2005), RS-II-3 (2011-2012) and RS-II-4 (2015-2016); and RS-III-2 (2012-2014)). Waist circumference in cm was measured at the level midway between the lower rib margin and the iliac crest with the participant in a standing position. Data on physical activity (PA) expressed in metabolic equivalent of task (MET) minutes per week were obtained using the Zutphen Physical Activity Questionnaire for RS-I-3 and RS-II-1 [28] and using the LASA Physical Activity Questionnaire (LAPAQ) for RS-III-1 [29, 30]. Diet quality was expressed as adherence to 14 food groups of the Dutch Dietary Guidelines 2015 [24, 31]. Information on medication use was obtained from both home interviews and pharmacy dispensing records. Blood pressure was calculated as the mean of two consecutive measurements with a random-zero sphygmomanometer while

subjects were in a sitting position and had rested for 5 min. Hypertension was defined as: systolic blood pressure ≥ 140 mmHg; and/or diastolic blood pressure ≥ 90 mmHg; and/or use of antihypertensive medication. Serum total cholesterol and HDL cholesterol were measured with the use of an automatic enzymatic procedure. Information on family history of diabetes was available at RS-I-1 and RS-II-1 and was defined as having at least one parent or sibling with T2D. Coronary heart disease (CHD) at baseline was defined as having a medical record of myocardial infarction and at follow-up as myocardial infarction or definite coronary mortality [32].

Statistical Analysis

Descriptive data were presented as means and standard deviations (SD) for continuous variables, medians, and interquartile ranges (IQR) for non-normally distributed continuous variables and frequencies and percentages for categorical variables. Natural log-transformed values for HOMA-IR were used to approximate normal data distributions. To analyse associations between the various dairy types and prediabetes incidence, we used Cox proportional hazard models. Results were expressed as a Hazard Ratio and 95% confidence intervals (HR, 95%CI). To analyse associations between dairy types and longitudinal HOMA-IR, we performed linear mixed models with time as a fixed effect, a random intercept for participants and a random slope for the time of repeated HOMA-IR measures. Results were expressed as beta coefficients of log-transformed HOMA-IR and 95% confidence intervals (β, 95%CI), representing the overall effect of dairy consumption averaged across follow-up. Dairy types were analysed as categorical variables based on quartiles, comparing the highest versus lowest (reference) quartile of intake, and continuous variables (servings/day). Dairy types for which many participants reported no intake were categorized into a non-consumer category (reference) and consumers into tertiles. The *p* for the trend was calculated using the median values of dairy intake range categories as continuous variables in the model. For each model, we examined whether non-linear terms of continuous dairy types (2nd order polynomials or natural splines with 3 degrees of freedom, excluding outliers) significantly improved model fit compared to the linear model assessed by likelihood ratio tests.

Confounders were chosen based on previous research [31, 33, 34]. The basic model (model 1) was adjusted for age, sex, and daily energy intake. Model 2 was additionally adjusted for educational attainment, alcohol intake, smoking status, physical activity, family history of T2D (RS-I and RS-II only), intake of fruits, vegetables, whole grains, legumes, nuts, tea, coffee, red meat, and sugar-sweetened beverages (SSBs). We presented descriptive data stratified by the dairy types significantly associated with the outcomes to provide insight in characteristics of participants with high and low intakes. Effect modification by age, sex and waist circumference were examined in model 2, and stratified associations were presented in case of significant interactions ($p < 0.05$).

Multiple sensitivity analyses were performed in model 2 to examine the robustness of the findings. First, we additionally adjusted for longitudinal waist circumference to examine the potential confounding or mediating effect of obesity over time in associations of dairy, prediabetes incidence and longitudinal HOMA-IR. To adjust for longitudinal waist circumference in models of prediabetes incidence, we applied a joint modelling approach [35]. With this approach, model 2 was combined with a random slope linear mixed model, including the repeated measures of waist circumference before the onset of prediabetes as an outcome, time of waist circumference measurements and interactions between dairy types and time of waist circumference measurements. Second, we additionally adjusted for cholesterol, hypertension, and triglycerides, as these factors are potential mediators. Third, we additionally adjusted for consumption of other dairy types to assess whether associations of certain dairy types were independent of each other. Fourth, participants with prevalent or incident CHD were excluded to address reverse causation by the change of diet and lifestyle. Fifth, associations were calculated with energy-adjusted intake of dairy types in gram/day using the residual method. Sixth, repeated measures of dairy intake (baseline and measures 20 years after baseline in RS-I (*n* = 1,028, 34.6%) and 10 years after baseline in RS-II (*n* = 859, 60.8%)) were included in Cox models as time-dependent exposure and in linear mixed models as a fixed effect. Note that for most participants, the Cox model used baseline measures of dairy only because most prediabetes cases occurred before the repeated dietary assessment. Therefore, we additionally adjusted for dairy intake at follow-up in the subset with these data available to explore potential effects of altered dairy intake.

All analyses were performed separately for RS-I, RS-II and RS-III, and the results were pooled using a fixed-effects meta-analysis. To adjust for potential bias associated with missing data, a multiple imputation procedure (*n* = 10) was used to account for missing data on covariates (**Supplemental table 1**). No correction for multiple testing was made, as most exposures were correlated, and corrections may have resulted in a type II error [36]. Statistical procedures were performed using SPSS statistical software, version 21.0 (IBM Corp, Armonk, NY, USA) and R version 4.0.2. (The R Foundation for Statistical Computing, Vienna, Austria).

Results

Population Characteristics

The mean total dairy intake was 3.6 ± 1.2 servings/day, mostly consisting of low-fat milk (0.9 ± 0.6) and high-fat cheese consumption (1.5 ± 0.8) (**Figure 2**). The mean age was 61.7 ± 3.9 years, and 58.7% were female (**Table 2**). The mean waist circumference was 91.1 ± 6.7 cm, the mean BMI was 26.6 ± 2.2 kg/m2 and 16% were obese (BMI ≥ 30 kg/m2). In the highest (6.0 ± 1.1 servings/day) compared to the lowest quartile of dairy intake (1.5 ± 0.3 servings/day), participants were more often highly educated (21.4 versus 16.1%) and less often smokers (22.4 versus 27.0%). Furthermore, diet quality, energy intake

and intakes of vegetables, fruit, whole grains, sodium, and calcium were on average higher with increasing dairy intake. Characteristics by dairy consumption, sub-cohort and included versus excluded from the current analyses are presented in **Supplemental Table 2**. Stratified by cohort, RS-III was on average younger (56.8 ± 6.4, 65.5 ± 6.7, 63.6 ± 7.2 years in RS-III; -I; -II, respectively), with more highly educated participants (28.7% versus 10.1 and 18.1%). Participants included in this study (*n* = 6,770) compared to those who were excluded (*n* = 8162) were generally younger (62.0 ± 7.8 versus 69.2 ± 11.4), higher educated (18.3 versus 11.6%), less often smokers (22.7 versus 25.6%), had higher physical activity levels and lower HOMA-IR levels (2.9 ± 2.4 versus 5.6 ± 13.0).

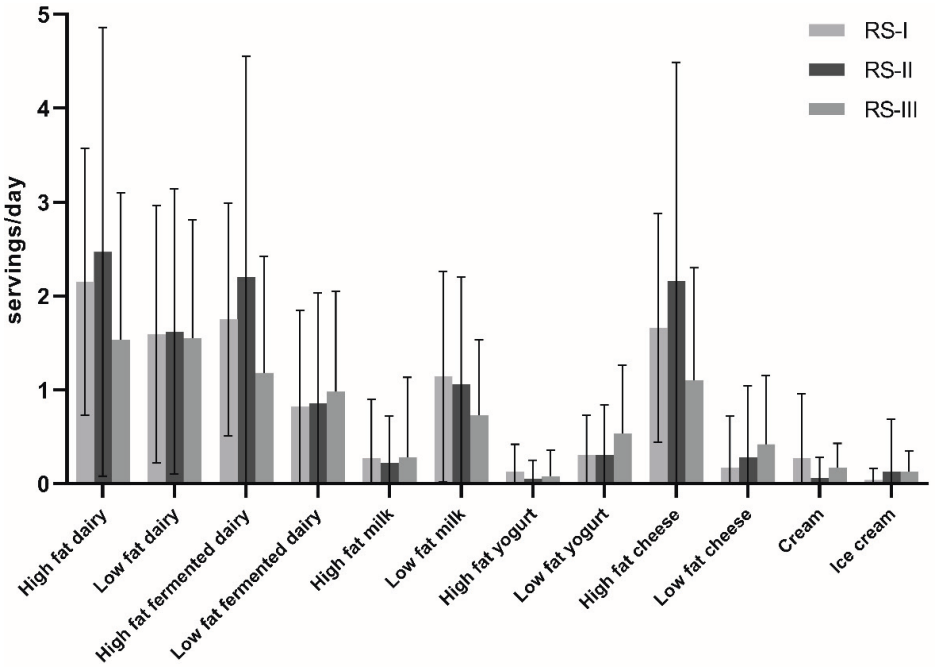


Figure 2. Dairy type intake by cohort of the Rotterdam Studies in servings/day (mean ± SD): milk, 200 mL; yogurt, 150 mL; cheese, 20 g; cream 3 g; ice cream, 50 g. Combined total dairy category: liquid dairy products, 200 mL; cheese, 20 g. Abbreviations: RS, Rotterdam Study.

Table 2. Baseline characteristics of participants in the Rotterdam Study in the total population and across quartiles (Q) of total dairy intake (*n* = 6770).

	Total	**Q1**	**Q2**	**Q3**	**Q4**
	n = 6770	*n* = 1692	*n* = 1697	*n* = 1688	*n* = 1693
Total dairy intake	3.6 ± 1.2	1.5 ± 0.3	2.8 ± 0.2	3.9 ± 0.2	6.0 ± 1.1
Range	0–15.1	0–1.2	0.9–1.9	1.4–2.7	2.2–15.1
Age at dietary assessment	61.7 ± 3.9	62.1 ± 4.0	62.0 ± 3.9	61.8 ± 3.8	60.9 ± 3.7
Sex, female (%)	58.7	59.9	61.7	59.6	53.8
BMI (kg/m2)	26.6 ± 2.2	26.7 ± 2.2	26.7 ± 2.2	26.6 ± 2.2	26.6 ± 2.2
Waist circumference (cm)	91.1 ± 6.7	91.1 ± 6.7	90.7 ± 6.6	91.0 ± 6.5	91.6 ± 6.8
Education level (%)					
Primary education	11.8	13.7	11.2	11.4	11.0
Lower education	40.9	42.2	43.8	40.1	37.6
Intermediate	28.9	28.0	28.0	29.7	30.0
Higher	18.3	16.1	17.0	18.8	21.4
Smoking (%)					
Never	32.2	31.3	33.5	32.9	31.3
Ever	45.0	41.7	45.9	46.3	46.3
Current	22.7	27.0	20.6	20.8	22.4
Physical activity (MET-hours/week)					
Zutphen Physical Activity Questionnaire (*n* = 4,328)	79.7 [54.7, 112.1]	77.7 [51.5, 110.8]	78.2 [53.0, 109.2]	79.4 [57.5, 111.9]	82.4 [56.6, 115.6]
LASA Physical Activity Questionnaire (*n* = 2,177)	42.9 [17.7, 82.5]	39.8 [15.0, 75.9]	40.5 [16.9, 79.9]	48.2 [21.0, 87.8]	45.0 [18.0, 86.4]
Family history diabetes	12.8	12.7	12.6	12.6	13.3
Dietary intake					
Energy (kcal/day)	2113 ± 333	1858 ± 293	2012 ± 283	2151 ± 285	2452 ± 365
Diet quality score (0–14)	6.6 ± 1.1	6.0 ± 1.0	6.6 ± 1.0	7.0 ± 1.1	7.0 ± 1.1
Total fat (E%)	35.1 ± 3.6	35.3 ± 4.1	35.0 ± 3.5	34.5 ± 3.3	35.4 ± 3.6
Total saturated fat (E%)	13.2 ± 1.6	12.4 ± 1.6	12.9 ± 1.5	13.1 ± 1.4	14.1 ± 1.8
Total protein (E%)	16.7 ± 1.7	15.8 ± 1.6	16.4 ± 1.6	16.8 ± 1.5	17.6 ± 1.7
Carbohydrate (E%)	44.5 ± 4.2	44.5 ± 4.9	44.9 ± 4.1	44.9 ± 3.8	43.8 ± 3.9
Calcium (mg/day)	1109 ± 251	688 ± 113	960 ± 95	1175 ± 101	1621 ± 245
Sodium (mg/day)	2344 ± 463	1979 ± 385	2203 ± 366	2398 ± 372	2814 ± 511
Alcohol (g/day)	6.6 [0.7, 18.8]	6.7 [0.5, 20.9]	6.6 [0.7, 18.6]	6.2 [0.7, 17.6]	6.7 [0.7, 17.6]
Vegetables (g/day)	211 ± 69	206 ± 62	208 ± 71	207 ± 59	220 ± 74
Fruit (g/day)	228 ± 98	209 ± 102	230 ± 93	235 ± 94	234 ± 101
Wholegrains (g/day)	116 ± 43	95 ± 41	111 ± 41	123 ± 42	133 ± 46
Legumes (g/day)	16.5 ± 12.5	15.7 ± 14.6	16.4 ± 11.2	15.4 ± 9.9	17.5 ± 11.8
Nuts (g/day)	8.5 ± 7.9	7.9 ± 8.0	8.3 ± 7.9	8.5 ± 7.3	9.1 ± 8.1
Red meat (g/day)	93 ± 36	91 ± 35	92 ± 34	93 ± 32	97 ± 40
Fish (g/day)	20 ± 13	19 ± 13	21 ± 12	20 ± 13	21 ± 13

Table 2. Baseline characteristics of participants in the Rotterdam Study in the total population and across quartiles (Q) of total dairy intake (*n* = 6770). (continued)

	Total	Q1	Q2	Q3	Q4
	n = 6770	*n* = 1692	*n* = 1697	*n* = 1688	*n* = 1693
Tea (g/day)	288 ± 155	286 ± 162	275 ± 146	304 ± 152	286 ± 159
Coffee (g/day)	471 ± 152	445 ± 155	462 ± 144	475 ± 146	502 ± 159
SSBs (g/day)	94 ± 74	92 ± 78	96 ± 74	86 ± 63	101 ± 79

Values are mean ± SD for continuous variables with a normal distribution (pooled), or median [IQR] for continuous variables with a skewed distribution; percentages for categorical variables, based on unimputed data. Abbreviations: E%, percentage of total energy intake; MET, metabolic equivalent of task; RS, Rotterdam Study; SD; standard deviation, SSBs, Sugar Sweetened Beverages.

Dairy Intake and Prediabetes Risk

During a mean follow-up of 11.4 ± 4.8 years, 1139 incident prediabetes cases were identified among 6,053 participants (18.8%). In pooled multivariable models (**Table 3**, Model 2), high-fat yogurt (19% of the sum of total yogurt intake) was associated with a lower prediabetes risk (HR_{Q4vsQ1} 0.70, 95%CI 0.54-0.91 and $HR_{serving/day}$ 0.67, 95%CI 0.51-0.89). Additionally, high-fat milk (21% of total milk) was associated with a lower prediabetes risk (HR_{Q4vsQ1} 0.81, 0.67-0.97 and $HR_{serving/day}$ 0.88, 95%CI 0.79-0.99). In contrast, low-fat dairy, low-fat milk, and total cheese were associated with a higher prediabetes risk when analysed on a continuous scale and not in quartiles ($HR_{serving/day}$, respectively, 1.05, 1.01-1.10; 1.07, 1.01-1.13; 1.05, 1.01-1.09). In addition, low-fat cheese was associated with a higher prediabetes risk when analysed in quartiles (15% of total cheese, HR_{Q4vsQ1} 1.17, 95%CI 0.95-1.44, P_{trend} = 0.04). Total and high-fat dairy; total, high-fat, and low-fat fermented dairy; total milk; total and low-fat yogurt; high-fat cheese; cream and ice cream were not associated with a prediabetes risk.

Dairy Intake and Longitudinal Insulin Resistance

The median HOMA-IR index was 2.3 (IQR 1.7-3.4) at baseline and 2.4 (1.7-3.8) at follow-up. In line with results for prediabetes, high-fat yogurt was associated with lower longitudinal log-transformed HOMA-IR (β_{Q4vsQ1} -0.10, 95%CI -0.16; -0.05, P_{trend} = 0.0003, $\beta_{serving/day}$ -0.08, 95%CI -0.13; -0.03) (**Table 4**, Model 2). Low-fat dairy and low-fat milk were associated with higher longitudinal log-HOMA-IR (β_{Q4vsQ1}, respectively, 0.06, 95%CI 0.03-0.10, P_{trend} = 0.0003 and 0.07, 0.03-0.11, P_{trend} = 0.001, and $\beta_{serving/day}$, respectively, 0.02, 0.01-0.03 and 0.02, 0.01-0.04). Total milk was significantly associated with higher longitudinal log-HOMA-IR only when comparing top versus bottom quartiles (β_{Q4vsQ1} 0.05, 0.01-0.09, P_{trend} = 0.02). In contrast to observations for prediabetes risk, total dairy, high-fat milk, and total and low-fat cheese were not associated with longitudinal HOMA-IR. A better fit of non-linear associations was found for low-fat fermented dairy, low-fat milk, high-fat milk, low-fat yogurt, and cream, but non-linear trends were inconclusive, as only a few participants had high intakes **(Supplemental Figure 1)**.

Table 3. Associations of dairy product types and prediabetes risk in the Rotterdam Studies (*n* = 6,053).

	Pooled hazard ratio (95%CI) across intake range categories[1]					Continuous[2]
	Q1	Q2	Q3	Q4	P_{trend}	HR (95% CI)
Total dairy						
n cases/n total	297/1512	287/1517	257/1513	298/1511		1139/6053
Median, servings/d	1.6	2.8	3.9	5.6		3.3
Model 1	1 (ref)	0.96 (0.81–1.13)	0.85 (0.71–1.00)	1.00 (0.84–1.19)	0.89	1.02 (0.99–1.05)
Model 2	1 (ref)	0.98 (0.83–1.16)	0.91 (0.76–1.08)	1.09 (0.91–1.31)	0.38	1.03 (1.00–1.07)
High-fat dairy						
n cases/n total	285/1512	302/1502	288/1532	264/1507		
Median, servings/d	0.4	1.4	2.1	3.6		1.7
Model 1	1 (ref)	1.11 (0.94–1.30)	0.98 (0.83–1.16)	0.91 (0.76–1.09)	0.11	1.00 (0.96–1.04)
Model 2	1 (ref)	1.11 (0.94–1.31)	0.97 (0.82–1.15)	0.94 (0.78–1.13)	0.22	1.00 (0.96–1.04)
Low-fat dairy						
n cases/n total	273/1517	288/1511	283/1510	295/1515		
Median, servings/d	0.1	1.0	1.8	3.1		1.4
Model 1	1 (ref)	1.07 (0.90–1.26)	1.06 (0.90–1.25)	1.09 (0.92–1.28)	0.31	1.03 (0.99–1.08)
Model 2	1 (ref)	1.08 (0.92–1.28)	1.10 (0.93–1.31)	1.17 (0.99–1.39)	0.06	1.05 (1.01–1.10) *
Fermented dairy						
n cases/n total	292/1525	284/1501	279/1514	284/1513		
Median, servings/d	1.0	1.8	2.7	4.2		2.2
Model 1	1 (ref)	0.96 (0.82–1.14)	0.93 (0.79–1.10)	0.94 (0.79–1.11)	0.48	1.02 (0.98–1.05)
Model 2	1 (ref)	0.98 (0.83–1.15)	0.95 (0.80–1.13)	1.00 (0.84–1.19)	0.94	1.03 (0.99–1.06)
High-fat fermented dairy						
n cases/n total	294/1522	279/1498	285/1520	281/1513		
Median, servings/d	0.2	1.1	1.8	3.0		1.3
Model 1	1 (ref)	0.99 (0.84–1.17)	0.95 (0.81–1.12)	0.94 (0.79–1.11)	0.44	1.03 (0.99–1.07)
Model 2	1 (ref)	0.99 (0.83–1.16)	0.94 (0.80–1.11)	0.93 (0.78–1.11)	0.41	1.03 (0.99–1.07)

Table 3. Associations of dairy product types and prediabetes risk in the Rotterdam Studies (*n* = 6,053) (continued).

	Pooled hazard ratio (95%CI) across intake range categories[1]					Continuous[2]
	Q1	Q2	Q3	Q4	P_{trend}	HR (95% CI)
Low-fat fermented dairy						
n cases/n total	309/1571	289/1523	258/1463	283/1496		
Median, servings/d	0.0	0.3	0.9	2.1		0.6
Model 1	1 (ref)	0.96 (0.81–1.12)	0.88 (0.75–1.04)	0.93 (0.79–1.10)	0.50	0.99 (0.94–1.05)
Model 2	1 (ref)	0.98 (0.83–1.16)	0.93 (0.79–1.11)	1.01 (0.85–1.19)	0.85	1.01 (0.96–1.07)
Total milk						
n cases/n total	276/1509	299/1499	271/1437	293/1608		
Median, servings/d	0.1	0.8	1.3	2.4		1.0
Model 1	1 (ref)	1.12 (0.95–1.31)	1.05 (0.89–1.25)	1.03 (0.87–1.22)	0.73	1.00 (0.95–1.06)
Model 2	1 (ref)	1.13 (0.95–1.33)	1.09 (0.92–1.30)	1.09 (0.92–1.29)	0.31	1.02 (0.97–1.08)
High-fat milk						
n cases/n total	600/3002	148/860	226/1141	165/1050		
Median, servings/d	0.0	0.1	0.2	1.0		0.0
Model 1	1 (ref)	0.90 (0.75–1.08)	1.01 (0.87–1.18)	0.79 (0.66–0.94)	0.02	0.87 (0.78–0.97) *
Model 2	1 (ref)	0.94 (0.78–1.13)	1.03 (0.88–1.21)	0.81 (0.67–0.97)	0.04	0.88 (0.79–0.99) *
Low-fat milk						
n cases/n total	274/1583	297/1488	294/1490	274/1492		
Median, servings/d	0.0	0.5	1.1	2.2		0.8
Model 1	1 (ref)	1.19 (1.01–1.40)	1.19 (1.00–1.40)	1.09 (0.92–1.29)	0.45	1.05 (0.99–1.11)
Model 2	1 (ref)	1.19 (1.01–1.41)	1.20 (1.02–1.43)	1.14 (0.96–1.36)	0.20	1.07 (1.01–1.13) *
Total yogurt						
n cases/n total	425/2050	279/1480	221/1231	214/1292		
Median, servings/d	0.0	0.3	0.6	1.0		0.4
Model 1	1 (ref)	0.89 (0.77–1.04)	0.93 (0.78–1.10)	0.80 (0.67–0.94)	0.00	0.88 (0.78–0.98) *
Model 2	1 (ref)	0.92 (0.79–1.08)	1.00 (0.84–1.19)	0.84 (0.71–0.99)	0.05	0.92 (0.82–1.02)

Table 3. Associations of dairy product types and prediabetes risk in the Rotterdam Studies (*n* = 6,053) (continued).

	Pooled hazard ratio (95%CI) across intake range categories[1]					Continuous[2]
	Q1	Q2	Q3	Q4	P_{trend}	HR (95% CI)
High-fat yogurt						
n cases/n total	908/4596	69/514	100/500	62/443		
Median, servings/d	0.0	0.1	0.3	0.7		0.0
Model 1	1 (ref)	0.69 (0.54–0.88)	1.04 (0.84–1.28)	0.68 (0.52–0.89)	0.003	0.66 (0.50–0.88) **
Model 2	1 (ref)	0.70 (0.54–0.89)	1.04 (0.84–1.28)	0.70 (0.54–0.91)	0.005	0.67 (0.51–0.89) **
Low-fat yogurt						
n cases/n total	457/2354	267/1377	205/1070	210/1252		
Median, servings/d	0.0	0.1	0.5	1.0		0.1
Model 1	1 (ref)	1.06 (0.91–1.24)	1.04 (0.87–1.23)	0.94 (0.79–1.11)	0.17	0.94 (0.84–1.06)
Model 2	1 (ref)	1.10 (0.93–1.28)	1.10 (0.92–1.31)	0.99 (0.83–1.17)	0.54	0.99 (0.88–1.11)
Total cheese						
n cases/n total	272/1501	287/1517	270/1501	310/1534		
Median, servings/d	0.5	1.2	2.0	3.1		1.5
Model 1	1 (ref)	1.04 (0.88–1.23)	0.97 (0.82–1.15)	1.09 (0.92–1.29)	0.46	1.05 (1.01–1.08) *
Model 2	1 (ref)	1.04 (0.88–1.23)	0.98 (0.83–1.17)	1.11 (0.94–1.33)	0.32	1.05 (1.01–1.09) *
High-fat cheese						
n cases/n total	281/1536	282/1502	286/1503	290/1512		
Median, servings/d	0.1	1.1	1.7	2.9		1.2
Model 1	1 (ref)	1.06 (0.90–1.25)	1.05 (0.89–1.24)	1.04 (0.88–1.24)	0.74	1.04 (1.00–1.07) *
Model 2	1 (ref)	1.05 (0.89–1.24)	1.03 (0.87–1.21)	1.05 (0.88–1.25)	0.75	1.03 (1.00–1.08)
Low-fat cheese						
n cases/n total	813/4260	111/625	104/560	111/608		
Median, servings/d	0.0	0.2	0.7	1.8		0.0
Model 1	1 (ref)	1.06 (0.86–1.30)	1.13 (0.91–1.41)	1.09 (0.89–1.35)	0.10	1.05 (0.97–1.14)
Model 2	1 (ref)	1.10 (0.90–1.36)	1.16 (0.93–1.45)	1.17 (0.95–1.44)	0.04	1.06 (0.97–1.14)

Table 3. Associations of dairy product types and prediabetes risk in the Rotterdam Studies (*n* = 6,053) (continued).

	Pooled hazard ratio (95%CI) across intake range categories[1]					Continuous[2]
	Q1	Q2	Q3	Q4	P_{trend}	HR (95% CI)
Cream						
n cases/n total	735/3604	140/825	128/849	136/775		
Median, servings/d	0.0	0.07	0.18	0.49		0.0
Model 1	1 (ref)	0.89 (0.74–1.09)	0.88 (0.71–1.07)	1.01 (0.83–1.23)	0.41	1.04 (0.93–1.17)
Model 2	1 (ref)	0.90 (0.74–1.10)	0.89 (0.72–1.09)	1.00 (0.82–1.22)	0.52	1.03 (0.92–1.16)
Ice cream						
n cases/n total	713/3615	137/792	122/765	167/881		
Median, servings/d	0.0	0.07	0.17	0.28		0.0
Model 1	1 (ref)	0.94 (0.78–1.13)	0.87 (0.71–1.07)	0.95 (0.80–1.13)	0.58	0.94 (0.71–1.26)
Model 2	1 (ref)	0.93 (0.77–1.12)	0.86 (0.70–1.05)	0.93 (0.78–1.11)	0.50	0.94 (0.70–1.26)

[1] Hazard ratios (95CIs) were estimated across four categories split by quartile values (Q1 to Q4) or non-consumers + tertile categories with the lowest category as the reference, adjusted for covariates as follows: Model 1 included age (continuous), sex and energy intake (continuous). Model 2 was additionally adjusted for education (3 categories), smoking (3 categories), physical activity (continuous), alcohol consumption (4 categories), family history of diabetes (yes/no, RS-I and RS-II only) and food groups associated with type 2 diabetes, including intakes of fruit, vegetables, wholegrains, legumes, nuts, tea, coffee, red meat, and sugar-sweetened beverages (SSB) (continuous). Linear trend across intake range categories was assessed by including median values of each category as a continuous variable in the model.
[2] Hazard ratios per 1 serving/day (see definition in Table 1) were estimated .* P=0.01 to 0.05, ** P<0.01.
Abbreviations: CI, Confidence Interval; Q, Quartile.

Table 4. Associations of dairy product types and longitudinal insulin resistance in the Rotterdam Studies (*n* = 6,593).

	Pooled betas from log-transformed HOMA-IR (95%CI) across intake range categories[1]					Continuous[2]
	Q1	Q2	Q3	Q4	P_{trend}	B (95% CI)
Total dairy						
N total	1650	1642	1647	1654		
Median[3]	1.6	2.8	3.9	5.6		3.3
Model 1	ref	0.01 (-0.03, 0.05)	-0.02 (-0.05, 0.02)	0.02 (-0.02, 0.06)	0.37	0.00 (-0.01, 0.01)
Model 2	ref	0.02 (-0.02, 0.05)	0.00 (-0.04, 0.04)	0.04 (0.00, 0.08)	0.07	0.00 (0.00, 0.01)
High-fat dairy						
N total	1644	1663	1642	1644		
Median[3]	0.4	1.4	2.2	3.6		1.7
Model 1	ref	0.02 (-0.02, 0.06)	-0.03 (-0.07, 0.00)	-0.03 (-0.08, 0.01)	0.03*	-0.01 (-0.01, 0.00)
Model 2	ref	0.02 (-0.02, 0.06)	-0.03 (-0.07, 0.01)	-0.03 (-0.07, 0.01)	0.06	-0.01 (-0.02, 0.00)
Low-fat dairy						
N total	1652	1655	1655	1631		
Median[3]	0.1	1	1.7	3		1.3
Model 1	ref	0.02 (-0.02, 0.06)	0.05 (0.02, 0.09)	0.05 (0.01, 0.09)	0.01**	0.01 (0.00, 0.02)*
Model 2	ref	0.02 (-0.01, 0.06)	0.06 (0.03, 0.10)	0.06 (0.03, 0.10)	0.0003***	0.02 (0.01, 0.03)**
Fermented dairy						
N total	1653	1645	1644	1651		
Median[3]	1	1.8	2.7	4.2		2.2
Model 1	ref	-0.03 (-0.07, 0.01)	-0.02 (-0.06, 0.02)	-0.04 (-0.08, 0.00)	0.10	-0.01 (-0.02, 0.00)
Model 2	ref	-0.02 (-0.06, 0.02)	0.00 (-0.04, 0.04)	-0.01 (-0.05, 0.03)	0.77	0.00 (-0.01, 0.00)
High-fat fermented dairy						
N total	1658	1625	1648	1662		
Median[3]	0.2	1.1	1.8	3		1.4
Model 1	ref	0.02 (-0.02, 0.06)	-0.02 (-0.06, 0.02)	-0.02 (-0.06, 0.02)	0.22	0.00 (-0.01, 0.01)
Model 2	ref	0.02 (-0.02, 0.06)	-0.01 (-0.05, 0.02)	-0.01 (-0.05, 0.03)	0.49	0.00 (-0.01, 0.01)

Table 4. Associations of dairy product types and longitudinal insulin resistance in the Rotterdam Studies (*n* = 6,593). (continued)

	Pooled betas from log-transformed HOMA-IR (95%CI) across intake range categories[1]					Continuous[2]
	Q1	Q2	Q3	Q4	P_{trend}	B (95% CI)
Low-fat fermented dairy						
N total	1718	1662	1624	1589		
Median[3]	0	0.3	0.9	2		0.5
Model 1	ref	0.04 (0.00, 0.08)	0.03 (-0.01, 0.06)	0.00 (-0.04, 0.04)	0.31	-0.01 (-0.02, 0.00)
Model 2	ref	0.05 (0.01, 0.08)	0.04 (0.00, 0.08)	0.02 (-0.01, 0.06)	0.70	0.00 (-0.01, 0.01)
Total milk						
N total	1633	1627	1752	1581		
Median[3]	0.1	0.8	1.3	2.4		1.0
Model 1	ref	0.03 (-0.01, 0.07)	0.03 (-0.01, 0.07)	0.05 (0.01, 0.09)	0.02*	0.01 (0.00, 0.02)
Model 2	ref	0.03 (-0.01, 0.07)	0.03 (0.00, 0.07)	0.05 (0.01, 0.09)	0.02*	0.01 (0.00, 0.02)
High-fat milk						
N total	3278	942	1146	1227		
Median[3]	0	0.05	0.2	0.9		0.003
Model 1	ref	0.01 (-0.03, 0.05)	-0.02 (-0.05, 0.02)	0.00 (-0.04, 0.04)	0.46	-0.02 (-0.04, 0.00)
Model 2	ref	0.00 (-0.04, 0.04)	-0.03 (-0.06, 0.01)	0.00 (-0.04, 0.03)	0.46	-0.02 (-0.04, 0.00)
Low-fat milk						
N total	1736	1604	1620	1633		
Median[3]	0	0.5	1.1	2.2		0.8
Model 1	ref	0.03 (-0.01, 0.06)	0.03 (-0.01, 0.06)	0.07 (0.03, 0.11)	0.002**	0.02 (0.01, 0.04) ***
Model 2	ref	0.02 (-0.02, 0.06)	0.02 (-0.02, 0.06)	0.07 (0.03, 0.11)	0.001**	0.02 (0.01, 0.04) ***
Total yogurt						
N total	2259	1606	1402	1326		
Median[3]	0	0.3	0.6	1.0		0.4
Model 1	ref	0.04 (0.00, 0.07)	-0.01 (-0.05, 0.03)	-0.03 (-0.07, 0.00)	0.02*	-0.03 (-0.05, -0.01)**
Model 2	ref	0.04 (0.01, 0.08)	0.00 (-0.04, 0.04)	-0.01 (-0.05, 0.02)	0.18	-0.02 (-0.04, 0.00)

Table 4. Associations of dairy product types and longitudinal insulin resistance in the Rotterdam Studies (*n* = 6,593). (continued)

	Pooled betas from log-transformed HOMA-IR (95%CI) across intake range categories[1]					Continuous[2]
	Q1	Q2	Q3	Q4	P_{trend}	B (95% CI)
High-fat yogurt						
N total	5021	540	470	562		
Median[3]	0	0.1	0.3	0.7		0.0
Model 1	ref	-0.01 (-0.06, 0.03)	0.00 (-0.05, 0.05)	-0.11 (-0.17, -0.06)	0.0001***	-0.08 (-0.13, -0.03)**
Model 2	ref	-0.01 (-0.06, 0.04)	0.00 (-0.05, 0.05)	-0.10 (-0.16, -0.05)	0.0003***	-0.08 (-0.13, -0.03)**
Low-fat yogurt						
N total	2593	1412	1366	1222		
Median[3]	0	0.1	0.5	1.0		0.14
Model 1	ref	0.03 (-0.01, 0.06)	0.05 (0.02, 0.09)	0.00 (-0.04, 0.03)	0.71	-0.02 (-0.04, 0.01)
Model 2	ref	0.03 (0.00, 0.07)	0.07 (0.03, 0.10)	0.01 (-0.02, 0.05)	0.52	0.00 (-0.03, 0.02)
Total cheese						
N total	1650	1633	1670	1640		
Median[3]	0.5	1.2	2.0	3.1		1.5
Model 1	ref	0.01 (-0.03, 0.05)	-0.01 (-0.05, 0.02)	-0.01 (-0.05, 0.03)	0.62	0.00 (-0.01, 0.01)
Model 2	ref	0.02 (-0.02, 0.06)	0.00 (-0.04, 0.04)	0.01 (-0.03, 0.05)	0.62	0.00 (-0.01, 0.01)
High-fat cheese						
N total	1691	1607	1652	1643		
Median[3]	0.2	1.1	1.7	3.0		1.3
Model 1	ref	0.01 (-0.02, 0.05)	-0.01 (-0.04, 0.03)	-0.01 (-0.05, 0.03)	0.52	0.00 (-0.01, 0.01)
Model 2	ref	0.01 (-0.02, 0.05)	0.00 (-0.04, 0.04)	0.00 (-0.04, 0.04)	0.92	0.00 (-0.01, 0.01)
Low-fat cheese						
N total	4682	664	647	600		
Median[3]	0	0.2	0.7	1.8		0.0
Model 1	ref	0.02 (-0.03, 0.07)	0.00 (-0.05, 0.05)	-0.01 (-0.05, 0.04)	0.83	0.00 (-0.02, 0.02)
Model 2	ref	0.02 (-0.02, 0.07)	0.01 (-0.04, 0.06)	0.02 (-0.03, 0.07)	0.50	0.01 (-0.01, 0.03)

Table 4. Associations of dairy product types and longitudinal insulin resistance in the Rotterdam Studies (*n* = 6,593). (continued)

	Pooled betas from log-transformed HOMA-IR (95%CI) across intake range categories[1]					Continuous[2]
	Q1	Q2	Q3	Q4	P_{trend}	B (95% CI)
Cream						
N total	3977	880	829	907		
Median[3]	0	0.07	0.2	0.5		0.0
Model 1	ref	-0.03 (-0.07, 0.02)	-0.06 (-0.10, -0.01)	-0.07 (-0.11, -0.02)	0.89	-0.01 (-0.04, 0.01)
Model 2	ref	-0.03 (-0.07, 0.02)	-0.05 (-0.09, 0.00)	-0.07 (-0.12, -0.03)	0.77	-0.02 (-0.05, 0.01)
Ice cream						
N total	3963	851	956	823		
Median[3]	0	0.07	0.2	0.3		0.0
Model 1	ref	0.04 (0.00, 0.08)	0.04 (0.00, 0.08)	0.03 (-0.01, 0.07)	0.05	0.05 (0.00, 0.10)*
Model 2	ref	0.03 (-0.02, 0.07)	0.02 (-0.02, 0.07)	0.01 (-0.03, 0.05)	0.52	0.04 (-0.01, 0.08)

[1] Betas from log-transformed HOMA-IR (95CIs) were estimated across four categories split by quartile values (Q1 to Q4) or non-consumers + tertile categories with the lowest category as the reference, adjusted for covariates as follows: Model 1 included age (continuous), sex and energy intake (continuous). Model 2 was additionally adjusted for education (3 categories), smoking (3 categories), physical activity (continuous), alcohol consumption (4 categories), family history of diabetes (yes/no, RS-I and RS-II only) and intakes of fruit, vegetables, wholegrains, legumes, nuts, tea, coffee, red meat, and sugar-sweetened beverages (continuous). Model 4 was additionally adjusted for longitudinal waist circumference. Linear trend across intake range categories was assessed by including median values of each category as a continuous variable in the model.

[2] Betas from log-transformed HOMA-IR per 1 serving/day (see definition in Table 1) were estimated.

[3] Median intake per intake range category in servings/d. * P=0.01 to 0.05, ** P<0.01.

Abbreviations: CI, Confidence Interval; HOMA-IR, Homeostatic Model Assessment of Insulin Resistance; Q, Quartile. yogurt, and cream, but non-linear trends were inconclusive, as only a few participants had high intakes **(Supplemental Figure 1)**.

Associations in Sub-Cohorts

The association of a higher intake of high-fat yogurt with a lower prediabetes risk and longitudinal insulin resistance was consistently found in all three sub-cohorts (**Supplemental Tables 3 and 4**). For other dairy types, some discrepancies were found. Associations of low-fat dairy and low-fat milk with a higher prediabetes risk were found in RS-I and RS-II but not in RS-III, whereas the positive associations with HOMA-IR were observed in RS-II and RS-III but not in RS-I. The positive association for total cheese and prediabetes was observed in RS-I but not in RS-II and RS-III.

Sensitivity Analysis

We observed significant interactions of dairy consumption with sex and waist circumference on prediabetes risk (**Supplemental Table 5**) and with sex, age, and waist circumference on HOMA-IR (**Supplemental Table 6**). However, stratified analyses revealed no clear patterns. Associations of dairy types with prediabetes risk and longitudinal HOMA-IR were comparable, although for most, no longer statistically significant after additionally adjusting for longitudinal waist circumference (Model 3) (**Supplemental Table 7 and 8**). Furthermore, all associations were similar or only slightly attenuated after additional mutual adjustment for other dairy types; additional adjustment for cholesterol, hypertension, and triglycerides; exclusion of participants with prevalent and incident CHD; or using dairy intake as an energy adjusted variable with the residual method, instead. In a subsample of RS-I and RS-II with repeated dietary intake assessment, dairy intake at follow-up was similar to baseline - only total cheese intake was higher in RS-I (2.5 ± 2.1 versus 1.9 ± 1.1 servings/day) (**Supplemental Table 9**). After adjustment for dairy intake at follow-up, associations generally remained similar, except for low-fat dairy, total milk, and low-fat milk in RS-II (**Supplemental Table 10 and 11**).

Discussion

In this population-based cohort study, high-fat yogurt was consistently associated with a lower prediabetes risk and lower longitudinal insulin resistance. Additionally, high-fat milk was associated with a lower prediabetes risk but not with longitudinal insulin resistance. Higher intakes of low-fat dairy, low-fat milk and total and low-fat cheese were associated with a higher prediabetes risk but inconsistently across sub-cohorts and by variable type (continuous or quartiles). Higher intake of low-fat dairy and total and low-fat milk were associated with a higher longitudinal insulin resistance. Total dairy, fermented dairy, low-fat yogurt, high-fat cheese, cream, and ice cream were not associated with prediabetes risk or longitudinal insulin resistance.

Of the dairy types examined in our study, high-fat yogurt intake was most strongly associated with prediabetes and insulin resistance, consistent across all three sub-cohorts and robust in sensitivity analyses. Generally, prior studies lack information on the fat content of yogurt. Previous meta-analyses of observational studies showed

that a higher compared to a lower intake of total yogurt is significantly associated with a lower T2D risk (relative risks (RRs) ranging from 0.74-0.86 in five meta-analyses) [15, 16]. There are limited prospective [18, 19] and cross-sectional studies with prediabetes as the specific outcome [37, 38]. In the FHS Offspring cohort, yogurt intake was non-linearly associated with prediabetes, with the lowest risk at 2-4 servings/week but an increased risk with higher intakes [18]. In addition, the Dutch Maastricht cohort (*n* = 3,451) showed that the highest versus the lowest intake of yogurt was associated with lower odds of prediabetes (OR 0.67, 95%CI 0.50-0.90) [37]. On the contrary, no associations of yogurt with prediabetes were found in two other Dutch studies: the prospective Hoorn Studies [19] and Lifelines study [38]. An important explanation for heterogeneous associations between populations may be the potential confounding by unmeasured differences in population characteristics and health status. In Dutch populations, dairy foods are consumed by various population groups and within a wide range of diets, while, for example in the US, high dairy intake generally reflects overall healthier behaviour [40]. Furthermore, the quantity and composition of dairy-type categories vary by the availability of products and consumption habits in a region. FFQs do not assess the sugar content of yogurt products. The intake of plain and sugar-sweetened yogurt may differ between populations, both plausibly differentially associated with cardio metabolic outcomes. Overall, the evidence indicates a neutral or inverse association between yogurt and early phases of T2D, yet there is a need for studies explaining heterogeneity and further examining the role of fat content on cardio metabolic effects.

In our cohort, a higher intake of high-fat milk was associated with a lower prediabetes risk, but not with longitudinal insulin resistance. In the FHS Offspring Cohort, a non-linear association was found for high-fat milk, with moderate intakes associated with a lower prediabetes risk but a higher risk with higher intakes [18]. No associations with high-fat milk or milk were found in the Hoorn Studies [19] and two Dutch cross-sectional studies [37, 38]. Meta-analyses of observational studies indicate neutral associations of milk and T2D, confirmed by Mendelian randomization studies [39].

The associations found for higher intake of total dairy, low-fat dairy, low-fat milk and total and low-fat cheese and the outcomes were somewhat weaker and not found in the most recent Rotterdam sub-cohort. Previous analyses of the Rotterdam Study showed that protein from dairy was associated with a higher prediabetes risk (HR per 5% energy increment 1.26, 95%CI 1.06-1.49) and longitudinal insulin resistance (β 0.04, 95%CI 0.0003, 0.08), independent of other macronutrients and diet quality [40]. This suggests that specifically dairy protein intake might underlie the positive associations between low-fat dairy types and outcomes in the Rotterdam Study. Our results differ from the previous evidence of prospective studies [18, 19]. The FHS Offspring Cohort (*n* = 1,867) reported beneficial associations for total and low-fat dairy [18], and the Hoorn Studies (*n* = 2,262) found no associations between low-fat dairy types and prediabetes risk [19]. The intake of low-fat dairy and total protein was slightly higher in the Rotterdam Study

compared to the Hoorn Studies, plausibly explaining different associations. In addition, with higher low-fat dairy intake, energy intake and total carbohydrate intake was also higher, yet associations were independent of major sources of carbohydrates, suggesting that protein content could be responsible for the associations. These associations need further confirmation in trials. So far, the evidence from trials is inconclusive. A recent meta-analysis from our group of 54 controlled dietary intervention studies reported that higher protein diets led to greater weight loss, fat mass loss and beneficial reductions in systolic blood pressure and improved lipid and insulin outcomes compared to lower protein diets over a follow-up period of four to five months [41]. This study showed no detrimental effects and some beneficial effects of higher protein diets on body weight and markers of cardiometabolic health. Longer- term trials are warranted to give insights on the effects of specific dairy proteins [42]. Furthermore, our results contradict with meta-analyses of observational studies showing mostly neutral or slightly beneficial associations between these dairy types and T2D [15].

An additional source of inconsistencies in the associations of dairy types with prediabetes and insulin resistance may arise from the inclusion of participants with various metabolic states at baseline [18]. Potential non-linear associations, less pronounced effect estimates in insulin resistance models and potential measurement errors in dairy intake warrant replication studies. Recent meta-analyses of studies using experimental designs show contradicting results [43, 44]. Null associations (13 RCTS, n = 840) [43] as well as significant reductions in HOMA-IR (14 RCTs, n = 794) [44] have been reported when comparing diets high and low in dairy. In addition, a recent RCT showed that both high-fat and low-fat dairy diets, compared to a diet low in dairy, were associated with decreased insulin sensitivity [45]. Nevertheless, it is worthy to note that this study included participants with the metabolic syndrome, and the primary outcome, glucose tolerance, did not change. Overall, these interventions were heterogeneous in the study population (age, co-morbidities), duration and treatment and control diets. Furthermore, the authors did not specify the dairy type or fat content. It is unknown if these results translate to long-term risk of diabetes, for which future well-designed trials and long-term studies are needed.

Multiple explanations have been proposed linking dairy intake, especially yogurt, to a lower risk of T2D development, although causal molecular mechanisms remain unclear [9]. Yogurt contains probiotics originating from the fermentation process. Probiotics have been associated with lower weight gain, lower cholesterol, and blood glucose levels in animal models, possibly by compositional and functional changes in the gut microbiome, increased butyrate production and anti-inflammatory effects [42, 46, 47]. In a study in people with prediabetes, participants with daily intake of yogurt enriched with *Lactobacillus plantarum* showed greater reductions in HbA1c levels compared to participants with a daily intake of conventional yogurt [48]. Some dairy fats and proteins have been related to pathways linked to a lower risk of prediabetes in animal and in

vitro studies, albeit the content of these specific nutrients in yogurt is low. For example, branched-chain and ruminant trans fatty acids may inhibit hepatic de novo lipogenesis, improve insulin resistance, and reduce inflammation [46, 49]. In vitro, whey proteins have shown to upregulate hepatic glucose metabolism through gene expression regulation [50]. Branched-chain amino acids may activate the mammalian target of the rapamycin complex (mTOR) signalling pathway upregulating insulin secretion, resulting in enhanced glucose clearance [51]. Yet, prolonged increased insulin levels may lead to insulin resistance and T2D [52]. Nevertheless, high protein diets show favourable effects on weight loss in RCTs [41], related to effects on gut-derived hormones and thermogenesis promoting satiating, and preservation of fat-free mass during weight loss [53]. We did observe that both beneficial associations of high-fat yogurt and positive associations of low-fat dairy slightly attenuated after adjustment for longitudinal waist circumference, suggesting obesity may partly mediate some associations [40]. However, we observed that associations were independent of blood lipids and hypertension, suggesting that these factors did not play a role.

The current study has multiple strengths. First, we examined associations of several dairy types with both prediabetes and insulin resistance in a large population-based cohort. In addition, our study provided temporal associations with repeated measures of insulin resistance and a considerable follow-up duration. Second, to our knowledge, this is the first study examining associations of dairy with repeated measures of insulin resistance. Third, the associations were controlled for a wide range of confounders. These included major energy-providing food groups previously associated with development of T2D to prevent confounding by background diet, which is not widely done in dairy-diabetes research, thus improving the quality of the current evidence on dairy-diabetes research [54].

There are also some limitations to the current study. First, measurement errors in the habitual dairy intake assessment, for example due to recall bias, may result in bias towards the null. Furthermore, dairy intake might have changed over time. However, a sensitivity analysis incorporating repeated measures of dairy consumption and excluding participants likely to change their diet due to diagnosis of cardiovascular diseases showed similar associations for most dairy types. Second, the between-person variation in the intake of several dairy products, such as high-fat yogurt and low-fat cheese, were limited due to the observational nature of our study. Third, residual and unmeasured confounding can never be ruled out in observational studies, for example, by potential effects of meal frequency and timing and replacement choices for dairy consumption [55, 56]. Fourth, no 2 h plasma glucose levels were available, and using the FPG only to define prediabetes cases may lead to underestimation of prediabetes cases [2]. This possible non-differential misclassification of the outcome may have resulted in bias towards the null.

Conclusions

In this population-based cohort study, high-fat yogurt showed robust inverse associations with a prediabetes risk and longitudinal insulin resistance. Higher intake of high-fat milk was also associated with a lower prediabetes risk. Low-fat dairy, total milk, low-fat milk, and total and low-fat cheese were positively associated with the outcomes but inconsistently. With the current study, we extend the understanding of the role of dairy intake before clinical stages and decreasing the risk of reverse causation by the presence of disease. Well-designed prospective cohort studies and long-term trials are needed to confirm associations and to explore confounding factors.

List of supplementary materials chapter 3

Supplemental Table 1. Missing values of covariates in participants before imputation.

Supplemental Table 2. Baseline characteristics of study population across different population subgroups (total *n* = 6,770).

Supplemental Figure 1. Non-linear relationship of dairy product types and longitudinal HOMA-IR by sub-cohort.

Supplemental Table 3. Associations of dairy product types and prediabetes risk in the Rotterdam Studies, pooled and by sub-cohort (*n* = 6,053).

Supplemental Table 4. Associations of dairy product types and longitudinal insulin resistance in the Rotterdam Studies, pooled and by sub-cohort (*n* = 6,593).

Supplemental Table 5. Associations of dairy product types and prediabetes risk in the Rotterdam Studies, stratified by sex and baseline waist circumference (*n* = 6053).

Supplemental Table 6. Associations of dairy intake and longitudinal insulin resistance in the Rotterdam Studies, stratified by sex, age and waist circumference (*n* = 6,593).

Supplemental Table 7. Sensitivity analyses of associations of dairy product types and prediabetes risk in the Rotterdam Studies (*n* = 6,053).

Supplemental Table 8. Sensitivity analyses of associations of dairy product types and longitudinal insulin resistance in the Rotterdam Studies (*n* = 6,593).

Supplemental Table 9. Mean dairy intake at baseline and follow-up in a subsample of RS-I and RS-II (*n* = 1,887).

Supplemental Table 10. Associations of dairy product types and prediabetes risk using repeated measures of dairy intakes as time-dependent exposure (*n* = 6,053) or adjusting for dairy intake at follow-up (*n* = 1,707*).*

Supplemental Table 11. Associations of dairy product intakes and longitudinal insulin resistance with repeated measures of dairy intake as fixed effect in RS-I and RS-II (*n* = 4,274).

Scan this QR code to download the supplementary materials.

References

1. Stumvoll, M., B.J. Goldstein, and T.W. van Haeften, *Type 2 diabetes: principles of pathogenesis and therapy.* Lancet, 2005. **365**(9467): p. 1333-46.
2. World Health Organization (WHO), *Definition and diagnosis of diabetes mellitus and intermediate hyperglycaemia: report of a WHO/IDF consultation.* 2006.
3. Mutie, P.M., et al., *An investigation of causal relationships between prediabetes and vascular complications.* Nat Commun, 2020. **11**(1): p. 4592.
4. Cai, X., et al., *Association between prediabetes and risk of all cause mortality and cardiovascular disease: updated meta-analysis.* Bmj, 2020. **370**: p. m2297.
5. Zheng, Y., S.H. Ley, and F.B. Hu, *Global aetiology and epidemiology of type 2 diabetes mellitus and its complications.* Nat Rev Endocrinol, 2018. **14**(2): p. 88-98.
6. Hu, F.B., et al., *Diet, lifestyle, and the risk of type 2 diabetes mellitus in women.* N Engl J Med, 2001. **345**(11): p. 790-7.
7. Yang, Q., et al., *Trends in cardiovascular health metrics and associations with all-cause and CVD mortality among US adults.* Jama, 2012. **307**(12): p. 1273-83.
8. Long, G.H., et al., *Healthy behaviours and 10-year incidence of diabetes: a population cohort study.* Prev Med, 2015. **71**: p. 121-7.
9. Mozaffarian, D., *Dairy Foods, Obesity, and Metabolic Health: The Role of the Food Matrix Compared with Single Nutrients.* Advances in Nutrition, 2019. **10**(5): p. 917S-923S.
10. Li, Y., et al., *Effect of vitamin K2 on type 2 diabetes mellitus: A review.* Diabetes Res Clin Pract, 2018. **136**: p. 39-51.
11. Beulens, J.W., et al., *Dietary phylloquinone and menaquinones intakes and risk of type 2 diabetes.* Diabetes Care, 2010. **33**(8): p. 1699-705.
12. Comerford, K.B., et al., *Global Review of Dairy Recommendations in Food-Based Dietary Guidelines.* Front Nutr, 2021. **8**: p. 671999.
13. Drouin-Chartier, J.-P., et al., *Comprehensive review of the impact of dairy foods and dairy fat on cardiometabolic risk.* Advances in nutrition, 2016. **7**(6): p. 1041-1051.
14. Godos, J., et al., *Dairy foods and health: an umbrella review of observational studies.* International Journal of Food Sciences and Nutrition, 2020. **71**(2): p. 138-151.
15. Alvarez-Bueno, C., et al., *Effects of Milk and Dairy Product Consumption on Type 2 Diabetes: Overview of Systematic Reviews and Meta-Analyses.* Adv Nutr, 2019. **10**(suppl_2): p. S154-S163.
16. Gijsbers, L., et al., *Consumption of dairy foods and diabetes incidence: a dose-response meta-analysis of observational studies.* Am J Clin Nutr, 2016. **103**(4): p. 1111-24.
17. Brouwer-Brolsma, E.M., et al., *Intake of different types of dairy and its prospective association with risk of type 2 diabetes: The Rotterdam Study.* Nutr Metab Cardiovasc Dis, 2016. **26**(11): p. 987-995.
18. Hruby, A., et al., *Associations of Dairy Intake with Incident Prediabetes or Diabetes in Middle-Aged Adults Vary by Both Dairy Type and Glycemic Status.* J Nutr, 2017. **147**(9): p. 1764-1775.
19. Slurink, I.A.L., et al., *Dairy product consumption and incident prediabetes in Dutch middle-aged adults: the Hoorn Studies prospective cohort.* Eur J Nutr, 2022. **61**(1): p. 183-196.
20. Drehmer, M., 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): p. 775-82.
21. Turner, K.M., J.B. Keogh, and P.M. Clifton, *Dairy consumption and insulin sensitivity: a systematic review of short- and long-term intervention studies.* Nutr Metab Cardiovasc Dis, 2015. **25**(1): p. 3-8.
22. Ikram, M.A., et al., *The Rotterdam Study: 2018 update on objectives, design and main results.* Eur J Epidemiol, 2017. **32**(9): p. 807-850.
23. Neeley, W.E., *Simple automated determination of serum or plasma glucose by a hexokinase-glucose-6 -phosphate dehydrogenase method.* Clin Chem, 1972. **18**(6): p. 509-15.
24. Voortman, T., et al., *Adherence to the 2015 Dutch dietary guidelines and risk of non-communicable diseases and mortality in the Rotterdam Study.* Eur J Epidemiol, 2017. **32**(11): p. 993-1005.

25. Klipstein-Grobusch, K., et al., *Dietary assessment in the elderly: validation of a semiquantitative food frequency questionnaire.* Eur J Clin Nutr, 1998. **52**(8): p. 588-96.
26. Feunekes, G.I., et al., *Relative and biomarker-based validity of a food-frequency questionnaire estimating intake of fats and cholesterol.* Am J Clin Nutr, 1993. **58**(4): p. 489-96.
27. Goldbohm, R.A., et al., *Validation of a dietary questionnaire used in a large-scale prospective cohort study on diet and cancer.* Eur J Clin Nutr, 1994. **48**(4): p. 253-65.
28. Caspersen, C.J., et al., *The prevalence of selected physical activities and their relation with coronary heart disease risk factors in elderly men: the Zutphen Study, 1985.* Am J Epidemiol, 1991. **133**(11): p. 1078-92.
29. Stel, V.S., et al., *Comparison of the LASA Physical Activity Questionnaire with a 7-day diary and pedometer.* J Clin Epidemiol, 2004. **57**(3): p. 252-8.
30. Ainsworth, B.E., et al., *2011 Compendium of Physical Activities: a second update of codes and MET values.* Med Sci Sports Exerc, 2011. **43**(8): p. 1575-81.
31. Dutch Health Council (Gezondheidsraad), *Dutch dietary guidelines 2015 (Richtlijnen goede voeding 2015). Publication nr. 2015/24. ISBN 978-94-6281-089-1. The Hague.* 2015.
32. Leening, M.J., et al., *Methods of data collection and definitions of cardiac outcomes in the Rotterdam Study.* Eur J Epidemiol, 2012. **27**(3): p. 173-85.
33. Soedamah-Muthu, S.S. and J. Guo, *Dairy consumption and cardiometabolic diseases: Evidence from prospective studies*, in *Milk and Dairy Foods*. 2020. p. 1-28.
34. Schwingshackl, L., et al., *Food groups and risk of type 2 diabetes mellitus: a systematic review and meta-analysis of prospective studies.* Eur J Epidemiol, 2017. **32**(5): p. 363-375.
35. Rizopoulos, D., *Joint Models for Longitudinal and Time-to-Event Data: With Applications in R, ed. C.a. Hall/CRC. 2012.*
36. Rothman, K.J., *No adjustments are needed for multiple comparisons.* Epidemiology, 1990. **1**(1): p. 43-6.
37. Eussen, S.J., et al., *Consumption of dairy foods in relation to impaired glucose metabolism and type 2 diabetes mellitus: the Maastricht Study.* Br J Nutr, 2016. **115**(8): p. 1453-61.
38. Brouwer-Brolsma, E.M., et al., *Dairy product consumption is associated with pre-diabetes and newly diagnosed type 2 diabetes in the Lifelines Cohort Study.* Br J Nutr, 2018. **119**(4): p. 442-455.
39. Soedamah-Muthu, S. and J. Guo, *Dairy consumption and cardiometabolic diseases: Evidence from prospective studies. In D. I. Givens (Ed.), Milk and dairy foods: Their functionality in human health and disease (pp. 1-28). Academic Press*. 2020.
40. Chen, Z., et al., *Associations of specific dietary protein with longitudinal insulin resistance, prediabetes and type 2 diabetes: The Rotterdam Study.* Clin Nutr, 2020. **39**(1): p. 242-249.
41. Vogtschmidt, Y.D., et al., *Is protein the forgotten ingredient: Effects of higher compared to lower protein diets on cardiometabolic risk factors. A systematic review and meta-analysis of randomised controlled trials.* Atherosclerosis, 2021. **328**: p. 124-135.
42. Fekete Á, A., D.I. Givens, and J.A. Lovegrove, *Can milk proteins be a useful tool in the management of cardiometabolic health? An updated review of human intervention trials.* Proc Nutr Soc, 2016. **75**(3): p. 328-41.
43. O'Connor, S., et al., *Increased Dairy Product Intake Modifies Plasma Glucose Concentrations and Glycated Hemoglobin: A Systematic Review and Meta-Analysis of Randomized Controlled Trials.* Advances in Nutrition, 2019. **10**(2): p. 262-279.
44. Sochol, K.M., et al., *The Effects of Dairy Intake on Insulin Resistance: A Systematic Review and Meta-Analysis of Randomized Clinical Trials.* Nutrients, 2019. **11**(9).
45. Schmidt, K.A., et al., *The impact of diets rich in low-fat or full-fat dairy on glucose tolerance and its determinants: a randomized controlled trial.* Am J Clin Nutr, 2021. **113**(3): p. 534-547.
46. Mozaffarian, D. and J.H.Y. Wu, *Flavonoids, Dairy Foods, and Cardiovascular and Metabolic Health: A Review of Emerging Biologic Pathways.* Circ Res, 2018. **122**(2): p. 369-384.
47. Li, X., et al., *A comparative study of the antidiabetic effects exerted by live and dead multi-strain probiotics in the type 2 diabetes model of mice.* Food Funct, 2016. **7**(12): p. 4851-4860.
48. Toshimitsu, T., et al., *Effects of 12-Week Ingestion of Yogurt Containing Lactobacillus plantarum OLL2712 on Glucose Metabolism and Chronic Inflammation in Prediabetic Adults: A Randomized Placebo-Controlled Trial.* Nutrients, 2020. **12**(2).

49. Tremblay, B.L. and I. Rudkowska, *Nutrigenomic point of view on effects and mechanisms of action of ruminant trans fatty acids on insulin resistance and type 2 diabetes.* Nutr Rev, 2017. **75**(3): p. 214-223.
50. Da Silva, M.S., et al., *Dairy Product Consumption Interacts with Glucokinase (GCK) Gene Polymorphisms Associated with Insulin Resistance.* J Pers Med, 2017. **7**(3).
51. Jakubowicz, D. and O. Froy, *Biochemical and metabolic mechanisms by which dietary whey protein may combat obesity and Type 2 diabetes.* J Nutr Biochem, 2013. **24**(1): p. 1-5.
52. Kahn, B.B. and J.S. Flier, *Obesity and insulin resistance.* The Journal of Clinical Investigation, 2000. **106**(4): p. 473-481.
53. Moon, J. and G. Koh, *Clinical Evidence and Mechanisms of High-Protein Diet-Induced Weight Loss.* J Obes Metab Syndr, 2020. **29**(3): p. 166-173.
54. Guo, J., et al., *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. **32**(4): p. 269-287.
55. Paoli, A., et al., *The Influence of Meal Frequency and Timing on Health in Humans: The Role of Fasting.* Nutrients, 2019. **11**(4).
56. Fewell, Z., G. Davey Smith, and J.A. Sterne, *The impact of residual and unmeasured confounding in epidemiologic studies: a simulation study.* Am J Epidemiol, 2007. **166**(6): p. 646-55.

Chapter 4

Dairy intake and prediabetes risk in the Australian Diabetes, Obesity, and Lifestyle Study

Manuscript based on this chapter

Isabel A.L. Slurink, Lei Chen, Dianna J. Magliano, Nina Kupper, Tom Smeets, & Sabita S. Soedamah-Muthu (2023). *Dairy product consumption and incident prediabetes in the Australian Diabetes, Obesity, and Lifestyle study with 12 years of follow-up.* The Journal of Nutrition, 153(6), 1742–1752.

Abstract

Background

Investigating modifiable risk factors for the early stages of the development of type 2 diabetes is essential for effective prevention. Some studies show protective associations between dairy and prediabetes; however, associations are heterogeneous by the type and fat content of dairy foods.

Objective

To examine the relationship between the consumption of dairy, including different types of dairy products and risk of prediabetes.

Methods

The study included 4891 participants with normal glucose tolerance (aged 49.0 ± 12.3 years, 57% female) of the Australian Diabetes, Obesity, and Lifestyle (AusDiab) study, a longitudinal population-based study. Dairy intake was measured at baseline using a food frequency questionnaire. Prediabetes at the 5-y and 12-y follow-ups was defined according to the WHO criteria as fasting plasma glucose levels of 110-125 mg/dL or 2-h plasma glucose levels of 140-199 mg/dL. Associations were analysed using Poisson regression, adjusted for social demographics, health behaviours, a family history of diabetes, and food group intake.

Results

In total, 765 (15.6%) incident cases of prediabetes were observed. The mean intake of dairy foods was 2.4 ± 1.2 servings/day, mostly consisting of low-fat milk (0.70 ± 0.78 servings/day) and high-fat milk (0.47 ± 0.72 servings/day). A higher intake of high-fat dairy ($RR_{serving/day}$ 0.92, 95%CI 0.85, 1.00), high-fat milk (0.89, 95%CI 0.80, 0.99), and total cheese (0.74, 95%CI 0.56, 0.96) was associated with a lower risk of prediabetes. Low-fat milk intake was associated nonlinearly with prediabetes risk. Low-fat dairy foods, total milk, yogurt, low-fat cheese, and ice cream were not associated with prediabetes risk.

Conclusion

In this large Australian cohort, protective associations were found for high-fat dairy types, whereas neutral associations were seen for low-fat dairy types. Studies with more detail on sugar content of types of dairy foods and products eaten with dairy foods (e.g., cereals or jam), and studies into potential causal mechanisms of the health effects of dairy intake are required.

Introduction

Prediabetes is defined as the intermediate stage between normal glucose tolerance (NGT) and type 2 diabetes (T2D), including impaired fasting glucose and/or impaired glucose tolerance [1]. People in this early risk stage of T2D already display metabolic disturbances and are prone to develop microvascular and macrovascular complications [2-5]. The prevalence of prediabetes is increasing worldwide [6, 7]; in particular, a high prevalence is observed in people with obesity and of older age [8]. Prevention is needed because a significant proportion of people with prediabetes will develop T2D over time (cumulative incidence of 9%-84%, depending on the follow-up duration and prediabetes definition) [9], and the incidence of cardiovascular disease among people with prediabetes is substantial [relative risk (RR): 1.15, 95%CI 1.11, 1.18 compared with NGT] [5]. Furthermore, people with prediabetes may revert to NGT [9, 10]. Lifestyle modification is the recommended approach to prevent and treat prediabetes, and its effectiveness has been shown in randomized controlled trials (RCTs) [11, 12]. Dairy food is a key component in many diets and has, therefore, received ample attention in the literature. Nevertheless, the health effects of dairy foods are heterogeneous, partly underlined by differences in study populations and variations in consumed dairy types [13, 14].

Few prospective cohort studies have presented associations between dairy intake and incident prediabetes. In the Framingham Offspring Cohort (FHS-OC) (*n* = 1,867, 12 years of follow-up), total, low-fat, and high-fat dairy were associated with a 39%, 32%, and 25% lower risk of incident prediabetes, respectively, for the top compared with the bottom intakes [15]. In addition to this study in the United States, our research group conducted analyses in 2 large Dutch cohort studies because dairy consumption is much higher and more widespread in the Netherlands compared with that in other countries. In the Hoorn Studies (pooling 2 cohorts, *n* = 2,262, 6.4 years of follow-up), high-fat fermented dairy, total cheese, and high-fat cheese were associated with a lower risk of prediabetes, but total dairy and other types of dairy were not associated with prediabetes [16]. On the contrary, in the Rotterdam studies (pooling 3 cohorts, *n* = 6053, 11.4 years of follow-up), high-fat yogurt and high-fat milk intake were strongly associated with a lower prediabetes risk, but low-fat dairy and low-fat milk were associated with a higher prediabetes risk [17].

In a previous analysis of the Australian Diabetes, Obesity, and Lifestyle (AusDiab) study (*n* = 5582, 5-y follow-up), a nonsignificant association of total dairy and T2D incidence was found (OR: 0.71, 95%CI 0.48, 1.05 for the third compared with the first tertile) [18]. Low-fat milk was significantly associated with lower odds of diabetes incidence (OR: 0.65, 95%CI 0.44, 0.94), whereas there was no association with full-fat milk, yogurt, and cheese. The relationship between dairy intake and prediabetes has not yet been investigated in the AusDiab population. Investigating this early stage is essential because potential associations between dairy and prediabetes have important implications for effective early-stage prevention of diabetes and cardiovascular disease.

Therefore, the aim of this study was to examine the relationship between the consumption of dairy, including different types of dairy products, and prediabetes risk in the nationwide, on a population-based longitudinal AusDiab study with 12 years of follow-up.

Methods

Study design and population

The AusDiab study is a national, population-based survey of 11,247 adults aged older than 25 years in 1999-2000, with follow-up measurements in 2004-2005 and 2011-2012; details are described previously [19]. The AusDiab study aimed to provide national benchmark data on the prevalence, incidence and risk factors of diabetes, obesity, hypertension, and kidney disease in Australia. In short, a stratified cluster sample was drawn from 42 randomly selected census collector districts across Australia, including mostly participants with an Australian, New Zealand or British background (85%).

Participants were interviewed at home after which they received a biomedical examination at the centre. All eligible participants were invited to attend the follow-up measurement, excluding those who were deceased, had moved overseas or into a nursing facility classified for high care, or had a terminal illness. Of those completing the baseline household interview, 55% completed the biomedical examination. Differences between responders and non-responders have been described previously [19]. The baseline measurements were repeated at 5 years (response rate 60.6% of eligible participants) and 12 years of follow-up (response rate, 59.8% of eligible participants) [20-22]. The study was approved by the human research ethics committee of the International Diabetes Institute, and the Alfred Hospital (Melbourne, Australia). All participants provided written informed consent.

For this analysis, we excluded pregnant participants (*n* = 60), participants with missing baseline dietary data (*n* = 203), or those with implausible energy intakes (defined as men <800 or >4,200 kcal, women <500 or >3,500 kcal) (*n* = 276) [23, 24] (**Figure 1**). Furthermore, we excluded participants with prediabetes or diabetes at baseline (*n* = 2,823) or with missing information related to glucose testing at baseline (*n* = 94), resulting in 7791 participants without prediabetes or diabetes at baseline. Participants with complete follow-up information for the 5-y and/or 12-y examination and without diabetes or missing prediabetes information at the follow-up were included in the analysis, resulting in an analytical sample of 4891 participants.

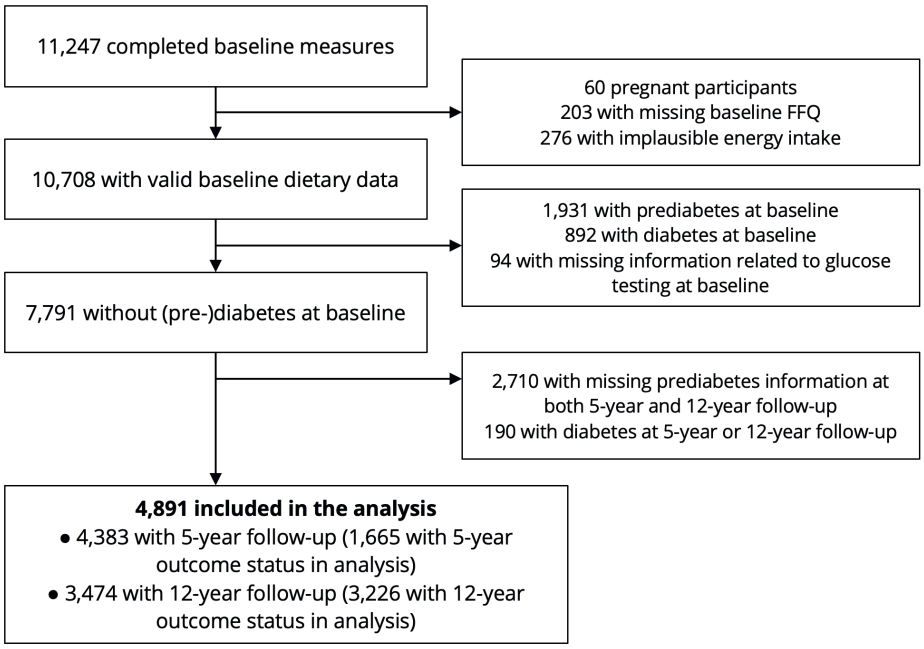


Figure 1. Flow-chart for inclusion of participants for the present analysis of the Ausdiab Study.

Ascertainment of prediabetes and type 2 diabetes

At baseline and both follow-up measurements, blood samples were collected after an overnight fast (≥8 h) [25]. All participants except those on diabetes medication or who were pregnant underwent a standard 75-g oral glucose tolerance test (OGTT) [1]. Fasting plasma glucose (FPG) levels and 2-h glucose levels (2hPG) based on the OGTT were determined with a glucose oxidase method at baseline and with a spectrophotometric-hexokinase method at both follow-up measurements [21]. Prevalent prediabetes at baseline and incident prediabetes at the follow-up was defined as FPG levels in the range 110 and 125 mg/dL (6.1 and 6.9 mmol/L) or 2hPG between 140 and 199 mg/dL (7.8 and 11.0 mmol/L) based on the WHO criteria [1, 20]. Additional sensitivity analyses were per- formed with FPG levels in the range 100 and 125 mg/dL (5.6 and 6.9 mmol/L) based on the American Diabetes Association (ADA) criteria [26]. T2D was defined as FPG levels of ≥126 mg/dL (≥7.0 mmol/L) or 2hPG levels of ≥200 mg/dL (≥11.1 mmol/L) or current treatment with insulin or oral hypoglycaemic agents. Incident prediabetes was defined as occurrence of the outcome at 5-y or 12-y follow-up. Thus, participants with prediabetes at 5-y follow-up were coded as prediabetes regardless of their 12-y follow-up status (n = 255) or if they had missing 12-y follow-up status (n = 153). Participants with NGT at 12 y and NGT or missing information at 5 y were coded as having NGT (n = 2,869).

Dietary assessment

Dietary intake in the last 12 months was measured without the interviewer's assistance using a self-administered 74-item FFQ designed by the Cancer Council Victoria [27]. This FFQ was compared with 7-d weighted food records in 63 women of childbearing age, with correlation coefficients of 0.39 for protein, 0.64 for saturated fat, and 0.30 for sodium. Diet quality was measured with the food-based Australian Guideline Index, comprising 15 items, with scores ranging from 0 to 150, with higher scores reflecting better adherence to dietary guidelines [28]. Questions on dairy intake included milk (full-fat milk, reduced fat milk, and skimmed milk), flavoured milk, yogurt, cheese (hard cheese, firm cheese, soft cheese, low-fat cheese, ricotta/cottage cheese, and cream cheese), and ice cream. For milk intake, participants were asked to report quantity of intake per day (from "none" to "3 cups or more"). For other dairy food types, participants were asked to select from 10 frequency responses ("never" to "3 or more times per day") for each item on the FFQ. Daily intakes in grams were calculated using sex-specific standard portion sizes derived from weighted food records. Nutrient intakes were calculated using the NUTTAB95 food composition table [29]. Intake in grams were converted to standard serving size, that is, milk: 250 g, yogurt: 200 g, cheese: 40 g, and ice cream: 50 g [30]. In the combined total dairy and fermented dairy category, a serving of liquid dairy products was defined as 200 mL and cheese as 20 g. Dairy types analysed included total dairy, high-fat dairy (liquid products ≥2%, cheese ≥20%; including full-fat milk, hard cheese, firm cheese, soft cheese, cream cheese, and ice cream), low-fat dairy (liquid products <2%, cheese <20%; including reduced-fat milk, skimmed milk, low-fat cheese, and ricotta/cottage cheese), and fermented dairy (yogurt and hard cheese, firm cheese, soft cheese, and low-fat cheese) (**Table 1**). Because the FFQ did not distinguish between high-fat and low-fat yogurt types, yogurt intake was divided equally across the high-fat and low-fat dairy category according to the Australian Health Survey showing that ~ 45% of yogurt consumption was low-fat and ~48% was regular or high-fat yogurt [31].

Table 1. Food items and their fat content included in total dairy and dairy types and consumption in the Ausdiab (*n* = 4,781).

Dairy product	Included dairy types	Consumers	Intake (servings/day)	
		N (%)	Mean ± SD	Median [IQR]
Total dairy	All dairy products	4874 (99.7)	2.38 ± 1.20	2.29 [1.52, 3.04]
High-fat	High-fat milk, high-fat cheese, ice cream, yogurt (50% of intake)	4843 (99.0)	0.76 ± 0.92	0.33 [0.08, 1.12]
Low-fat	Low-fat milk, low-fat cheese, yogurt (50% of intake)	4236 (86.6)	0.98 ± 1.01	1.00 [0.03, 1.88]
Total fermented dairy	Yogurt, hard cheese, firm cheese, soft cheese, low-fat cheese	4739 (96.9)	0.80 ± 0.65	0.72 [0.33, 1.13]
Total milk	High-fat milk, low-fat milk, flavoured milk (2 g)	4552 (93.1)	1.17 ± 0.72	0.82 [0.80, 1.50]

Table 1. Food items and their fat content included in total dairy and dairy types and consumption in the Ausdiab (*n* = 4,781). (continued)

Dairy product	Included dairy types	Consumers	Intake (servings/day)	
		N (%)	Mean ± SD	Median [IQR]
High-fat	Full-fat milk (full cream milk) (3.3 g)	1811 (37.0)	0.47 ± 0.72	0 [0, 0.80]
Low-fat	Reduced fat milk (1.2 g), skimmed milk (0 g)	2699 (55.2)	0.70 ± 0.78	0.80 [0, 1.50]
Yogurt	Yogurt (3.8 g)	3696 (75.6)	0.19 ± 0.28	0.05 [0.01, 0.20]
Cheese	All cheeses	4555 (93.1)	0.33 ± 0.29	0.26 [0.10, 0.45]
High-fat	Hard cheese (33.5 g), firm cheese (31.3 g), soft cheese (21.7 g), cream cheese (34.2 g)	3840 (78.5)	0.25 ± 0.27	0.20 [0.04, 0.36]
Low-fat	Low-fat cheese (15.3 g), ricotta (11.8 g) or cottage cheese (5.7 g)	1165 (23.8)	0.08 ± 0.19	0 [0, 0]
Ice cream	Ice cream (5.8 g)	4318 (88.3)	0.26 ± 0.43	0.11 [0.03, 0.30]

[1] Values are mean ± SD. Consumers were defined as consuming >0 servings/day of a specific dairy type. Serving sizes were milk, 250 mL; yogurt, 200 mL; cheese, 40 g; and ice cream, 50 g. Combined total dairy category: liquid dairy products, 200 mL; cheese, 20 g.

Covariates

Interviewer-administered questionnaires were used to collect data on demographic and health-related information. Educational level was categorized into primary school/never attended, some secondary school, completed secondary school, or university/further higher education. A smoking history was assessed using a validated questionnaire, and participants were categorized as current, past, and never smoker [32]. Total leisure-time physical activity was measured using the Active Australia questionnaire [33, 34]. Total physical activity was calculated as the sum of time spent on walking (if continuous and for ≥10 min) or moderate-intensity activities, plus double the time spent in vigorous-intensity activities in the past week [35]. Physical activity was categorized as none, insufficient, 1-149 min/wk or sufficient, ≥150 min/wk. A family history of diabetes was defined as having a parent diagnosed with diabetes. The presence of cardiovascular disease (CVD) was obtained during the interviewer-administered questionnaire. BMI was calculated as weight (in kilograms) divided by height (in meters) squared.

Waist circumference (WC) was measured in duplicate halfway between the lower border of the ribs and the iliac crest on a horizontal plane [36]. Serum triacylglycerol, total cholesterol, LDL-cholesterol, and HDL-cholesterol concentrations were measured using standard enzymatic methods (Olympus AU600 analyzer; Olympus Optical) in serum fasting samples. Systolic and diastolic blood pressure levels were measured in a seated position using a Dinamap oscillometric blood pressure recorder (GE Healthcare), except in Victoria, where a standard mercury phygmomanometer was used with appropriate

adjustments to calculate blood pressure levels [37]. Hypertension was defined as systolic blood pressure of ≥140 mm Hg, diastolic blood pressure of ≥90 mm Hg, and/or reporting the use of antihypertensive medication.

Statistical analysis

Descriptive data were presented as means and SDs for continuous variables, medians and IQRs for nonnormally distributed continuous variables, and frequencies and percent- ages for categorical variables. Poisson regression models with robust variance were used because of the high incidence of prediabetes, in which case an odds ratio would overestimate the strength of the association [38]. Relative risks and 95% CIs for associations between dairy types and prediabetes incidence were calculated for dairy types in tertiles (reference lowest) and continuously (servings/day). For dairy types with many zero-consumers, a non-consumer category (reference) was made, and consumers were dichotomized at the median value. The P_{trend} was calculated by incorporating the median values of dairy tertiles as continuous variables in the model. For each model, we examined whether nonlinear terms of continuous dairy types (second-order polynomials or natural splines with 3 to 5 knots depending on the intake range, excluding outliers) significantly improved model fit compared with the linear model assessed by likelihood ratio tests. To adjust for potential bias associated with missing data, a multiple imputation procedure (n = 10) was used for missing data on covariates (**Supplemental Table 1**). No corrections for multiple testing were made as most exposures were correlated, and corrections may result in a type II error [39]. Statistical procedures were performed with the software STATA (version 15.1).

Confounders were selected based on the literature [40-42]. The basic model (model 1) adjusted for age, sex, and energy intake. Model 2 was additionally adjusted for educational level, alcohol intake, smoking status, physical activity, and genetic background with a family history of T2D. Model 3 additionally adjusted for food group intake associated with T2D including fruits, vegetables, whole grains, legumes, nuts, red and processed meat, and fruit juice intake. Model 4 additionally adjusted for baseline WC, change in WC from baseline to follow- up, LDL-cholesterol and hypertension as these variables are potential mediators.

To further examine possible confounding of associations by a healthy life- style, we presented descriptive data stratified by the dairy food types significantly associated with the outcomes. Potential effect modification by age, sex, and WC were explored in model 3, and stratified associations were presented in case of significant interactions ($P < 0.05$).

Multiple sensitivity analyses were performed in model 3 to examine the robustness of the findings. First, we additionally adjusted for intake of all other dairy types to assess whether associations of certain dairy types were independent of each other. Second, participants with prevalent CVD were excluded to address reverse causation by change of diet and lifestyle (n = 211). Third, associations were calculated with energy-adjusted

intake of dairy types using the residual method [24]. Fourth, analyses were repeated using the ADA cutoff levels for prediabetes [26].

Results

Participant characteristics

In 4,891 participants with NGT as measured by blood glucose at baseline, the mean age was 49.0 ± 12.3 y, 57% were female, and 12% were current smokers (**Table 2**). The mean WC was 88.0 ± 12.8 cm, the mean BMI was 26.1 ± 4.3 kg/m2, and 15.7% were obese (BMI ≥ 30 kg/m2). The mean total dairy intake was 2.4 ± 1.2 servings/day, mostly consisting of low-fat milk (0.70 ± 0.78, consumed by 55% of participants) and high-fat milk (0.47 ± 0.72, consumed by 37% of participants) intake (**Table 1**). Participants in the highest (3.7 ± 0.9 servings/day) compared with the lowest (1.2 ± 0.5 servings/day) tertile of total dairy intake recorded a higher educational level (47.9 versus 39.1% with university/further education level), were more physically active (58.8 versus 52.2% with sufficient (≥ 150 min/wk) level) and were less likely to be hypertensive (19.9 versus 24.7%) (**Table 2**). Furthermore, the average diet quality, energy intake, and intake of fruits, vegetables, grains, and fruit juice were higher in participants with the highest dairy intake that those with the lowest intake. Characteristics by the intake of specific dairy food types and by included versus excluded from the current analyses are presented in **Supplemental Table 2**.

Table 2. Baseline characteristics of participants in the AusDiab study in the total population and across tertiles (T) of total dairy intake (*n* = 4,891).

	Total (*n* = 4891)	**T1** (*n* = 1646)	**T2** (*n* = 1624)	**T3** (*n* = 1621)
Total consumption (serving/day)	2.4 ± 1.2	1.2 ± 0.5	2.3 ± 0.3	3.7 ± 0.9
Range	0.0-9.1	0.0-1.8	1.8-2.7	2.7-9.1
Follow-up time (5 y/12 y)	34.0/66.0	35.7/64.3	34.0/66.0	32.5/67.6
Age at baseline (y)	49.0 ± 12.3	49.4 ± 11.7	49.0 ± 12.7	48.7 ± 12.4
Sex, female	56.7	56.7	56.5	56.9
Educational level				
Primary school/never attended school	3.4	4.0	3.3	2.8
Completed some high school	33.1	37.4	32.3	29.5
Completed high school	19.6	19.6	19.3	19.8
University/further education	44.0	39.1	45.1	47.9
Smoking				
Current	11.5	12.3	11.2	11.2
Former	28.1	28.3	29.4	26.7
Never	60.4	59.4	59.5	62.2

Table 2. Baseline characteristics of participants in the AusDiab study in the total population and across tertiles (T) of total dairy intake (*n* = 4,891). (continued)

	Total (*n* = 4891)	**T1** (*n* = 1646)	**T2** (*n* = 1624)	**T3** (*n* = 1621)
Physical activity level				
Inactive (0 min/wk)	14.6	17.9	13.0	13.0
Insufficient (1-149 min/wk)	29.8	29.9	31.1	28.2
Sufficient (≥150 min/wk)	55.6	52.2	56.0	58.8
Family history of diabetes	17.7	18.4	17.8	16.8
BMI (kg/m2)	26.1 ± 4.3	26.1 ± 4.4	26.1 ± 4.3	26.0 ± 4.3
Waist circumference (cm)	88.0 ± 12.8	88.3 ± 12.9	88.1 ± 12.8	87.8 ± 12.7
Total cholesterol (mmol/L)	5.6 ± 1.0	5.6 ± 1.1	5.6 ± 1.0	5.6 ± 1.0
LDL-cholesterol (mmol/L)	3.5 ± 0.9	3.5 ± 0.9	3.5 ± 0.9	3.5 ± 0.9
HDL-cholesterol (mmol/L)	1.5 ± 0.4	1.5 ± 0.4	1.5 ± 0.4	1.5 ± 0.4
TAG (mmol/L)	1.1 (0.8-1.7)	1.2 (0.8-1.7)	1.1 (0.8-1.7)	1.1 (0.8-1.6)
Hypertension	22.6	24.7	23.4	19.9
Dietary intake				
Energy intake (kcal/d)	1910 ± 657	1689 ± 590	1876 ± 610	2170 ± 677
Diet quality	84.6 ± 13.9	80.3 ± 14.5	85.0 ± 13.0	88.6 ± 12.9
Fruit (g/day)	198.0 ± 136.5	193.8 ± 144.8	191.9 ± 129.8	208.5 ± 133.9
Vegetables (g/day)	104.1 ± 49.1	100.6 ± 50.9	104.2 ± 47.7	107.6 ± 48.4
Grains (g/day)	164.3 ± 122.3	152.1 ± 120.6	158.3 ± 118.3	182.7 ± 125.7
Legumes (g/day)	27.7 ± 19.3	27.4 ± 20.0	27.5 ± 19.2	28.2 ± 18.6
Nuts (g/day)	3.4 ± 6.4	3.6 ± 7.0	2.7 ± 4.6	3.9 ± 7.2
Meat (red and processed) (g/day)	97.8 ± 73.8	93.5 ± 75.6	96.8 ± 72.8	103.3 ± 72.5
Fruit juice (g/day)	86.3 ± 120.8	80.3 ± 128.4	83.3 ± 109.2	95.5 ± 123.3
Total fat (en%)	36.0 ± 5.7	36.4 ± 5.6	35.8 ± 5.6	35.8 ± 5.8
Saturated fat (en%)	14.4 ± 3.5	13.7 ± 3.3	14.4 ± 3.4	14.9 ± 3.6
Carbohydrates (en%)	45.2 ± 6.0	45.1 ± 6.4	45.4 ± 5.9	45.0 ± 5.5
Protein (en%)	19.3 ± 3.0	19.0 ± 3.1	19.2 ± 3.0	19.6 ± 2.8
Calcium (mg/day)	908 ± 326	619 ± 172	875 ± 148	1233 ± 277
Sodium (mg/day)	2612 ± 991	2318 ± 909	2575 ± 932	2950 ± 1026
Alcohol (g/day)	13.0 ± 16.7	13.8 ± 17.9	12.9 ± 16.6	12.3 ± 15.5

Values are mean ± SD for continuous variables with a normal distribution, or median [IQR] for continuous variables with a skewed distribution, percentages for categorical variables, based on unimputed data. Abbreviations: en%, percentage of total energy intake; LDL, low-density lipoprotein; HDL, high-density lipoprotein; TAG, triacylglycerol.

Dairy intake and prediabetes risk

A total of 765 incident prediabetes cases were identified (15.6%): 408 at the 5-y follow-up (of 4,383, 9.3%) and 357 at the 12-y follow-up (of 3,474, 10.3%). A higher intake of total dairy was significantly associated with a lower prediabetes risk in model 2 ($RR_{serving/day}$ 0.94, 95%CI 0.89-1.00), but this association was no longer significant after additional adjustment for dietary intake in model 3 (**Table 3**). Furthermore, a higher intake of fermented dairy was associated with a lower prediabetes risk in model 2 ($RR_{serving/day}$ 0.88, 95%CI 0.78-0.99) but non-significantly in model 3 ($RR_{serving/day}$ 0.91, 95%CI 0.80-1.02). A higher intake of high-fat dairy (44.4% of total dairy intake) was significantly associated with a lower prediabetes risk in fully adjusted models when analysed on a continuous scale ($RR_{serving/day}$ 0.92, 95%CI 0.85-1.00) but not when analysed in tertiles. In line with the results of high-fat dairy intake, a higher intake of high-fat milk (39.8% of total milk) was associated with a lower prediabetes risk on a continuous scale in model 3 ($RR_{serving/day}$ 0.89, 95%CI 0.80-0.99, p = 0.03) and borderline significant in the third compared with those of the first intake category (RR_{T3vsT1} 0.79, 95%CI 0.65-0.97, P_{trend} = 0.05). Furthermore, a higher intake of total cheese was associated with a lower prediabetes risk ($RR_{serving/day}$ 0.74, 95%CI 0.56-0.96). A higher intake of high-fat cheese (77.1% of total cheese) was associated with a lower prediabetes risk in model 2 ($RR_{serving/day}$ 0.71, 95%CI 0.54-0.95), but this association was similar but no longer significant in model 3 (0.75, 95%CI 0.57-1.00). The third compared with the first intake category of low-fat milk (59.8% of total milk) intake was associated with a higher prediabetes risk in model 3 (RR_{T3vsT1} 1.15, 95%CI 0.97-1.35, P_{trend} = 0.04), although continuous associations were not significant. A better fit of a nonlinear association was found for low-fat milk (P = 0.03) consumption, showing an increased risk of up to 1.5 servings/ d and a lower risk with higher intakes (**Figure 2**). Low-fat dairy, total milk, total yogurt, low-fat cheese, and ice cream were not associated with risk of prediabetes in multivariate-adjusted models. With additional adjustment for the potential mediators baseline WC, change in WC from baseline to follow-up, LDL-cholesterol, and hypertension in model 4, only total cheese remained associated with a lower prediabetes risk ($RR_{serving/day}$ 0.74, 95%CI 0.57-0.96).

Table 3. The associations of dairy intake and prediabetes risk in the AusDiab study (*n* = 4,891).

	Relative risk (95% CI) across intake range categories[1]				Continuous[2]
	T1	T2	T3	P_{trend}	RR (95%CI)
Total dairy					
n cases/n total	280/1646	250/1624	235/1621		765/4891
Median, servings/d	1.3	2.3	3.4		
Model 1	1 (ref)	0.92 (0.78-1.07)	0.88 (0.75-1.04)	0.14	0.93 (0.88-0.99)*
Model 2	1 (ref)	0.93 (0.80-1.09)	0.91 (0.77-1.07)	0.26	0.94 (0.89-1.00)*
Model 3	1 (ref)	0.97 (0.82-1.13)	0.97 (0.82-1.15)	0.72	0.96 (0.91-1.02)
Model 4	1 (ref)	0.97 (0.83-1.13)	0.97 (0.82-1.15)	0.73	0.96 (0.90-1.02)

Table 3. The associations of dairy intake and prediabetes risk in the AusDiab study (*n* = 4,891). (continued)

	Relative risk (95% CI) across intake range categories[1]				Continuous[2]
	T1	T2	T3	P_{trend}	RR (95%CI)
High-fat dairy					
n cases/n total	258/1637	280/1645	227/1609		
Median, servings/d	0.05	0.3	1.9		
Model 1	1 (ref)	1.06 (0.90-1.24)	0.90 (0.76-1.06)	0.08	0.91 (0.84-0.98)*
Model 2	1 (ref)	1.07 (0.92-1.25)	0.89 (0.75-1.06)	0.06	0.90 (0.83-0.98)*
Model 3	1 (ref)	1.10 (0.94-1.28)	0.92 (0.77-1.09)	0.12	0.92 (0.85-1.00)*
Model 4	1 (ref)	1.12 (0.96-1.31)	0.96 (0.81-1.14)	0.29	0.94 (0.87-1.02)
Low-fat dairy					
n cases/n total	237/1608	270/1664	258/1619		
Median, servings/d	0.01	1.0	1.9		
Model 1	1 (ref)	1.13 (0.96-1.32)	1.09 (0.93-1.28)	0.29	1.03 (0.97-1.10)
Model 2	1 (ref)	1.15 (0.98-1.35)	1.13 (0.96-1.34)	0.14	1.05 (0.98-1.11)
Model 3	1 (ref)	1.16 (0.98-1.36)	1.16 (0.98-1.37)	0.08	1.06 (0.99-1.13)
Model 4	1 (ref)	1.13 (0.96-1.32)	1.10 (0.94-1.30)	0.24	1.04 (0.97-1.11)
Fermented dairy					
n cases/n total	294/1656	231/1530	240/1703		
Median, servings/d	0.2	0.7	1.4		
Model 1	1 (ref)	0.89 (0.76-1.04)	0.83 (0.71-0.98)	0.03*	0.87 (0.78-0.98)*
Model 2	1 (ref)	0.89 (0.76-1.04)	0.85 (0.72-0.99)	0.05	0.88 (0.78-0.99)*
Model 3	1 (ref)	0.91 (0.78-1.07)	0.88 (0.75-1.04)	0.16	0.91 (0.80-1.02)
Model 4	1 (ref)	0.90 (0.77-1.05)	0.92 (0.78-1.08)	0.34	0.92 (0.82-1.04)
Total milk					
n cases/n total	269/1568	165/1017	331/2306		
Median, servings/d	0.8	0.8	1.5		
Model 1	1 (ref)	1.05 (0.87-1.25)	0.89 (0.76-1.03)	0.04*	0.94 (0.86-1.03)
Model 2	1 (ref)	1.05 (0.88-1.25)	0.90 (0.78-1.05)	0.08	0.95 (0.87-1.04)
Model 3	1 (ref)	1.06 (0.89-1.27)	0.94 (0.81-1.10)	0.22	0.98 (0.89-1.08)
Model 4	1 (ref)	1.01 (0.85-1.21)	0.92 (0.79-1.07)	0.17	0.97 (0.88-1.06)
High-fat milk					
n cases/n total	502/3080	153/923	110/888		
Median, servings/d	0.0	0.8	1.5		
Model 1	1 (ref)	1.02 (0.87-1.21)	0.77 (0.64-0.94)	0.03*	0.88 (0.78-0.99)*
Model 2	1 (ref)	1.00 (0.85-1.18)	0.77 (0.63-0.94)	0.02*	0.88 (0.79-0.97)*
Model 3	1 (ref)	0.99 (0.84-1.17)	0.79 (0.65-0.97)	0.05*	0.89 (0.80-0.99)*
Model 4	1 (ref)	1.03 (0.88-1.22)	0.83 (0.69-1.01)	0.15	0.91 (0.82-1.01)
Low-fat milk					
n cases/n total	312/2192	242/1342	211/1357		
Median, servings/d	0.0	0.8	1.5		
Model 1	1 (ref)	1.27 (1.09-1.48)	1.10 (0.94-1.29)	0.11	1.04 (0.97-1.13)
Model 2	1 (ref)	1.28 (1.10-1.49)	1.13 (0.96-1.33)	0.06	1.06 (0.98-1.14)
Model 3	1 (ref)	1.27 (1.09-1.49)	1.15 (0.97-1.35)	0.04*	1.07 (0.99-1.16)
Model 4	1 (ref)	1.20 (1.03-1.39)	1.08 (0.92-1.27)	0.20	1.04 (0.96-1.13)

Table 3. The associations of dairy intake and prediabetes risk in the AusDiab study (*n* = 4,891). (continued)

	Relative risk (95% CI) across intake range categories[1]				Continuous[2]
	T1	T2	T3	P_{trend}	RR (95%CI)
Total yogurt					
n cases/n total	219/1195	220/1537	326/2159		
Median, servings/d	0.0	0.01	0.36		
Model 1	1 (ref)	0.91 (0.76-1.08)	0.94 (0.80-1.10)	0.83	1.05 (0.84-1.32)
Model 2	1 (ref)	0.92 (0.77-1.10)	0.97 (0.82-1.15)	0.89	1.10 (0.87-1.38)
Model 3	1 (ref)	0.93 (0.78-1.11)	0.99 (0.84-1.17)	0.69	1.14 (0.90-1.43)
Model 4	1 (ref)	0.93 (0.78-1.10)	1.01 (0.85-1.19)	0.50	1.20 (0.95-1.51)
Total cheese					
n cases/n total	278/1583	297/1894	190/1414		
Median, servings/d	0.05	0.3	0.7		
Model 1	1 (ref)	0.91 (0.79-1.06)	0.81 (0.68-0.97)	0.02*	0.69 (0.53-0.91)**
Model 2	1 (ref)	0.92 (0.79-1.06)	0.82 (0.69-0.97)	0.02*	0.69 (0.53-0.91)**
Model 3	1 (ref)	0.94 (0.81-1.09)	0.85 (0.72-1.01)	0.07	0.74 (0.56-0.96)*
Model 4	1 (ref)	0.93 (0.81-1.08)	0.85 (0.72-1.01)	0.07	0.74 (0.57-0.96)*
High-fat cheese					
n cases/n total	293/1656	332/2218	140/1017		
Median, servings/d	0.0	0.2	0.7		
Model 1	1 (ref)	0.86 (0.74-0.99)	0.81 (0.67-0.98)	0.04*	0.72 (0.54-0.95)*
Model 2	1 (ref)	0.85 (0.74-0.99)	0.81 (0.67-0.98)	0.04*	0.71 (0.54-0.95)*
Model 3	1 (ref)	0.87 (0.75-1.01)	0.84 (0.70-1.02)	0.11	0.75 (0.57-1.00)
Model 4	1 (ref)	0.85 (0.74-0.98)	0.85 (0.70-1.02)	0.12	0.78 (0.59-1.03)
Low-fat cheese					
n cases/n total	591/3726	71/457	103/708		
Median, servings/d	0.0	0.1	0.4		
Model 1	1 (ref)	1.02 (0.81-1.29)	0.97 (0.80-1.18)	0.77	0.85 (0.59-1.22)
Model 2	1 (ref)	1.02 (0.82-1.29)	0.98 (0.81-1.19)	0.87	0.86 (0.60-1.22)
Model 3	1 (ref)	1.02 (0.81-1.28)	0.99 (0.82-1.20)	0.94	0.87 (0.61-1.25)
Model 4	1 (ref)	1.00 (0.80-1.25)	0.96 (0.80-1.17)	0.71	0.84 (0.58-1.20)
Ice cream					
n cases/n total	273/1676	242/1581	250/1634		
Median, servings/d	0.02	0.11	0.5		
Model 1	1 (ref)	0.95 (0.81-1.11)	0.87 (0.74-1.04)	0.14	1.02 (0.86-1.20)
Model 2	1 (ref)	0.95 (0.81-1.12)	0.87 (0.74-1.04)	0.13	1.02 (0.86-1.20)
Model 3	1 (ref)	0.95 (0.81-1.11)	0.88 (0.75-1.05)	0.18	1.03 (0.87-1.21)
Model 4	1 (ref)	0.94 (0.80-1.10)	0.88 (0.74-1.04)	0.16	1.03 (0.88-1.21)

[1] Relative risks (95CIs) were estimated across four categories split by tertile values (T1 to T3) or non-consumers + median categories with the lowest category as the reference, adjusted for covariates as follows: Model 1 included age (continuous) and sex and energy intake (continuous). Model 2 was additionally adjusted for education (3 categories), smoking (3 categories), physical activity (3 categories), alcohol consumption (continuous), and a family history of diabetes (yes/no). Model 3 was additionally adjusted for food groups associated with T2D, including intakes of fruit, vegetables, grains, legumes, nuts, red and processed meat, and fruit juice (continuous). Model 4 included baseline WC, change in WC from baseline to follow-up, LDL-cholesterol, and hypertension. Linear trend across intake range categories was assessed by including median values of each category as a continuous variable in the model.

[2] Relative risks per 1 serving/day (see definition in Table 1) were estimated. P-value significance level: *0.05, **0.01, ***0.001.

Abbreviations: CI, Confidence Interval; Q, Quartile.

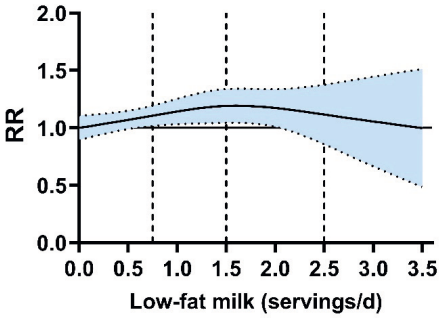


Figure 2. Non-linear association between low-fat milk intake and prediabetes risk in the AusDiab study (p for non-linearity = 0.04). The solid line indicates risk estimate fitted with a restricted cubic spline regression with 3 knots specified at the 5th, 50th, and 95th percentile of low-fat milk intake as indicated by the dashed vertical lines. The coloured area indicates the 95% confidence interval around the relative risk (RR). The model was adjusted for age, sex, energy intake, education, smoking, physical activity, alcohol consumption, a family history of diabetes, fruit, vegetables, grains, legumes, nuts, red and processed meat, and fruit juice intake.

Sensitivity analyses

None of the interactions of types of dairy food with sex, age, and baseline WC were statistically significant. The associations for the intake of high-fat dairy and high-fat milk were similar but no longer statistically significant. All associations were similar after adjustment for intake of all other dairy types, after excluding participants with prevalent CVD and with dairy types adjusted for energy intake (**Supplemental Table 3**). Using the ADA cutoffs resulted in attenuation of the associations for high-fat dairy and high-fat milk but not for total cheese.

Discussion

In this large, prospective Australian cohort, high-fat dairy, high-fat milk, and total cheese were associated with a lower incidence of prediabetes. These associations were independent of age, sex, energy intake, educational level, smoking status, physical activity, alcohol consumption, a family history of diabetes, and background dietary intake. By contrast, a nonlinear association for low-fat milk intake was found; risk was highest at 1.5 servings/day, with a decreasing risk at lower and higher intakes. Total dairy foods, fermented dairy, and high-fat cheese were associated with a lower incidence of prediabetes, but not when considering dietary intake of other food groups. Low-fat dairy foods, total milk, total yogurt, low-fat cheese, and ice cream were not associated with risk of prediabetes in multivariable models.

Results in context

Dairy should not be regarded as one single product but as a heterogeneous group of foods because associations with (markers of) disease risk vary by the product type [43]. Furthermore, our results showed that distinguishing the fat content is important for the assessment of the health effects of dairy. For instance, we found a protective association for high-fat milk and prediabetes and a nonlinear positive association for low-fat milk. These associations were in accordance with those of the Dutch Rotterdam studies, where we found significant associations for high-fat milk ($HR_{serving/day}$ 0.88, 95% CI 0.79-0.99) and in the other direction for low-fat milk ($HR_{serving/day}$ 1.07, 95% CI 1.01, 1.13) [17]. However, low-fat milk and high-fat milk were not associated with prediabetes in the Dutch Hoorn Studies [16]. In the FHS-OC, no contrast in associations by the fat content of milk was found; they reported significant dose-response associations for both skimmed milk (HR 2.14 versus 0-1 servings/wk, 0.82, 95%CI 0.61-1.10) and whole milk (HR 2.1 versus 0 servings/wk, 0.84, 95%CI 0.69-1.01) [15].

A higher intake of total and high-fat cheese was associated with a lower prediabetes risk, which is consistent with the Hoorn Studies (total cheese, RR top versus bottom quartile: 0.86, 95%CI 0.73-1.02), high-fat cheese, $RR_{serving/day}$ 0.94, 95%CI 0.88-1.00; and RR top versus bottom quartile: 0.79, 95%CI 0.66-0.94) [16] and the FHS-OC (total cheese, HR 2-4 versus 0-1 servings/wk, 0.86, 95%CI 0.69-1.07) [15] but not with the Rotterdam Study [17]. In addition, this association was reported in several cross-sectional studies with prediabetes outcomes [44, 45] and meta-analyses of prospective cohort studies on cheese and T2D [46].

Many guidelines worldwide recommend low-fat dairy types to limit the intake of saturated fat [47]. However, there is currently little evidence that high-fat dairy regardless of its high saturated fat content is harmful for health [48-50]. Differences in nutrient content are marginal: for example, high-fat milk contains 3.5% fat, semi skimmed milk contains 1%-1.5% fat, and skimmed milk contains no more than 0.15% fat, but high-fat milk contains higher concentrations of fat-soluble vitamin A (36, 15, and 1 µg, respectively) [29] and vitamin K (1.4, 0.7, and 0 µg, respectively) [51]. It is unlikely that these nutritional differences completely accounted for the sign reversal of our observed associations for milk. Possibly, low-fat foods have a lower satiety value, which could result in overconsumption of carbohydrates, particularly harmful to health if fats are substituted for refined starches and sugar [52, 53]. Furthermore, consumers may prefer sweetened low-fat milk to compensate for the reduced flavour by removing fat globules and cream, subsequently increasing sugar intake. Indeed, in our study, the diet of participants in the top compared with those in the bottom tertiles of low-fat milk and low-fat cheese contained higher proportions of carbohydrates. Compared with high glycaemic carbohydrates, saturated fat increases HDL-cholesterol levels, resulting in similar total cholesterol/HDL-cholesterol and lower triglyceride levels [54]. In addition, the effects of saturated fatty acids depend on the type. Dairy contains palmitic acid

associated with increased T2D risk but also various fatty acids with potential opposite effects. A meta-analysis of 16 studies showed that higher levels of odd-chained saturated fatty acids (C15:0 and C17:0), and natural ruminant trans-fats [t16:1(n-7)] were associated with a lower T2D risk [55].

We did not find that yogurt intake was associated with prediabetes in our study, in line with previous prospective studies [15, 16]. Only in the Rotterdam Study, we previously found that high-fat yogurt was associated with a lower prediabetes risk, but low-fat yogurt showed neutral associations [17]. A nonlinear inverse association was found between yogurt intake and T2D (at 80 g/day, RR: 0.86, 95%CI 0.83-0.90), with no additional benefit at higher intake levels [56]. Furthermore, in 3 large cohorts among US individuals and in the Iranian Tehran Lipid and Glucose study, increased intake of yogurt during the follow-up was associated with a lower risk of T2D [57, 58]. Compared with US and Western-European cohorts, intake levels of yogurt were considerably lower in our current study (0.19 ± 0.28 servings/d), which might explain the neutral as- sociation we observed. Low intake levels in this sample were in line with the Australian Health Survey (0.12 servings/d) [31]. Furthermore, in Australia, 76% of yogurt was flavoured or had added fruit, and 46% was of low fat. The FFQs used in many population-based studies do not consider the variety in sugar, protein, and fat content of commercially available yogurts [59, 60]. Moreover, many consumers add sweeteners such as sugar, jam, or honey to plain yogurt [61]. Unmeasured differences in nutrient content of consumed yogurts in each cohort might contribute to inconsistent findings. Future studies should collect more detailed information on yogurt composition and consumer behaviour to elucidate potential heterogeneity.

Our results are not in line with population-based studies showing inverse associations for low-fat dairy and yogurt intake with T2D [46]. Furthermore, our findings are not in line with the previous analysis of the AusDiab study, which found an inverse association between low-fat milk intake and T2D incidence at the 5 y follow-up (OR: 0.65, 95%CI 0.44, 0.94) and a nonsignificant association for full-fat milk (OR: 1.18, 95%CI 0.78, 1.79) [18]. Nevertheless, associations between yogurt and cheese were similar. We excluded approximately 2,000 participants with prediabetes at baseline, and thus, these different findings may result from a baseline sample with less variation in the glycaemic status compared with studies on T2D. It could be that dairy has differential effects according to individual's metabolic state or degree of insulin resistance [15], and more research is needed in that regard.

Current evidence from long-term RCTs with dairy consumption as the main intervention and diabetes related outcomes is inconclusive owing to differences in design, duration, and geographic location [62]. Results are further affected by the diet consumed parallel or in replacement of high dairy diets, physical activity levels, and weight variation during the study. Two recent RCTs studying the effects of diets high in dairy did not

find differences in glucose measurements after the intervention [63, 64]. However, one of these RCTs showed that insulin sensitivity was decreased in the high dairy diets compared with that in the low dairy control, possibly because of the insulinotropic effect of dairy and alterations in the gut microbiota [63]. Short-term controlled feeding trials showed that milk proteins attenuated acute hyperglycaemia [65] and regulated lipid changes induced by glucose ingestion [66]. Future studies are required to detangle the role of different dairy types and dairy fat content in T2D development, particularly RCTs.

Strengths and limitations

This study was performed within a large population-based study with up to 12 y of follow-up. This study adds to the body of evidence by differentiating associations by the fat content of dairy types. Prediabetes was defined both based on FPG and 2hPG, representing impaired fasting glucose and impaired glucose tolerance, respectively, 2 distinct states of T2D development [1]. The results should be interpreted carefully by considering the following limitations. First, we used only baseline data on dairy intake, and people might have changed their diet over time, resulting in misclassification of the exposure, biasing estimates toward the null. However, in our previous analyses of the Rotterdam Study, we showed that the inclusion of repeated measurements of dairy consumption did not change associations [17]. Second, the FFQ is useful in estimating the intake of frequently used foods such as dairy and ranking participants according to their food intake in observational studies; however, the FFQ relies on participants' memory and ability to estimate portion sizes and is, therefore, prone to recall bias. Thus, misclassification of exposure is possible, resulting in bias toward the null. Third, reverse causality might be an issue. Prediabetes is commonly an asymptomatic condition, making reverse causality due to a prediabetes diagnosis unlikely. Moreover, the exclusion of people with CVD at the baseline or follow-up in sensitivity analyses did not change the estimates. However, presence of obesity or other risk factors might have induced behavioural changes, such as a shift from high-fat to low-fat dairy consumption to reduce fat and caloric intake. Fourth, residual confounding cannot be ruled out considering the observational nature of our study; nevertheless, we carefully adjusted for a wide range of confounders, such as background diet and CVD risk factors. Finally, the AusDiab is not entirely representable of the general population, as socioeconomic status of responders was somewhat higher than of non-responders, and some healthy volunteer selection bias is likely [19]. Furthermore, there was a considerable loss to the follow-up in the AusDiab study.

Conclusions

In conclusion, in the long-term, population-based Australian cohort, associations of dairy and prediabetes differ by both type and fat content. High-fat dairy foods, high-fat milk, and total cheese were associated with a lower prediabetes incidence. Further prospective studies should collect more specific information on fat and sugar content of various milk and yogurt types and examine the influence of reverse causation.

List of supplementary materials chapter 4

Supplemental table 1. Missing values of covariates in participants before imputation.

Supplemental table 2. Baseline characteristics of participants in the AusDiab study across different population subgroups (*n* = 4,891).

Supplemental table 3. Sensitivity analyses of associations of dairy product types and prediabetes risk in the Ausdiab study (*n* = 4,891).

Scan this QR code to download the supplementary materials.

References

1. World Health Organization (WHO), *Definition and diagnosis of diabetes mellitus and intermediate hyperglycaemia: report of a WHO/IDF consultation.* 2006.
2. Gerstein, H.C., et al., *Annual incidence and relative risk of diabetes in people with various categories of dysglycemia: a systematic overview and meta-analysis of prospective studies.* Diabetes Res Clin Pract, 2007. **78**(3): p. 305-12.
3. Yeboah, J., et al., *Impaired fasting glucose and the risk of incident diabetes mellitus and cardiovascular events in an adult population: MESA (Multi-Ethnic Study of Atherosclerosis).* Journal of the American College of Cardiology, 2011. **58**(2): p. 140-146.
4. Mutie, P.M., et al., *An investigation of causal relationships between prediabetes and vascular complications.* Nat Commun, 2020. **11**(1): p. 4592.
5. Cai, X., et al., *Association between prediabetes and risk of all cause mortality and cardiovascular disease: updated meta-analysis.* Bmj, 2020. **370**: p. m2297.
6. Danaei, G., et al., *National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2·7 million participants.* Lancet, 2011. **378**(9785): p. 31-40.
7. *Effects of diabetes definition on global surveillance of diabetes prevalence and diagnosis: a pooled analysis of 96 population-based studies with 331,288 participants.* Lancet Diabetes Endocrinol, 2015. **3**(8): p. 624-37.
8. Liu, C., et al., *Trends in Self-reported Prediabetes and Metformin Use in the USA: NHANES 2005-2014.* J Gen Intern Med, 2020. **35**(1): p. 95-101.
9. Richter, B., et al., *Development of type 2 diabetes mellitus in people with intermediate hyperglycaemia.* Cochrane Database of Systematic Reviews, 2018(10).
10. Perreault, L., et al., *Effect of regression from prediabetes to normal glucose regulation on long-term reduction in diabetes risk: results from the Diabetes Prevention Program Outcomes Study.* The Lancet, 2012. **379**(9833): p. 2243-2251.
11. Glechner, A., et al., *Effects of lifestyle changes on adults with prediabetes: A systematic review and meta-analysis.* Primary Care Diabetes, 2018. **12**(5): p. 393-408.
12. Galaviz, K.I., et al., *Interventions for Reversing Prediabetes: A Systematic Review and Meta-Analysis.* American Journal of Preventive Medicine, 2022. **62**(4): p. 614-625.
13. Yu, E. and F.B. Hu, *Dairy Products, Dairy Fatty Acids, and the Prevention of Cardiometabolic Disease: a Review of Recent Evidence.* Curr Atheroscler Rep, 2018. **20**(5): p. 24.
14. Weaver, C.M., *Dairy matrix: is the whole greater than the sum of the parts?* Nutr Rev, 2021. **79**(Suppl 2): p. 4-15.
15. Hruby, A., et al., *Associations of Dairy Intake with Incident Prediabetes or Diabetes in Middle-Aged Adults Vary by Both Dairy Type and Glycemic Status.* J Nutr, 2017. **147**(9): p. 1764-1775.
16. Slurink, I.A.L., et al., *Dairy product consumption and incident prediabetes in Dutch middle-aged adults: the Hoorn Studies prospective cohort.* Eur J Nutr, 2022. **61**(1): p. 183-196.
17. Slurink, I.A.L., et al., *Dairy Product Consumption in Relation to Incident Prediabetes and Longitudinal Insulin Resistance in the Rotterdam Study.* Nutrients, 2022. **14**(3).
18. Grantham, N.M., et al., *The association between dairy food intake and the incidence of diabetes in Australia: the Australian Diabetes Obesity and Lifestyle Study (AusDiab).* Public Health Nutr, 2013. **16**(2): p. 339-45.
19. Dunstan, D.W., et al., *The Australian Diabetes, Obesity and Lifestyle Study (AusDiab)--methods and response rates.* Diabetes Res Clin Pract, 2002. **57**(2): p. 119-29.
20. Magliano, D.J., et al., *Glucose indices, health behaviors, and incidence of diabetes in Australia: the Australian Diabetes, Obesity and Lifestyle Study.* Diabetes Care, 2008. **31**(2): p. 267-72.
21. Tanamas, S.K., *The Australian diabetes, obesity and lifestyle study.* 2013.
22. Barr, E., et al., *AusDiab 2005: The Australian Diabetes.* Obesity and Lifestyle Study Melbourne: International Diabetes Institute, 2005.
23. Rhee, J.J., et al., *Comparison of methods to account for implausible reporting of energy intake in epidemiologic studies.* American journal of epidemiology, 2015. **181**(4): p. 225-233.

24. Willett, W.C., *Nutritional Epidemiology*. 1998, New York: Oxford University Press.
25. Dunstan, D.W., et al., *The Rising Prevalence of Diabetes and Impaired Glucose Tolerance: The Australian Diabetes, Obesity and Lifestyle Study.* Diabetes Care, 2002. **25**(5): p. 829-834.
26. Association, A.D., *2. Classification and Diagnosis of Diabetes.* Diabetes Care, 2016. **40**(Supplement_1): p. S11-S24.
27. Hodge, A., et al., *The Anti Cancer Council of Victoria FFQ: relative validity of nutrient intakes compared with weighed food records in young to middle-aged women in a study of iron supplementation.* Aust N Z J Public Health, 2000. **24**(6): p. 576-83.
28. McNaughton, S.A., et al., *An index of diet and eating patterns is a valid measure of diet quality in an Australian population.* J Nutr, 2008. **138**(1): p. 86-93.
29. Lewis, J.M.G.C.H.A.N.F.A., *NUTTAB 95 : nutrient data table for use in Australia*. 1995, Commonwealth of Australia: [Canberra].
30. Council., A.G.N.H.a.M.R. *Serve sizes*. 2015 [cited 2021 01-09]; Available from: eatforhealth.gov.au.
31. Australian Bureau of Statistics. *Australian Health Survey: Nutrition First Results-Foods and Nutrients. 2011-12 financial year.* 2014 [cited 2022 28-07]; Available from: https://www.abs.gov.au/statistics/health/health-conditions-and-risks/australian-health-survey-nutrition-first-results-foods-and-nutrients/2011-12.
32. Australian Institute of Health and Welfare, *Standard Questions on the Use of Tobacco Among Adults. Canberra: AIHW.* 1998.
33. Australian Institute of Health and Welfare, *The Active Australia Survey: A Guide and Manual for Implementation, Analysis and Reporting. Canberra, Australia, Australia Institute of Health and Welfare.* 2003.
34. Brown, W.J., et al., *Test-retest reliability of four physical activity measures used in population surveys.* J Sci Med Sport, 2004. **7**(2): p. 205-15.
35. Armstrong, T., A. Bauman, and J. Davies, *Physical Activity Patterns of Australian Adults: Results of the 1999 National Physical Activity Survey. Canberra, Australia, Australian Institute of Health and Welfare, 2000 (AIHW cat. no. CVD 10).*
36. Cameron, A.J., et al., *Overweight and obesity in Australia: the 1999-2000 Australian diabetes, obesity and lifestyle study (AusDiab).* Medical journal of Australia, 2003. **178**(9): p. 427-432.
37. Briganti, E.M., et al., *Untreated hypertension among Australian adults: the 1999-2000 Australian Diabetes, Obesity and Lifestyle Study (AusDiab).* Med J Aust, 2003. **179**(3): p. 135-9.
38. Knol, M.J., et al., *Overestimation of risk ratios by odds ratios in trials and cohort studies: alternatives to logistic regression.* CMAJ, 2012. **184**(8): p. 895-9.
39. Rothman, K.J., *No adjustments are needed for multiple comparisons.* Epidemiology, 1990. **1**(1): p. 43-6.
40. Soedamah-Muthu, S.S. and J. de Goede, *Dairy Consumption and Cardiometabolic Diseases: Systematic Review and Updated Meta-Analyses of Prospective Cohort Studies.* Curr Nutr Rep, 2018. **7**(4): p. 171-182.
41. Schwingshackl, L., et al., *Food groups and risk of type 2 diabetes mellitus: a systematic review and meta-analysis of prospective studies.* Eur J Epidemiol, 2017. **32**(5): p. 363-375.
42. Dutch Health Council (Gezondheidsraad), *Dutch dietary guidelines 2015 (Richtlijnen goede voeding 2015). Publication nr. 2015/24. ISBN 978-94-6281-089-1. The Hague.* 2015.
43. Thorning, T.K., et al., *Whole dairy matrix or single nutrients in assessment of health effects: current evidence and knowledge gaps.* Am J Clin Nutr, 2017. **105**(5): p. 1033-1045.
44. Drehmer, M., 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): p. 775-82.
45. Eussen, S.J., et al., *Consumption of dairy foods in relation to impaired glucose metabolism and type 2 diabetes mellitus: the Maastricht Study.* Br J Nutr, 2016. **115**(8): p. 1453-61.
46. Alvarez-Bueno, C., et al., *Effects of Milk and Dairy Product Consumption on Type 2 Diabetes: Overview of Systematic Reviews and Meta-Analyses.* Adv Nutr, 2019. **10**(suppl_2): p. S154-S163.
47. Comerford, K.B., et al., *Global Review of Dairy Recommendations in Food-Based Dietary Guidelines.* Front Nutr, 2021. **8**: p. 671999.

48. Mozaffarian, D. and J.H.Y. Wu, *Flavonoids, Dairy Foods, and Cardiovascular and Metabolic Health: A Review of Emerging Biologic Pathways.* Circ Res, 2018. **122**(2): p. 369-384.
49. Drehmer, M., et al., *Total and Full-Fat, but Not Low-Fat, Dairy Product Intakes are Inversely Associated with Metabolic Syndrome in Adults.* J Nutr, 2016. **146**(1): p. 81-9.
50. Drouin-Chartier, J.P., et al., *Systematic Review of the Association between Dairy Product Consumption and Risk of Cardiovascular-Related Clinical Outcomes.* Adv Nutr, 2016. **7**(6): p. 1026-1040.
51. The Dutch National Institute for Public Health and the Environment (RIVM), *NEVO-Table. Dutch Food Composition Database 2021/7.0.*, Dutch Nutrition Centre, Editor. 2021: Den Haag.
52. Ludwig, D.S. and W.C. Willett, *Three daily servings of reduced-fat milk: an evidence-based recommendation?* JAMA Pediatr, 2013. **167**(9): p. 788-9.
53. Ardisson Korat, A.V., et al., *Dairy fat intake and risk of type 2 diabetes in 3 cohorts of US men and women.* The American journal of clinical nutrition, 2019. **110**(5): p. 1192-1200.
54. Micha, R. and D. Mozaffarian, *Saturated fat and cardiometabolic risk factors, coronary heart disease, stroke, and diabetes: a fresh look at the evidence.* Lipids, 2010. **45**(10): p. 893-905.
55. Imamura, F., 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. **15**(10): p. e1002670.
56. Gijsbers, L., et al., *Consumption of dairy foods and diabetes incidence: a dose-response meta-analysis of observational studies.* Am J Clin Nutr, 2016. **103**(4): p. 1111-24.
57. Drouin-Chartier, J.P., et al., *Changes in dairy product consumption and risk of type 2 diabetes: results from 3 large prospective cohorts of US men and women.* Am J Clin Nutr, 2019. **110**(5): p. 1201-1212.
58. Yuzbashian, E., et al., *Changes in dairy product consumption and subsequent type 2 diabetes among individuals with prediabetes: Tehran Lipid and Glucose Study.* Nutrition Journal, 2021. **20**(1): p. 88.
59. Rybicka, I. and A. Gliszczyńska-Świgło, *Sugars in dairy products of different flavours.* International Dairy Journal, 2021. **114**: p. 104933.
60. Moore, J.B., A. Horti, and B.A. Fielding, *Evaluation of the nutrient content of yogurts: a comprehensive survey of yogurt products in the major UK supermarkets.* BMJ Open, 2018. **8**(8): p. e021387.
61. Saint-Eve, A., et al., *How much sugar do consumers add to plain yogurts? Insights from a study examining French consumer behavior and self-reported habits.* Appetite, 2016. **99**: p. 277-284.
62. O'Connor, S., et al., *Increased dairy product intake alters serum metabolite profiles in subjects at risk of developing type 2 diabetes.* Molecular nutrition & food research, 2019. **63**(19): p. 1900126.
63. Schmidt, K.A., et al., *Impact of low-fat and full-fat dairy foods on fasting lipid profile and blood pressure: exploratory endpoints of a randomized controlled trial.* The American Journal of Clinical Nutrition, 2021. **114**(3): p. 882-892.
64. Eelderink, C., et al., *The effect of high compared with low dairy consumption on glucose metabolism, insulin sensitivity, and metabolic flexibility in overweight adults: a randomized crossover trial.* The American journal of clinical nutrition, 2019. **109**(6): p. 1555-1568.
65. Fekete Á, A., D.I. Givens, and J.A. Lovegrove, *Can milk proteins be a useful tool in the management of cardiometabolic health? An updated review of human intervention trials.* Proc Nutr Soc, 2016. **75**(3): p. 328-41.
66. Chen, L., et al., *Dairy Milk Casein and Whey Proteins Differentially Alter the Postprandial Lipidome in Persons with Prediabetes: A Comparative Lipidomics Study.* Journal of Agricultural and Food Chemistry, 2022. **70**(33): p. 10209-10220.

Chapter 5

Dairy intake and prediabetes risk in the Lifelines study

Manuscript based on this chapter:

Isabel A.L. Slurink, Eva Corpeleijn, Stephan J. Bakker, Joran Jongerling, Nina Kupper, Tom Smeets & Sabita S. Soedamah-Muthu, (2023). *Dairy consumption and incident prediabetes: prospective associations and network models in the large population-based Lifelines study.* The American journal of clinical nutrition, 118(6), 1077-1090.

Abstract

Background

Evidence on associations between dairy consumption and incident prediabetes is inconsistent. One potential explanation for heterogeneity is that health behaviour and food intake covary with the consumption of various high-fat and low-fat dairy types. The objective was to investigate the associations of total dairy and dairy types with incident prediabetes and to assess how dairy intake is linked with metabolic risk factors, health behaviours, and foods, as potential explanations for these associations.

Methods

Overall, 74,132 participants from the prospective population-based Lifelines study were included (mean age, 45.5 ± 12.3 y; 59.7% female). Baseline dairy intake was measured using a validated food frequency questionnaire. Prediabetes at follow-up was defined based on the World Health Organization/International Expert Committee criteria as fasting plasma glucose of 110–125 mg/dL or glycated haemoglobin concentrations of 6.0%–6.5%. Associations were analysed using Poisson regression models adjusted for social demographics, health behaviours, family history of diabetes, and food group intake. Interconnections were assessed with mixed graphical model networks.

Results

At a mean follow-up of 4.1 ± 1.1 years, 2746 participants developed prediabetes (3.7%). In regression analyses, neutral associations were found for most dairy types. Intake of plain milk and low-fat milk were associated with a higher risk of prediabetes in the top compared with bottom quartiles (RR 1.17; 95%CI 1.05-1.30; P_{trend} = 0.04 and RR 1.18; 95%CI 1.06-1.31; P_{trend} = 0.01). Strong but non-significant effect estimates for high-fat yogurt in relation to prediabetes were found ($RR_{serving/day}$ 0.80; 95%CI 0.64-1.01). The network analysis showed that low-fat milk clustered with energy-dense foods, including bread, meat, and high-fat cheese, whereas high-fat yogurt had no clear link with health risk factors and food intake.

Conclusions

In this large cohort of Dutch adults, low-fat milk intake was associated with higher prediabetes risk. Heterogeneous associations by dairy type and fat content might partly be attributed to confounding caused by behaviours and food intake related to dairy intake.

Introduction

Prediabetes, or intermediate hyperglycaemia, is characterized by blood glucose levels above the normal range but below the diagnostic threshold for diabetes [1]. Prediabetes increases the risk of type 2 diabetes (T2D) and cardiovascular disease (CVD) [2-5], emphasizing the need to identify potentially modifiable risk factors, including diet, that could prevent this condition. Dairy products are widely consumed and contain nutrients with beneficial health effects, including minerals and vitamins, but also contain nutrients linked to adverse cardiometabolic health, including saturated fat, added sugars, and sodium [6]. Observational studies regarding the relationship between dairy intake and prediabetes reported inconsistent results [7-13], calling for a better understanding of these associations and exploration of potential sources of heterogeneity.

One potential explanation for heterogeneity is that health behaviour and food intake may covary with consumption of different types of dairy intake. For example, yogurt consumption may relate to a relatively healthy diet and healthy lifestyle behaviour [14-17], and residual confounding because of unmeasured or incomplete adjustment of this healthier diet and lifestyle behaviour may explain the inverse associations found with prediabetes [10, 12]. As such, the association between one food and prediabetes relies on that foods' covariation with health behaviours or other foods with potentially different health effects. Multivariable regression models allow for the estimation of a hypothesis-driven independent effect of an exposure on an outcome. In addition to this traditional reductionism method, a holistic network modelling approach may aid the interpretation of regression models by accounting for this interrelatedness of risk factors. In network analyses, the conditional interdependencies between all exposures and confounding factors are analysed and visualized at once [18]. Network models fit with the multifactorial aspect of prediabetes because it describes clusters of risk factors and dietary patterns as potential causes. In the field of nutritional epidemiology, network modelling has been used to identify dietary and meal-specific patterns [19-22] and to describe relations between demographics, dietary behaviours, and clinical markers [23].

The current study aimed to investigate prospective associations of total dairy and dairy types with prediabetes in the Lifelines Cohort. We additionally aimed to assess how dairy intake is linked with metabolic risk factors, health behaviours, and food intake as potential factors underlying associations of total dairy and dairy types with prediabetes. We hypothesize that risk of prediabetes at baseline may be related to intake of specific dairy types, as potential awareness of individual's risk might relate to certain dietary choices, for example, a preference for low-fat dairy types to adhere to dietary guidelines [24]. Furthermore, we hypothesize that the concept that dairy intake is part of a network of metabolic risk factors, individual health behaviours and intake of other food groups may explain the heterogeneity in previous findings.

Methods

Design and participants

Lifelines is a multidisciplinary prospective population-based cohort study examining in a unique 3-generation design the health and health- related behaviours of 167,729 persons living in the North of the Netherlands [25]. It employs a broad range of investigative procedures in assessing the biomedical, sociodemographic, behavioural, physical, and psychological factors that contribute to the health and disease of the general population, with a special focus on multimorbidity and complex genetics. Baseline measurements took place between 2006 and 2013. All participants were invited for the second assessment between 2014 and 2018 and will be invited for follow-up assessments every 5 years. Participants were recruited by their general practitioner, through family members, or by self-registration via the website. Exclusion criteria included having a severe psychiatric or physical illness, limited life expectancy (<5 years) and insufficient knowledge of the Dutch language to complete a Dutch questionnaire. The Lifelines study is conducted according to the principles of the Declaration of Helsinki and in accordance with the research code of the University Medical Centre Groningen. The Lifelines protocol was approved by the University Medical Centre Groningen Medical ethical committee under number 2007/152.

In total, 147,182 participants were aged >18 y and were not pregnant during baseline or follow-up (**Figure 1**). We excluded participants without dietary data at baseline or the food frequency questionnaire (FFQ) judged as unreliable by research dieticians, for example, owing to nutrient or food group reports below the possible under or upper limit (*n* = 8,662) or with implausible energy intakes (defined as men <800 or >4200 kcal and women <500 or >3,500 kcal) (*n* = 14,515) [26, 27]. We excluded participants with diabetes (*n* = 4,695) or prediabetes (*n* =10,400) at baseline or missing (pre-)diabetes information at baseline (*n* = 690). Exclusion of diabetes was based on self-reported diabetes, diabetes medication use, or a fasting plasma glucose (FPG) ≥126 mg/dL or glycated haemoglobin (HbA1c) concentration of ≥6.5% [1]. Of the participants, 1,714 (1.2% of 147,182) had a FPG of ≥126 mg/dL or HbA1c concentration of ≥6.5% without a self-report diagnosis or medication use for diabetes. Exclusion of prediabetes was based on FPG between 110 and 125 mg/dL or HbA1c concentrations between 6.0% and 6.4% [1, 28]. Furthermore, we excluded 27,710 participants without follow-up measurements, 5,272 participants with missing prediabetes information at follow-up, and 461 participants with diabetes at follow-up. A total of 74,132 participants were included in our analysis.

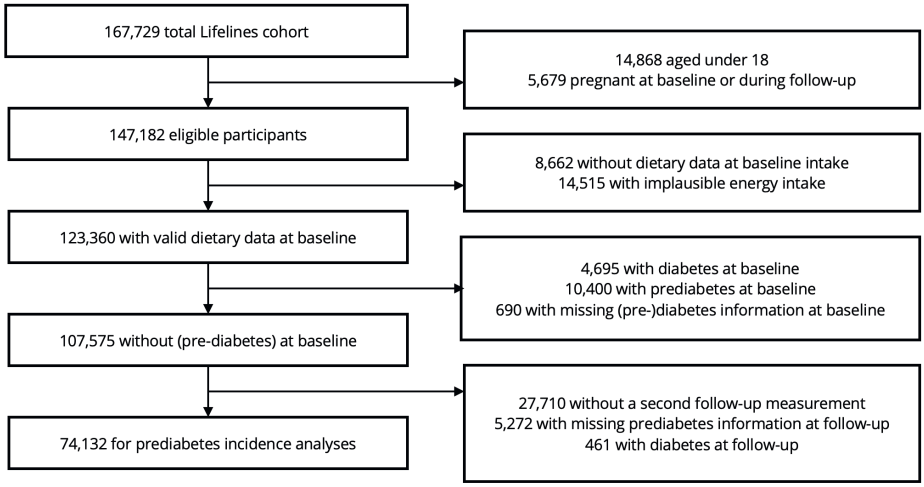


Figure 1. Flowchart for inclusion of participants for the present analysis of the Lifelines Cohort.

Dietary assessment

Dietary intake was assessed with the "Flower-FFQ," which contains 1 main questionnaire on energy and macronutrient intake (heart) assessed at baseline, and 4 complementary food questionnaires (petals) on micronutrients assessed during follow-up. Information about the development and validation of the Flower-FFQ has been previously published [29]. For this analysis, the baseline data of the flower heart were used, comprising 110 food items, including all major dairy types, bread, pasta, rice, potatoes, fruit, vegetables, legumes, meat, fish, coffee, tea, and soda/juice, selected based on the Dutch National Food Consumption Survey. Participants indicated the frequency of consumption in the past months, ranging from "never" to "6–7 d/wk." Portion sizes were estimated using natural portions and commonly used household measures. FFQ data were converted into nutrient intake by researchers from Wageningen University & Research using the Dutch Food Composition table 2011 (NEVO) [30, 31]. The Flower-FFQ was compared with a conventional FFQ in 2048 participants showing comparable mean absolute intake estimates for major food groups (Spearman correlations $r \geq 0.40$) and good ranking agreement (≥80% in the same or adjacent quartile) [29]. Furthermore, the Flower-FFQ was validated in 242 and 361 participants with data on urinary sodium and potassium excretions, showing acceptable ranking abilities of respectively 75% ($r = 0.40$) and 73% ($r = 0.37$) of participants in the same or adjacent quartile [29]. The Flower-FFQ does not differentiate between whole grain and refined cereal products. Whole meal and brown bread contribute to ~70% of bread consumption and is the largest contributor to whole grain intake in the Netherlands [32, 33]. Therefore, bread intake was used as a proxy for whole grain consumption in this study. Dairy types included total dairy, fermented dairy,

milk, yogurt, cheese, cream, and ice cream (**Table 1**). Each dairy category was further divided into low-fat (liquid products <2%, cheese <20%) and high-fat (liquid products >2%, cheese >20%). Intakes were expressed in servings/day enforcing equal water content of each type: milk, yogurt, cream, and ice cream: 150 mL; cheese, 20 g.

Table 1. Food items included in total dairy and dairy types and consumption in the Lifelines study (*n* = 74,132).

Dairy product[1]	Included dairy types	Consumers[2]	Intake (servings/day)[3]	
		%	Mean ± SD	Median [IQR]
Total dairy	All dairy products	99.9%	3.52 ± 1.83	3.26 [2.27, 4.45]
High-fat	All high-fat dairy products	99.6%	1.63 ± 1.33	1.31 [0.70, 2.20]
Low-fat	All low-fat dairy products	98.0%	1.88 ± 1.35	1.65 [0.93, 2.58]
Fermented dairy		98.9%	2.07 ± 1.47	1.79 [1.03, 2.78]
High-fat	High-fat yogurt, high-fat cheese, curd cheese	96.9%	1.30 ±1.22	0.98 [0.44, 1.82]
Low-fat	Low-fat yogurt, low-fat cheese, buttermilk	76.7%	0.78 ± 0.95	0.46 [0.04, 1.12]
Milk	Full fat milk, skimmed milk, semi skimmed milk, buttermilk, chocolate milk, coffee milk, plain milk in coffee	91.6%	1.12 ± 1.06	0.93 [0.32, 1.86]
Plain milk	Full fat milk, skimmed milk, semi skimmed milk	70.9%	0.66 ± 0.85	0.36 [0, 0.93]
High-fat	Full fat milk	9.8%	0.06 ± 0.28	0 [0, 0]
Low-fat	Skimmed milk, semi skimmed milk	66.8%	0.60 ± 0.82	0.29 [0, 0.93]
Yogurt		59.8%	0.23 ± 0.32	0.12 [0, 0.37]
High-fat	Full fat natural yogurt	20.4%	0.06 ± 0.17	0 [0, 0]
Low-fat	Skimmed yogurt, skimmed fruit yogurt	47.4%	0.18 ± 0.29	0 [0, 0.27]
Cheese		96.9%	1.50 ± 1.28	1.19 [0.64, 2.03]
High-fat	40+ (spreadable) cheese, 48+ (spreadable) cheese, cream cheese, foreign cheeses, cheese cubes, (cream) cheese on baguette and on pieces of toast, grated cheese, diced cheese, feta cheese, cheese fondue	94.6%	1.15 ± 1.18	0.82 [0.31, 1.60]
Low-fat	20+/30+ (spreadable) cheese	46.4%	0.36 ± 0.69	0 [0, 0.45]
Cream	Whip cream, coffee cream	71.9%	0.01 ± 0.01	0.01 [0, 0.01]
Ice cream	Milk-based ice cream	62.6%	0.04 ± 0.06	0.02 [0, 0.06]

[1] Low-fat (liquid products <2%, cheese <20%); high-fat (liquid products > 2%, cheese > 20%).
[2] Consumers were defined as consuming >0 servings/day of a specific dairy type.
[3] Intakes are expressed in servings/day enforcing equal water content of each type: milk, yogurt, cream, and ice cream: 150 mL; cheese, 20 g.

Prediabetes incidence

Fasting blood samples were collected at baseline and at the follow- up measurement. FPG in venous plasma was determined using the glucose hexokinase method. HbA1c was determined in whole blood (EDTA-anticoagulated) using the turbidimetric inhibition immuno-assay on a Cobas Integra 800 CTS analyser (Roche Diagnostics Netherland BV). This assay was standardized against the reference method of the International Federation of Clinical HbA1c and Laboratory Medicine and had a variation coefficient of 2.1% for a mean HbA1c of 5.5% and 1.9% for a mean HbA1c of 10.6% [34]. Prediabetes was defined according to the WHO as FPG concentrations between 110 and 125 mg/dL (6.1 and 6.9 mmol/L) and the International Expert Committee (IEC) as HbA1c levels between 42 and 46 mmol/mol (6.0 and 6.4%) [1, 28], similar to our previous publications on this topic [11-13]. Additional sensitivity analyses were performed with prediabetes defined by the American Diabetes Association (ADA) criteria, defined as an FPG between 100 mg/dL and 125 mg/dL (5.6 and 6.9 mmol/L) and/or HbA1c concentrations between 39 and 46 mmol/mol (5.7 and 6.4%) [35]. We presented the main results using the international WHO/IEC cutoffs because this is European data, and we aimed to assess dairy in relation to an intermediate stage at which individuals are at high risk of progression of T2D and CVD. Using ADA compared with the WHO/IEC cutoffs for prediabetes results in a higher number of prediabetes cases, which are generally healthier people. Consequently, this broader definition of prediabetes is associated with a lower incidence of T2D and CVD [36, 37].

Other variables

A self-administered baseline questionnaire included data on sociodemographic, health, and clinical factors. Educational level was assessed in 9 categories and subsequently categorized into primary, secondary, higher, or other education. Smoking status was categorized as current, former, or never. Physical activity was measured using the validated Short QUestionnaire to ASsess Health-enhancing physical activity (SQUASH) and expressed in moderate to vigorous leisure time and commuting physical activity in min/wk (nonoccupational) [38, 39]. The SQUASH is fairly reliable with a Spearman correlation for overall reproducibility of 0.58 in 50 adults compared with an activity monitor [39]. Alcohol consumption was categorized as non-drinker, ≤10, 10–30, and ≥30 g/day. A family history of diabetes was defined as having at least a parent, sibling, or child with diabetes. Physical measurements were performed by research assistants at 1 of the 12 Lifelines research sites located in the North of the Netherlands. BMI was calculated as weight divided by height squared (kg/m2). Waist circumference in centimetres was measured at the level midway between the lower rib margin and the iliac crest at the end of gentle expiration. Blood pressure in mmHg was measured automatically 10 times during 10 min in a lying position, and the final 3 readings were averaged. Medication use was assessed by questionnaires and by medication wrapper brought to the baseline visit. Hypertension was defined based on self-report or as systolic blood pressure ≥140 mmHg and/or diastolic blood pressure ≥90 mmHg and/or use of antihypertensive medication.

Total cholesterol and HDL cholesterol were assessed in serum using an enzymatic colorimetric method. LDL cholesterol was determined in serum with a colorimetric method. Serum triglyceride (TAG) concentrations were measured with a colorimetric UV method. At baseline and follow-up, participants were asked to indicate the presence of various cardiac and vascular diseases, including treatments and/or cancer. Baseline diabetes risk was calculated using the Prospective Cardiovascular Münster (PROCAM) risk algorithm, including major risk factors but no lifestyle or dietary components [40]. Participants indicated with one question if they desired to lose weight (yes/no).

Statistical analysis

Statistical analyses were performed using R version 4.1.2 (R Foundation for Statistical Computing). Descriptive data are reported as means and SDs for normally distributed continuous variables, medians, and interquartile ranges (IQRs) for skewed continuous variables, and frequencies and percentages for categorical variables.

Multivariable regression analyses

Poisson regression with a robust variance was used to examine the prospective associations between baseline dairy type intakes and prediabetes at the follow-up measurement [41]. This approach is suited to directly calculate a relative risk (RR) and 95% confidence intervals (CIs) for common binary outcome data [42, 43]. The dairy types were modelled both in quartiles of intake (reference lowest), as the FFQ is suited for ranking participants, and continuously in servings/day. Dairy products for which many participants reported no intake were divided into a non-consumer category (reference) and consumers in tertiles. Linear trends across intake range categories were assessed by including the median values of each category as a continuous variable in the model. For each model, we examined whether nonlinear terms of continuous dairy types (second-order polynomials or restricted cubic splines, excluding outliers) significantly improved model fit compared with the linear model assessed by likelihood ratio tests. A multiple imputation procedure (n = 10) was used to account for missing data on covariates (Supplemental Table 1). Regression coefficients for each of the imputed datasets were pooled.

Potential confounders were chosen based on the literature [6, 44, 45]. Multiple models were constructed to allow for the comparison of the associations estimated from the various models. Model 1 included age, sex, energy intake, and follow-up duration. Model 2 additionally adjusted for educational level, smoking behaviour, alcohol use, physical activity level, and family history of diabetes. Model 3 additionally adjusted for food groups associated with T2D, including fruit, vegetables, bread, legumes, nuts, red and processed meat, coffee, tea, and sugar-sweetened beverages (SSBs). Model 4 was additionally adjusted for potential mediating or otherwise confounding factors, including TAGs, LDL cholesterol, waist circumference, and hypertension.

Potential effect modification by age, sex, waist circumference, and educational level was explored, and associations were stratified in case of significance [46]. A series of sensitivity analyses were conducted using model 3. First, the independence of the associations of specific dairy types was evaluated by mutually adjusting for the intake of other dairy types. Second, the analysis was repeated excluding participants with self-reported CVD or cancer at baseline or follow-up (*n* = 12,977). Third, associations were calculated with energy-adjusted intake of dairy types in grams/day using the residual method [47]. Fourth, analyses were repeated using sex-specific cutoffs for quartiles of dairy type intake to account for sex differences in dairy and energy intake. Lastly, analyses were repeated using the ADA cutoff levels for prediabetes [35].

Next, we explored if there was a potential influence of participants who choose a certain dairy type because of higher risk of T2D, and in line with this, a desire to lose weight. For each dairy type, a multivariable linear regression model was performed with baseline T2D risk, as calculated with the PROCAM risk algorithm [40], or desire to lose weight (yes/no), as an independent variable, and intake of the dairy type as the dependent variable. Covariates from model 3 were used, except for energy intake as this was not considered a confounder. Furthermore, models were additionally adjusted for desire to lose weight (yes/no).

Network estimation

Two exploratory mixed graphical model (mgm) networks were estimated in participants with complete data (*n* = 67,206) using the mgm and qgraph packages [48, 49]. The first network included dairy types and food groups. A second, complete network additionally included incident prediabetes and baseline sociodemographic characteristics, health risk factors, and clinical markers. The primary objective of network analysis is to characterize conditional dependencies between pairs of variables in multivariate data [50, 51]. The network model used is a Mixed Graphical Model (MGM), which allows for the combination of different variable types [48]. Mixed refers to the possibility to include variables with continuous, categorical and Poisson distributions in the same model. In an MGM, the nodes represent variables and the connections (edges) represent dependencies controlled for all other variables. Thus, the nodes that are unconnected are conditionally independent [48, 50]. In the case of continuous variables only, the edge weights represent partial correlations, similar to Gaussian Graphical Models (GGMs) [52]. The partial correlations range from -1 to 1. In the context of a MGM with both continuous and categorical variables, the edge weights represent a summary of several parameters involved in the regression of the variables, considering the appropriate measures for each variable including conditional variances or conditional probabilities. Prior studies have shown good sensitivity (the ability to identify true edges) and specificity (the ability to ignore spurious edges) of GGMs, meaning that networks are able to provide an accurate depiction of true connections [50].

Network analysis may provide insight into complex diet-disease relations, complementary to regression analysis. Confounding factors are visualized in a network when they are (in)directly connected with the exposure and outcome, indicating that these should be controlled for in the regression analysis. In addition, network models offer a valuable approach to exploring mediating effects, as multiple indirect pathways are visualized at once, which helps to interpret the complex pathways [50]. Furthermore, network models are suitable for highly correlated data, as the high correlation is informative of interconnectedness and clustering. In regression models, it can become difficult to disentangle the associations of the two strongly correlated variables with an outcome. The variable of interest might only be strongly associated with the outcome because it is strongly related to another variable, or to a cluster of variables, and thereby accounts for all the variance of these related variables in the regression model. In this case, the interrelatedness of the variable is associated with the outcome in the regression model, rather than the variable of interest itself. Thus, network analysis is suitable for highly interrelated factors such as food groups and clinical markers, while adjusting for all these factors in regression models could result in spurious associations in regression models. Careful consideration of the variables in the network model is needed as network models are prone to collider bias. This bias might occur when a variable is caused both by the exposure and outcome (a collider) and is included in the network model. By conditioning this variable an artificial association between the exposure and the outcome is created [53].

We used the mgm and qgraph package in RStudio to estimate a k-order MGM in participants with complete data [48, 49]. To obtain a joint distribution, a l1-penalized (LASSO) regression withing the Generalized Linear Model (GLM) framework is performed [54]. This LASSO regularization is applied to shrink small edges to zero to retain meaningful associations. This regularization parameter λ was selected using the Extended Bayesian Information Criterion (EBIC) with tuning parameter $\gamma = 0.5$. This tuning parameter is set at 0.5 to obtain a parsimonious model with higher specificity [55].

The nodes are connected with blue (positive, higher intake or risk factor), orange (negative, lower intake or protective factor) or grey (relationship between at least one categorical node). For example, for food groups, positive edges indicate that intake of one food group correlates with higher intake of another food group, while negative edges mean the opposite. The edge with the highest absolute weight has full-colour saturation at the widest width. The nodes are positioned using the Fruchterman-Reingold layout algorithm, organizing the network according to strength of the connections [56]. This visual aspect of network models enhances interpretation of underlying data structures and complex associations [57]. To note, with each estimation of a network figure, the edges are identical, however node placement may differ considerably, thus placement of nodes based on visual assessment only should not be overinterpreted.

Node predictability

For continuous variables, predictability of nodes was specified as the proportion of explained variance (R2) [58]. The predictability for categorical variables is slightly different. For categorical variables, the proportion of correct classification (accuracy) is calculated, divided into the accuracy of the intercept model (the accuracy achieved by just using the marginal of the variable, e.g. 90% correct if the other nodes do not predict this variable and the prevalence is 90%), and the additionally achieved accuracy achieved by all other remaining variables.

Clustering of nodes

The spinglass algorithm from the igrpah package was used to identify clusters of nodes in the network model [59]. With this procedure, clustering nodes are identified, with each node belonging to one cluster only. The spinglass algorithm provides different results with each run, therefore we performed the algorithm 100 times and extracted the number of clusters with the highest frequency.

Centrality

For each node in the network, the centrality was estimated and standardized by calculating the strength (i.e., sum of the absolute connection weight), closeness (i.e., inverse sum of the distance between one node and all other nodes) and betweenness (i.e., number of times the node is in the shortest paths between other nodes) [60]. From a social network perspective, nodes that are more central could be more influential in the network. Therefore, nodes that represent modifiable risk factors with high centrality could be targets for interventions. Another interpretation is that central nodes may contribute to the development and maintenance of the outcome. Centrality indices yield insight in the structure and behaviour of the related variables, a unique feature of network modelling which would not be possible when assessing the variables separately [58]. However, centrality measures are relative metrics, and their value is dependent on all other factors in the network [61]. Furthermore, the assumptions on which centrality indices are based on may not hold for networks with conditional independent relationships instead of observable connections in social networks [62]. Some studies showed that closeness and betweenness were somewhat unstable in cross-sectional networks [63]. Furthermore, a study on symptoms of social anxiety disorder showed that the most central symptoms were simply those that were more commonly reported [64]. The application of centrality in networks with nutritional intake data is understudied. Therefore, the centrality indices should not be overinterpreted, and always be regarded with other properties of the network such as clustering and prediction.

Edge-weight accuracy

The accuracy of the edges in the network were assessed by bootstrapping the 95%CI around the edge weights [50], using non-parametric bootstrapping (*n* boots = 100) with the resample function in the mgm package [48]. Hereby, the conditional independent

relationships are recalculated 100 times in randomly drawn sample subsets with replacement, meaning that some participants may be included more than once. For each edge weight, the proportion of nonzero estimates across all bootstrap models were obtained, plotted at the mean with lines indicating the 5% lower and 95% upper quantile. Thus, if the edge between two nodes varies widely in each bootstrapped sample, the CI will be wide [48, 50]. Smaller CIs around the edge-weights indicate more accurate estimates. Wide, overlapping CIs indicate that there are little differences between edge-weights even though some might appear stronger than others, and interpreting the order of the edges should be done with care.

Centrality stability

The stability of centrality indices was investigated by estimating the networks in subsets of the data using case-dropping bootstrapping (*n* boots = 1000) with the bootnet package [50]. Hereby, the network and centrality indices are estimated in 1000 bootstrapped samples created by randomly excluding a percentage of the sample. This procedure is repeated excluding 10% up to 99% of the sample. The average correlation between the centrality incidence in the networks sampled with the participants dropped and the original sample were plotted. The Correlation-Stability Coefficient (CS-coefficient) was calculated, describing the mean percentage of the sample that can be dropped withholding a correlation of r=0.9 between the case-dropped centrality indices and the original sample indices. A high CS-coefficient indicates that participant characteristics are not influencing the centrality indices. The CS-coefficient should not be lower than 0.25, and preferably be higher than 0.5.

Non-normality

The GGM can be estimated if data is assumed to be jointly normally distributed as an estimate of the covariance matrix is required as input. Consequences of violating this assumption within the GGM network are to be investigated in more detail, and methods to deal with multivariate non-normality are under development. As such, data-driven bootstrap methods to assess network stability are recommended [65]. A model that can be used to relax the normality assumption is the semi-parametric Gaussian copula, using a nonparanormal transformation described by Liu *et al.*, 2009, 2012 [66-68]. An advantage of this nonparanormal transformation is the fast computation and it requires one pass of the data matrix only, making it easy to implement within the procedure for mixed graphical models.

The dairy types, food groups, physical activity and TAG were not-normally distributed, as assessed with histograms and QQ-plots. We applied a nonparanormal transformation to these variables using the huge package [68]. The shrunken empirical cumulative distribution function (ECDF) option was selected to estimate the transformation function. With this function the empirical parameters were shrunken towards a uniform distribution (a continuous probability distribution in which all values in the range of

the distribution are equally likely to occur) [66, 67]. Thus, the averages of the empirical quantiles were weighted towards a normal distribution. These transformed variables were merged with the remaining normally distributed continuous variables and the categorical variables. Next, the network model was estimated as described above.

Results

Participant characteristics

In the study sample of 74,132 participants, the mean dairy intake was 3.5 ± 1.8 servings/d, mainly consisting of high-fat cheese intake (1.15 ± 1.18 servings/day), low-fat milk (0.60 ± 0.82 servings/day) and low-fat yogurt (0.18 ± 0.29 servings/day) (**Table 1**). Dairy intake contributed to a median of 10.0 weight% of the total diet (**Supplemental Figure 1**). The mean age was 45.5 ± 12.3 y, 59.7% were female, and 17.6% were current smokers (**Table 2**). The mean waist circumference was 89.1 ± 11.6 cm, the mean BMI was 25.7 ± 3.9 kg/m2, and 12.4% were obese (BMI ≥30 kg/m2). Participants with a high dairy intake (top quartile, mean 5.9 ± 1.7 servings/day) compared with low dairy intake (bottom quartile, mean 1.6 ± 0.5 servings/day) were on average older (47.8 ± 11.7 compared with 42.5 ± 12.4 years), more often male (52.8% compared with 63.4% females), more physically active (220 min/wk, IQR: 90–420 compared with 180 min/wk, IQR: 60–360), had a higher alcohol intake (7.9 ± 8.9 compared with 7.0 ± 9.0), were more often hypertensive (24.7% compared with 21.3%), and had a slightly higher waist circumference and total cholesterol. Furthermore, the mean total energy intake was higher, as well as the intake of fruits, vegetables, bread, nuts, meat, coffee, added sugar, and calcium. Participant characteristics according to quartiles of milk, yogurt, and cheese intake, sex, by included versus excluded from the current analysis and by having a follow-up measurement versus not are presented in **Supplemental Table 2**. Participants included in the study were more often highly educated (32.2% compared with 26.5%, $P < 0.001$), less often smokers (17.6% compared with 24.2%, $P < 0.001$), and had a generally more beneficial CVD risk factor profile compared to those excluded (all $P < 0.001$). Participants lacking a follow-up measurement were slightly younger (40.8 ± 12.5 compared with 45.5 ± 12.3, $P < 0.001$), more often smokers (26.1% compared with 17.6%, $P < 0.001$), and had a slightly lower diet quality (23.2 ± 6.1 compared with 24.4 ± 6.0, $P < 0.001$) than those with a follow-up measurement.

Table 2. Baseline characteristics of participants of the Lifelines Cohort in the total cohort and according to quartiles of total dairy intake (*N* = 74,132).

	Total	**Q1**	**Q2**	**Q3**	**Q4**
	(*n* = 74,132)	(*n* = 18,600)	(*n* = 18,485)	(*n* = 18,503)	(*n* = 18,544)
Total dairy intake, (servings/day)	3.5 ± 1.8	1.6 ± 0.5	2.8 ± 0.3	3.8 ± 0.3	5.9 ± 1.7
Range	0–22.4	0–2.3	2.3–3.3	3.3–4.5	4.5–22.4
Follow-up time	4.1 ± 1.1	4.1 ± 1.1	4.1 ± 1.1	4.1 ± 1.1	4.1 ± 1.1
Sex, female (%)	59.7	63.4	61.6	61.0	52.8
Age at baseline (y)	45.5 ± 12.3	42.5 ± 12.4	45.1 ± 12.4	46.5 ± 12.1	47.8 ± 11.7
Educational level (%)					
Low	27.6	26.9	27.4	27.7	28.2
Intermediate	40.2	41.2	40.7	39.7	39.4
High	32.2	31.9	31.9	32.7	32.4
Smoking (%)					
Never	48.2	48.1	48.5	48.6	47.5
Former	34.2	31.0	34.2	35.0	36.6
Current	17.6	20.8	17.4	16.4	15.9
Pack years	12.0 [5.6, 20.2]	11.0 [4.8, 19.2]	12.0 [5.9, 19.6]	12.4 [5.9, 20.0]	13.5 [6.5, 21.9]
Alcohol (g/day)	7.3 ± 8.7	7.0 ± 9.0	7.0 ± 8.4	7.2 ± 8.4	7.9 ± 8.9
Physical activity (min/wk)	200 [75, 375]	180 [60, 360]	195 [75, 360]	210 [85, 380]	220 [90, 420]
Family history of diabetes (%)					
No	29.5	32.0	29.5	28.6	27.9
Yes	8.2	8.1	8.6	8.3	7.7
Unknown/missing	62.3	59.9	61.9	63.1	64.3
BMI (kg/m2)	25.7 ± 3.9	25.6 ± 4.1	25.7 ± 4.0	25.7 ± 3.9	25.7 ± 3.8
Waist circumference (cm)	89.1 ± 11.6	88.1 ± 12.0	88.9 ± 11.6	89.2 ± 11.4	90.0 ± 11.3
Total cholesterol (mmol/L)	5.1 ± 1.0	5.0 ± 1.0	5.1 ± 1.0	5.2 ± 1.0	5.2 ± 1.0
LDL cholesterol (mmol/L)	3.3 ± 0.9	3.2 ± 0.9	3.2 ± 0.9	3.3 ± 0.9	3.3 ± 0.9
HDL cholesterol (mmol/L)	1.5 ± 0.4	1.5 ± 0.4	1.5 ± 0.4	1.5 ± 0.4	1.5 ± 0.4
TAG (mmol/L)	0.9 [0.7, 1.3]	0.9 [0.7, 1.3]	1.0 [0.7, 1.3]	0.9 [0.7, 1.3]	1.0 [0.7, 1.3]
Hypertension (%)	23.4	21.3	23.4	24.5	24.7
Dietary intake					
Energy intake (kcal/day)	2054 ± 583	1773 ± 537	1954 ± 511	2097 ± 513	2394 ± 585
Diet quality[1]	24.4 ± 6.0	23.4 ± 6.2	24.4 ± 6.0	24.9 ± 5.9	24.8 ± 5.8
Fruit (g/day)	141 ± 112	126 ± 113	136 ± 106	145 ± 108	156 ± 117
Vegetables (g/day)	105 ± 58	100 ± 61	102 ± 57	106 ± 56	112 ± 60
Bread (g/day)	113 ± 60	98 ± 60	109 ± 57	116 ± 56	129 ± 63
Legumes (g/day)	9.7 ± 15.4	9.2 ± 16.7	9.2 ± 14.2	9.5 ± 14.2	10.9 ± 16.3
Nuts (g/day)	12.8 ± 14.8	11.5 ± 14.6	12.3 ± 14.3	13.0 ± 14.2	14.5 ± 15.9
Meat (red and processed) (g/day)	66.7 ± 33.3	63.2 ± 33.9	65.2 ± 31.4	67.0 ± 31.4	71.2 ± 35.7
Fish (g/day)	12.5 ± 12.7	12.1 ± 13.1	12.3 ± 12.2	12.5 ± 12.2	13.0 ± 13.0

Table 2. Baseline characteristics of participants of the Lifelines Cohort in the total cohort and according to quartiles of total dairy intake (*N* = 74,132). (continued)

	Total	Q1	Q2	Q3	Q4
	(*n* = 74,132)	(*n* = 18,600)	(*n* = 18,485)	(*n* = 18,503)	(*n* = 18,544)
Coffee (g/day)	423 ± 275	360 ± 285	416 ± 271	438 ± 261	478 ± 268
Tea (g/day)	256 ± 248	261 ± 266	254 ± 244	256 ± 238	252 ± 240
SSBs (g/day)	142 ± 172	149 ± 188	137 ± 164	136 ± 159	144 ± 173
Total fat (en%)	34.5 ± 4.8	33.9 ± 5.2	34.3 ± 4.7	34.5 ± 4.5	35.3 ± 4.8
Saturated fat (en%)	15.3 ± 2.3	14.6 ± 2.4	15.0 ± 2.2	15.4 ± 2.1	16.1 ± 2.2
Carbohydrates (en%)	46.0 ± 5.7	47.0 ± 6.3	46.5 ± 5.5	46.0 ± 5.2	44.7 ± 5.5
Added sugar (g/day)	52.5 ± 32.5	47.8 ± 32.7	51.1 ± 30.9	53.2 ± 31.2	57.8 ± 34.3
Calcium (mg/day)	996 ± 352	634 ± 149	854 ± 124	1044 ± 132	1432 ± 314

Values are mean ± SD for continuous variables with a normal distribution, or median [IQR] for continuous variables with a skewed distribution, percentages for categorical variables, based on unimputed data.
[1] Diet quality was measured using the Lifelines Diet Score (LLDS), reflecting adherence to the 2015 Dutch Dietary Guidelines for prevention of chronic disease, with higher scores indicating better adherence [30]. The LLDS consists of 12 food groups, and possible scores range from 0 to 48.
Abbreviations: BMI, body mass index; en%, percentage of total energy intake; HDL, high-density lipoprotein; LDL, low-density lipoprotein; SD, standard deviation; SSB, sugar-sweetened beverage; TAG, triglyceride.

Dairy intake and incident prediabetes

At a mean follow-up of 4.1 ± 1.1 years (median: 4.1; IQR: 3.3–4.8), 2,746 incident cases of prediabetes were identified among 74,132 participants (3.7%). In multivariable-adjusted models, the top compared with bottom quartiles of plain milk (59.6% of total milk) and low-fat plain milk (91.0% of total plain milk) intake were associated with a 17% and 18% higher risk of prediabetes, respectively ($RR_{Q4vs.Q1}$ 1.17, 95%CI 1.05-1.30, P_{trend} = 0.04 and 1.18, 95%CI 1.06-1.31, P_{trend} = 0.01) (**Table 3**). Associations with these exposures on a continuous scale were not significant ($RR_{serving/day}$ respectively 1.04, 95%CI 0.99-1.08 and 1.04; 95%CI 1.00-1.09). Better fit of a nonlinear model was found for plain milk (P for nonlinearity = 0.02), with the highest risk at low intake (0.25–0.6 serving/day) but no additional risk at higher intakes (**Figure 2**). High-fat yogurt (24.5% of total yogurt intake) was associated with a lower risk of prediabetes in model 1 ($RR_{serving/day}$ 0.73, 95%CI 0.58-0.93, $RR_{Q4vs.Q1}$ 0.81, 95%CI 0.69-0.95, P_{trend} = 0.02) and model 2 ($RR_{serving/day}$ 0.78, 95%CI 0.62-0.98, quartiles P_{trend} = 0.08). However, this association slightly attenuated after additional adjustment for intake of other food groups related to diabetes (model 3, $RR_{serving/day}$ 0.80, 95%CI 0.64-1.01, $R_{Q4vs.Q1}$ 0.86, 95%CI 0.74-1.01, P_{trend} = 0.13). A potential nonlinear association was observed for cream, with lower risk already at low intakes (P for nonlinearity = 0.03) (**Figure 2**). Total dairy, fermented dairy, and cheese, irrespective of fat content, milk, high-fat plain milk, total and low-fat yogurt, and ice cream were not associated with prediabetes (**Table 3**). With additional adjustment for clinical markers including TAGs, LDL cholesterol, waist circumference, and hypertension, the associations for total plain milk, low-fat plain milk, and high-fat yogurt were attenuated.

Table 3. Associations of dairy product types and prediabetes risk in the Lifelines Cohort (*n* = 74,132).

	Relative risk (95% CI) across intake range categories[1]					Continuous[2]
	Q1	Q2	Q3	Q4	P_{trend}	RR (95% CI)
Total dairy						
n cases/n total	643/18,600	662/18,485	701/18,503	740/18,544		
Median, servings/d	1.7	2.8	3.8	5.5		
Model 1	1 (ref)	0.91 (0.82-1.01)	0.90 (0.81-1.01)	0.90 (0.80-1.01)	0.13	0.99 (0.97-1.01)
Model 2	1 (ref)	0.93 (0.84-1.03)	0.94 (0.84-1.05)	0.95 (0.84-1.06)	0.51	1.00 (0.98-1.02)
Model 3	1 (ref)	0.94 (0.84-1.04)	0.95 (0.85-1.06)	0.96 (0.86-1.08)	0.73	1.00 (0.98-1.03)
Model 4	1 (ref)	0.93 (0.84-1.03)	0.94 (0.84-1.05)	0.95 (0.85-1.07)	0.56	1.00 (0.98-1.02)
High-fat dairy						
n cases/n total	673/18,469	665/18,637	685/18,516	723/18,510		
Median, servings/d	0.4	1.0	1.7	3.0		
Model 1	1 (ref)	0.96 (0.87-1.07)	0.93 (0.84-1.04)	0.92 (0.83-1.03)	0.16	0.98 (0.95-1.01)
Model 2	1 (ref)	0.98 (0.88-1.09)	0.96 (0.86-1.06)	0.95 (0.85-1.06)	0.36	0.99 (0.96-1.02)
Model 3	1 (ref)	0.97 (0.87-1.08)	0.95 (0.85-1.05)	0.94 (0.84-1.06)	0.34	0.99 (0.96-1.02)
Model 4	1 (ref)	0.97 (0.87-1.07)	0.96 (0.87-1.07)	0.98 (0.88-1.10)	0.90	1.00 (0.97-1.03)
Low-fat dairy						
n cases/n total	702/19,090	607/18,010	711/18,481	726/18,551		
Median, servings/d	0.5	1.3	2.1	3.4		
Model 1	1 (ref)	0.90 (0.82-1.00)	0.98 (0.88-1.08)	0.93 (0.84-1.03)	0.40	1.00 (0.97-1.03)
Model 2	1 (ref)	0.93 (0.84-1.03)	1.01 (0.91-1.12)	0.97 (0.88-1.08)	0.95	1.01 (0.98-1.04)
Model 3	1 (ref)	0.93 (0.84-1.04)	1.02 (0.92-1.13)	0.99 (0.89-1.10)	0.76	1.01 (0.99-1.04)
Model 4	1 (ref)	0.92 (0.83-1.02)	0.99 (0.89-1.10)	0.94 (0.85-1.04)	0.47	1.00 (0.97-1.03)
Fermented dairy						
n cases/n total	602/18,476	657/18,555	698/18,562	789/18,539		
Median, servings/d	0.6	1.4	2.2	3.7		
Model 1	1 (ref)	0.94 (0.84-1.04)	0.89 (0.79-0.99)	0.91 (0.82-1.02)	0.13	0.98 (0.95-1.01)
Model 2	1 (ref)	0.96 (0.86-1.07)	0.93 (0.84-1.04)	0.98 (0.87-1.09)	0.79	0.99 (0.97-1.02)
Model 3	1 (ref)	0.97 (0.87-1.09)	0.95 (0.85-1.06)	1.01 (0.90-1.13)	0.73	1.00 (0.98-1.03)
Model 4	1 (ref)	0.95 (0.86-1.06)	0.93 (0.84-1.04)	1.00 (0.89-1.12)	0.80	1.00 (0.97-1.03)

Table 3. Associations of dairy product types and prediabetes risk in the Lifelines Cohort (*n* = 74,132). (continued)

	Relative risk (95% CI) across intake range categories[1]					Continuous[2]
	Q1	Q2	Q3	Q4	P_{trend}	RR (95% CI)
High-fat fermented dairy						
n cases/n total	709/18,587	613/18,471	691/18,566	733/18,508		
Median, servings/d	0.2	0.7	1.3	2.5		
Model 1	1 (ref)	0.88 (0.79-0.97)	0.91 (0.82-1.00)	0.89 (0.80-0.99)	0.14	0.98 (0.94-1.01)
Model 2	1 (ref)	0.90 (0.81-1.00)	0.95 (0.85-1.05)	0.94 (0.85-1.05)	0.57	0.99 (0.96-1.02)
Model 3	1 (ref)	0.90 (0.81-1.00)	0.94 (0.85-1.05)	0.95 (0.85-1.05)	0.66	0.99 (0.96-1.02)
Model 4	1 (ref)	0.90 (0.81-1.00)	0.94 (0.85-1.04)	0.97 (0.87-1.08)	0.93	1.00 (0.97-1.03)
Low-fat fermented dairy						
n cases/n total	617/18,279	627/17,964	722/19,389	780/18,500		
Median, servings/d	0.0	0.2	0.7	1.8		
Model 1	1 (ref)	1.10 (0.99-1.23)	1.02 (0.92-1.13)	0.99 (0.90-1.10)	0.31	0.99 (0.96-1.03)
Model 2	1 (ref)	1.12 (1.01-1.25)	1.06 (0.95-1.18)	1.04 (0.94-1.16)	0.89	1.01 (0.97-1.04)
Model 3	1 (ref)	1.13 (1.01-1.26)	1.07 (0.97-1.19)	1.07 (0.97-1.19)	0.60	1.02 (0.98-1.06)
Model 4	1 (ref)	1.08 (0.97-1.20)	1.02 (0.92-1.13)	1.01 (0.91-1.12)	0.67	1.00 (0.97-1.04)
Total milk						
n cases/n total	621/18,500	803/21,680	708/18,186	614/15,766		
Median, servings/d	0.1	0.6	1.3	2.3		
Model 1	1 (ref)	1.07 (0.97-1.19)	1.04 (0.93-1.16)	1.11 (0.99-1.24)	0.15	1.02 (0.99-1.06)
Model 2	1 (ref)	1.07 (0.96-1.18)	1.05 (0.94-1.16)	1.12 (1.00-1.26)	0.08	1.03 (0.99-1.07)
Model 3	1 (ref)	1.07 (0.96-1.18)	1.05 (0.94-1.17)	1.13 (1.01-1.27)	0.06	1.03 (0.99-1.07)
Model 4	1 (ref)	1.04 (0.94-1.16)	1.03 (0.93-1.15)	1.11 (0.99-1.24)	0.11	1.02 (0.98-1.06)
Plain milk						
n cases/n total	812/21,541	643/17,814	683/18,421	608/16,356		
Median, servings/d	0.0	0.1	0.7	1.9		
Model 1	1 (ref)	1.16 (1.05-1.28)	1.04 (0.94-1.15)	1.18 (1.06-1.31)	0.03*	1.04 (0.99-1.08)
Model 2	1 (ref)	1.14 (1.03-1.26)	1.04 (0.94-1.15)	1.17 (1.06-1.31)	0.03*	1.04 (0.99-1.08)
Model 3	1 (ref)	1.13 (1.02-1.25)	1.04 (0.94-1.14)	1.17 (1.05-1.30)	0.04*	1.04 (0.99-1.08)
Model 4	1 (ref)	1.11 (1.01-1.23)	1.02 (0.92-1.13)	1.13 (1.01-1.25)	0.13	1.02 (0.98-1.06)

Table 3. Associations of dairy product types and prediabetes risk in the Lifelines Cohort (*n* = 74,132). (continued)

	Relative risk (95% CI) across intake range categories[1]					Continuous[2]
	Q1	Q2	Q3	Q4	P_{trend}	RR (95% CI)
High-fat plain milk						
n cases/n total	2,476/66,901	87/2321	96/2479	87/2431		
Median, servings/d	0.0	0.1	0.4	0.9		
Model 1	1 (ref)	1.10 (0.89-1.36)	1.12 (0.92-1.37)	1.01 (0.81-1.24)	0.66	1.00 (0.87-1.15)
Model 2	1 (ref)	1.06 (0.86-1.30)	1.10 (0.90-1.34)	0.95 (0.77-1.17)	0.85	0.97 (0.84-1.12)
Model 3	1 (ref)	1.04 (0.85-1.29)	1.08 (0.89-1.32)	0.93 (0.76-1.15)	0.72	0.96 (0.83-1.11)
Model 4	1 (ref)	1.08 (0.87-1.32)	1.11 (0.91-1.35)	1.00 (0.81-1.24)	0.71	1.01 (0.88-1.16)
Low-fat plain milk						
n cases/n total	935/24,594	608/17,155	642/17,517	561/14,866		
Median, servings/d	0.0	0.1	0.7	1.9		
Model 1	1 (ref)	1.12 (1.01-1.24)	1.02 (0.92-1.12)	1.17 (1.06-1.30)	0.02*	1.04 (0.99-1.08)
Model 2	1 (ref)	1.11 (1.00-1.23)	1.03 (0.93-1.13)	1.18 (1.07-1.31)	0.01**	1.04 (1.00-1.09)
Model 3	1 (ref)	1.10 (1.00-1.22)	1.03 (0.93-1.13)	1.18 (1.06-1.31)	0.01**	1.04 (1.00-1.09)
Model 4	1 (ref)	1.08 (0.98-1.19)	1.01 (0.91-1.11)	1.12 (1.01-1.25)	0.08	1.02 (0.98-1.07)
Yogurt						
n cases/n total	1,099/29,786	486/13,356	613/16,155	548/14,835		
Median, servings/d	0.0	0.1	0.3	0.6		
Model 1	1 (ref)	1.11 (1.00-1.23)	1.05 (0.95-1.15)	0.93 (0.84-1.03)	0.10	0.90 (0.80-1.02)
Model 2	1 (ref)	1.12 (1.01-1.24)	1.08 (0.98-1.19)	0.98 (0.88-1.08)	0.61	0.95 (0.85-1.07)
Model 3	1 (ref)	1.12 (1.01-1.24)	1.08 (0.98-1.20)	1.00 (0.90-1.11)	0.96	0.98 (0.87-1.10)
Model 4	1 (ref)	1.09 (0.98-1.21)	1.08 (0.98-1.19)	1.02 (0.92-1.13)	0.71	1.00 (0.89-1.13)
High-fat yogurt						
n cases/n total	2,199/59,002	181/4,958	211/5,465	155/4,707		
Median, servings/d	0.0	0.1	0.2	0.5		
Model 1	1 (ref)	1.01 (0.87-1.17)	1.02 (0.88-1.17)	0.81 (0.69-0.95)	0.02*	0.73 (0.58-0.93)
Model 2	1 (ref)	0.99 (0.86-1.15)	1.03 (0.89-1.18)	0.85 (0.72-0.99)	0.08	0.78 (0.62-0.98)
Model 3	1 (ref)	0.99 (0.85-1.14)	1.03 (0.89-1.18)	0.86 (0.74-1.01)	0.13	0.80 (0.64-1.01)
Model 4	1 (ref)	1.01 (0.87-1.17)	1.08 (0.94-1.24)	0.96 (0.82-1.13)	0.96	0.94 (0.76-1.18)

Table 3. Associations of dairy product types and prediabetes risk in the Lifelines Cohort (*n* = 74,132). (continued)

	Relative risk (95% CI) across intake range categories[1]					Continuous[2]
	Q1	Q2	Q3	Q4	P_{trend}	RR (95% CI)
Low-fat yogurt						
n cases/n total	1,427/39,011	438/11,870	425/11,557	456/11,694		
Median, servings/d	0.0	0.1	0.3	0.6		
Model 1	1 (ref)	1.14 (1.02-1.26)	1.04 (0.94-1.16)	1.01 (0.91-1.12)	0.93	0.97 (0.86-1.10)
Model 2	1 (ref)	1.13 (1.02-1.26)	1.07 (0.96-1.18)	1.05 (0.95-1.16)	0.34	1.02 (0.90-1.16)
Model 3	1 (ref)	1.13 (1.02-1.26)	1.07 (0.96-1.19)	1.07 (0.96-1.19)	0.19	1.04 (0.92-1.18)
Model 4	1 (ref)	1.09 (0.98-1.21)	1.04 (0.93-1.15)	1.05 (0.94-1.16)	0.41	1.02 (0.90-1.16)
Cheese						
n cases/n total	601/18,473	664/18,583	697/18,451	784/18,625		
Median, servings/d	0.3	0.9	1.5	2.8		
Model 1	1 (ref)	0.93 (0.83-1.04)	0.92 (0.83-1.03)	0.97 (0.87-1.08)	0.94	0.99 (0.96-1.02)
Model 2	1 (ref)	0.95 (0.85-1.05)	0.95 (0.85-1.05)	1.01 (0.91-1.13)	0.53	1.00 (0.97-1.03)
Model 3	1 (ref)	0.96 (0.86-1.07)	0.96 (0.86-1.07)	1.03 (0.93-1.16)	0.33	1.01 (0.98-1.04)
Model 4	1 (ref)	0.95 (0.85-1.05)	0.94 (0.84-1.05)	1.01 (0.90-1.13)	0.59	1.00 (0.97-1.03)
High-fat cheese						
n cases/n total	672/18,539	672/18,384	661/18,613	741/18,596		
Median, servings/d	0.1	0.5	1.1	2.3		
Model 1	1 (ref)	1.02 (0.92-1.13)	0.91 (0.82-1.02)	0.96 (0.86-1.06)	0.25	0.98 (0.95-1.02)
Model 2	1 (ref)	1.05 (0.94-1.16)	0.95 (0.86-1.06)	1.00 (0.90-1.11)	0.69	0.99 (0.96-1.03)
Model 3	1 (ref)	1.04 (0.94-1.16)	0.95 (0.85-1.05)	1.00 (0.90-1.11)	0.73	1.00 (0.96-1.03)
Model 4	1 (ref)	1.03 (0.93-1.15)	0.94 (0.85-1.05)	1.01 (0.91-1.13)	0.97	1.00 (0.97-1.03)
Low-fat cheese						
n cases/n total	1,398/39,752	393/11,523	447/11,396	508/11,461		
Median, servings/d	0.0	0.2	0.5	1.2		
Model 1	1 (ref)	1.07 (0.96-1.19)	1.08 (0.97-1.20)	1.04 (0.94-1.15)	0.39	1.02 (0.97-1.07)
Model 2	1 (ref)	1.08 (0.97-1.21)	1.08 (0.98-1.20)	1.05 (0.95-1.16)	0.34	1.02 (0.97-1.06)
Model 3	1 (ref)	1.08 (0.97-1.20)	1.09 (0.99-1.21)	1.07 (0.97-1.18)	0.16	1.03 (0.98-1.07)
Model 4	1 (ref)	1.03 (0.93-1.15)	1.02 (0.92-1.14)	1.01 (0.91-1.11)	0.94	1.00 (0.96-1.05)

Table 3. Associations of dairy product types and prediabetes risk in the Lifelines Cohort (*n* = 74,132). (continued)

	Relative risk (95% CI) across intake range categories[1]					Continuous[2]
	Q1	Q2	Q3	Q4	P_{trend}	RR (95% CI)
Cream						
n cases/n total	855/20,831	951/26,458	538/15,307	402/11,536		
Median, servings/d	0.0	0.005	0.01	0.02		
Model 1	1 (ref)	0.85 (0.78-0.93)	0.87 (0.78-0.97)	0.88 (0.78-0.99)	0.05*	0.75 (0.03-16.7)
Model 2	1 (ref)	0.89 (0.81-0.97)	0.91 (0.82-1.02)	0.91 (0.80-1.02)	0.17	0.79 (0.04-14.5)
Model 3	1 (ref)	0.88 (0.80-0.96)	0.90 (0.81-1.00)	0.88 (0.78-0.99)	0.05	0.40 (0.02-8.44)
Model 4	1 (ref)	0.91 (0.83-1.00)	0.93 (0.83-1.03)	0.92 (0.81-1.03)	0.21	0.94 (0.05-16.48)
Ice cream						
n cases/n total	1,077/27,713	621/17,494	674/18,680	374/10,245		
Median, servings/d	0.0	0.02	0.1	0.1		
Model 1	1 (ref)	0.92 (0.83-1.01)	0.95 (0.87-1.05)	0.91 (0.81-1.03)	0.15	0.97 (0.56-1.98)
Model 2	1 (ref)	0.97 (0.88-1.07)	1.01 (0.92-1.11)	0.96 (0.85-1.08)	0.73	1.16 (0.64-2.09)
Model 3	1 (ref)	0.97 (0.88-1.07)	1.00 (0.91-1.10)	0.94 (0.83-1.06)	0.46	1.03 (0.56-1.89)
Model 4	1 (ref)	0.97 (0.88-1.07)	0.99 (0.90-1.09)	0.92 (0.81-1.03)	0.22	0.93 (0.49-1.75)

[1] Relative risks (95CIs) were estimated across four categories split by quartile values (Q1 to Q4) or non-consumers + tertile or median categories with the lowest category as the reference, adjusted for covariates as follows: Model 1 included age, sex, energy intake and follow-up duration. Model 2 additionally adjusted for educational level, alcohol use, smoking behaviour, physical activity level and family history of diabetes. Model 3 additionally adjusted for food groups associated with type 2 diabetes, including fruit, vegetables, bread, legumes, nuts, red and processed meat, coffee, tea, and sugar-sweetened beverages. Model 4 additionally adjusted for TAGs, LDL cholesterol, waist circumference, and hypertension. Linear trend across intake range categories was assessed by including median values of each category as a continuous variable in the model.

[2] Relative risks per 1 serving/day (see definition in Table 1) were estimated.* P=0.01 to 0.05, ** P<0.01.

Abbreviations: CI, Confidence Interval; Q, Quartile.

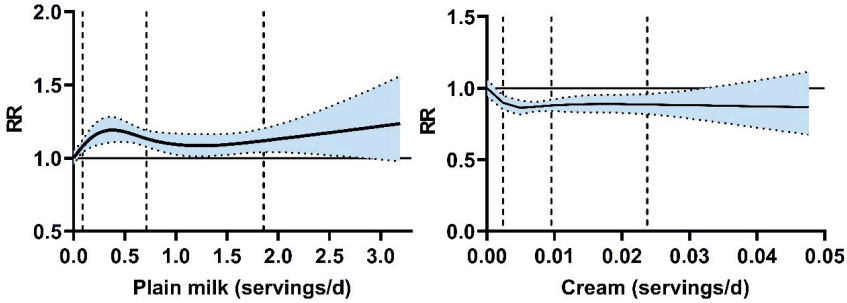


Figure 2. Non-linear associations between plain milk (p for nonlinearity = 0.02) and cream intake (p for nonlinearity = 0.03) and prediabetes risk in the Lifelines study. The solid line indicates the risk estimate fitted with restricted cubic spline regression with 4 knots specified at the 5th percentile at intake of 0 gram/day, and at the 33th, 67th and 95th percentile of intake as indicated by dotted vertical lines. The coloured area indicates the 95% confidence interval. The model was adjusted for age, sex, energy intake, follow-up duration, educational level, alcohol use, smoking behaviour, physical activity level, family history of diabetes, fruit, vegetables, bread, legumes, nuts, red and processed meat, coffee, tea, and sugar-sweetened beverages intake.

Subgroup analysis

Significant interactions were observed for total dairy and sex, high-fat dairy and high-fat cheese and age, and low-fat fermented dairy, and plain milk with waist circumference and low-fat fermented dairy, high- fat yogurt, low-fat cheese with educational level (**Supplemental Table 3**). In participants aged ≥60 y, high-fat dairy was associated with a lower risk of prediabetes ($RR_{serving/day}$ 0.93, 95%CI 0.88-0.99). However, there were no significant associations in other age groups ($RR_{serving/day}$ ranging from 0.93–1.04). High-fat yogurt was associated with lower risk of prediabetes in participants with high educational level ($RR_{serving/day}$ 0.45 95%CI 0.27-0.76) but not in participants with intermediate ($RR_{serving/day}$ 0.78, 95%CI 0.51-1.19) or low ($RR_{serving/day}$ 1.11, 95%CI 0.81-1.54) educational level. In contrast, low-fat cheese was only associated with higher risk of prediabetes among participants with high educational levels ($RR_{serving/day}$ 1.16, 95%CI 1.06-1.26) but not among intermediate and low educational levels ($RR_{serving/day}$ ~1). Other stratified associations were not significant.

Sensitivity analysis

In sensitivity analyses, all associations were similar after additional adjustment for intake of all other dairy types; in models excluding participants with prevalent and incident CVD or cancer; with models using energy- adjusted dairy intake as exposure; and with models using sex- specific quartiles of dairy as exposure (**Supplemental Table 4**). When using the ADA cutoffs, 8,705 incident cases of prediabetes were identified among 74,132 participants (11.7%). Using this cutoff resulted in attenuation of associations for total plain milk, low-fat plain milk, and high-fat yogurt.

Secondary analysis with diabetes risk score and "desire to lose weight"
The diabetes risk score at baseline (PROCAM risk score) was associated with the intake of total dairy and all dairy types (**Supplemental Tables 5, 6**), i.e., a higher baseline risk was related to a lower intake of high-fat dairy ($\beta_{serving/day}$ -0.17, 95%CI -0.20; -0.15) and a higher intake of low-fat dairy ($\beta_{serving/day}$ 0.12, 95%CI 0.10; 0.14). This was seen for all types of dairy (fermented dairy, milk, yogurt, and cheese). A similar pattern was observed for "desire to lose weight," which was associated with lower intake of high-fat dairy ($\beta_{serving/day}$ -0.15, 95%CI -0.16; -0.13) and higher intake of low-fat dairy ($\beta_{serving/day}$ 0.10, 95%CI 0.08; 0.11) (**Supplemental Tables 5, 7**). Furthermore, additional adjustment for "desire to lose weight" slightly attenuated nonsignificant associations for high-fat plain milk, high-fat yogurt, and low-fat yogurt (**Supplemental Table 8**).

Network structure

The network structure, including food groups and dairy types, showed several foods that are frequently consumed together within an eating pattern, as indicated by positive (blue) partial correlations (i.e., correlations corrected for all other factors in the network), or not consumed together, as indicated by negative (orange) partial correlations (**Figure 3**). High-fat dairy types were connected by positive edges (i.e., high-fat milk, high-fat yogurt, and high-fat cheese), which were negatively connected with their low-fat counterpart (e.g., high-fat yogurt correlated negatively with low-fat yogurt), reflecting a preference of participants for either high-fat or low-fat dairy. Meat, bread, high-fat cheese, and coffee were positively connected and somewhat connected with low-fat milk. In contrast, negative connections between beverages, tea, coffee, and SSBs were observed. Food groups considered beneficial for health were connected with positive edges, including vegetables, fruit, fish, legumes, tea, and nuts. A considerable positive connection between low-fat yogurt and fruit was observed.

The complete network structure showed that prediabetes was placed centrally in a cluster of age and clinical markers (**Figure 4; Supplemental Figure 2**). Prediabetes connected strongly with age and waist circumference, which in turn connected with the sociodemographic characteristics and health risk factors, and from there, further connecting with the food groups. Including energy intake in this network resulted in the clustering of food groups with high energy density around energy intake, including bread, meat, nuts and SSBs, as well as milk, cheese, cream, and ice cream. Furthermore, a cluster of fruit, vegetables, fish, legumes, and physical activity was observed. Low-fat yogurt was connected with this cluster in a network without the other dairy types where the positioning of low-fat yogurt was not affected by that of high-fat yogurt (**Supplemental Figure 3**).

Next, predictability was determined, describing the extent to which a node can be predicted by all other factors in the network. Thus, nodes with low predictability are determined by other factors not included in the network *or* self-determined.

Furthermore, a node with many connections but low predictability may indicate that the relevance of connections is limited. The degree of predictability is indicated by the rings around each node in the network plot. Of continuous nodes, energy intake (mean explained variance 0.69 as indicated by the blue rings in the figures), age (0.43), bread (0.41), waist circumference (0.31), and coffee (0.30) had the highest predictability, meaning that these nodes were somewhat determined by all other factors in the network (**Figure 4**). The individual dairy types and physical activity had the lowest predictability, meaning that they were determined to a lesser extent by factors in the network. For categorical nodes, the predictability was specified as the accuracy, divided into the proportion of correct classification by just using the marginal of the variable, and the accuracy achieved by all other remaining variables. The latter part is more relevant as it indicates if other nodes in the network provided sufficient information to classify the node. Prediabetes and low, but for sex, education, and smoking, the network provides some additional information.

Furthermore, the centrality of nodes was considered, quantifying the nodes that have a more central position in the network. Somewhat in line with predictability, energy intake, sex, age, smoking, and waist circumference showed the highest centrality in the network, meaning that they were related more closely to other nodes and were potentially more influential (**Figure 5**). Food groups with high centrality included SSBs, meat, coffee, and tea, reflective of high- intake levels in this population, yet direct edges with clinical markers were small. Of dairy types, high-fat cheese had the highest centrality. LDL cholesterol, physical activity, and other dairy types, especially high-fat milk, and high-fat yogurt, showed low centrality compared to other nodes.

Some dairy types were connected to the cluster of clinical markers, age, and prediabetes, although the visibility of these connections is limited in the network plot (**Figure 4**). Dairy types positively connected with waist circumference included low-fat milk, low-fat cheese, low-fat yogurt, ice cream, and high-fat cheese, and dairy types with negative connections included high-fat yogurt and high-fat milk.

Results on edge–weight accuracy, centrality stability and the semi-parametric copula network can be found in **Supplemental Figures 4 to 6**.

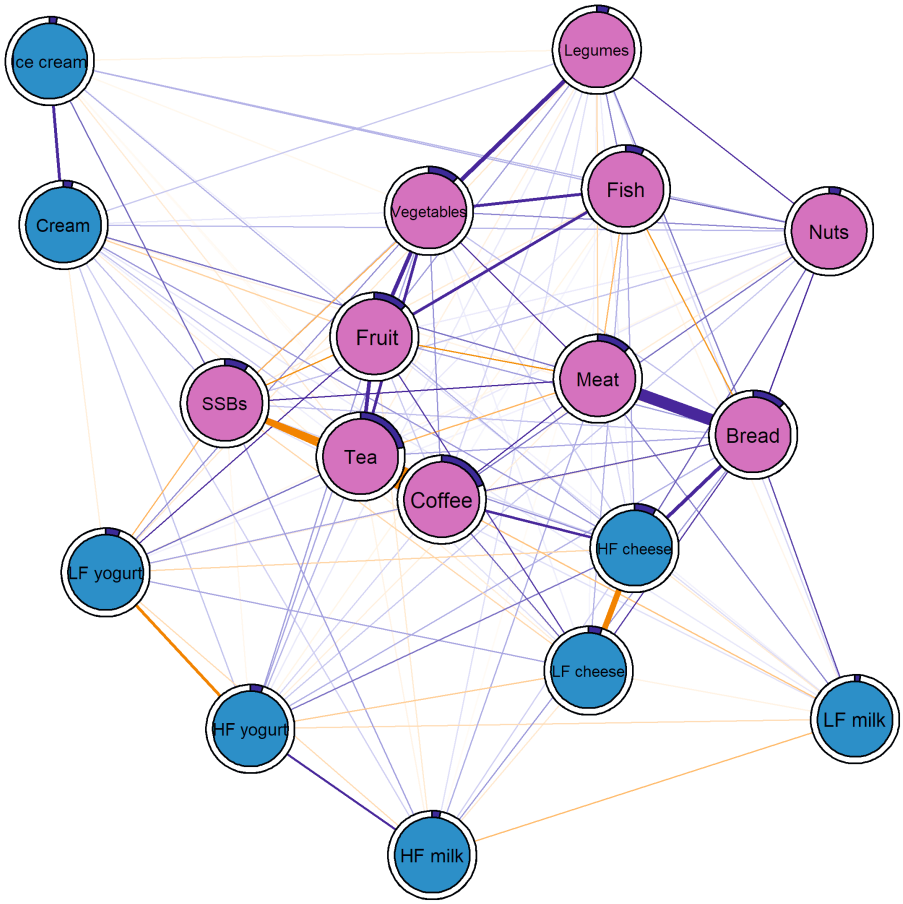


Figure 3. Network structure of dairy intake (blue) and food groups (pink) of the study population with complete data for variables in the model (*n* = 74,132). The edges between nodes (variables) represent conditional independent relationships; blue and orange edges indicate positive and negative relationships respectively between 2 continuous nodes, with the highest edge weight being 0.38 (between nodes coffee and tea). Edge thickness is proportional to the strength of the relationships between the nodes. The absence of an edge indicates that two nodes are conditional independent in the network. The predictability is indicated by the rings around each node; blue rings indicate the proportion variance explained by neighbouring nodes with the full circle indicating a r2 of 1.0. Abbreviations: HF, high-fat; LF, low-fat; SSB, sugar-sweetened beverage.

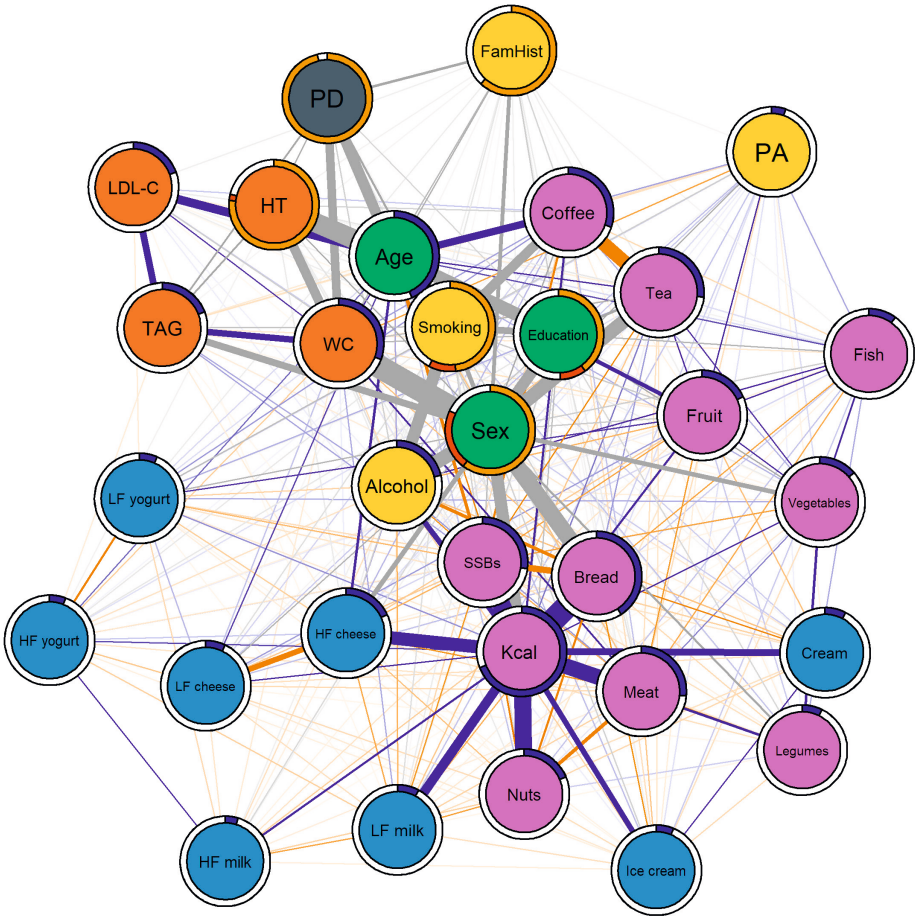


Figure 4. Network structure of dairy intake (blue), food groups and energy intake (pink), health risk factors (yellow), sociodemographic characteristics (green), clinical markers (orange), and prediabetes (grey) of the study population with complete data for variables in the model (*n* = 67,206). The edges between nodes (variables) represent conditional independent relationships; blue and orange edges indicate positive and negative relationships respectively between 2 continuous nodes, and grey edges indicate a relationship between at least 1 categorical variable. Edge thickness is proportional to the strength of the relation between the nodes, with the highest edge weight being 0.51 (between nodes kcal and bread). The absence of an edge indicates that two nodes are conditional independent in the network. The predictability is indicated by the rings around each node; blue rings indicate the proportion variance explained by neighbouring nodes with the full circle indicating a r^2 of 1.0; the range/red rings indicate the accuracy for the categorical nodes, respectively the marginal of the variable and the additionally achieved accuracy by all other remaining variables, with the full circle indicating an accuracy of 100%. Abbreviations: HF, high-fat; HT, hypertension; LDL-C, low-density lipoprotein cholesterol; LF, low-fat; PA, physical activity; PD, prediabetes; SSB, sugar-sweetened beverages; TAG, tri- glycerides; WC, waist circumference.

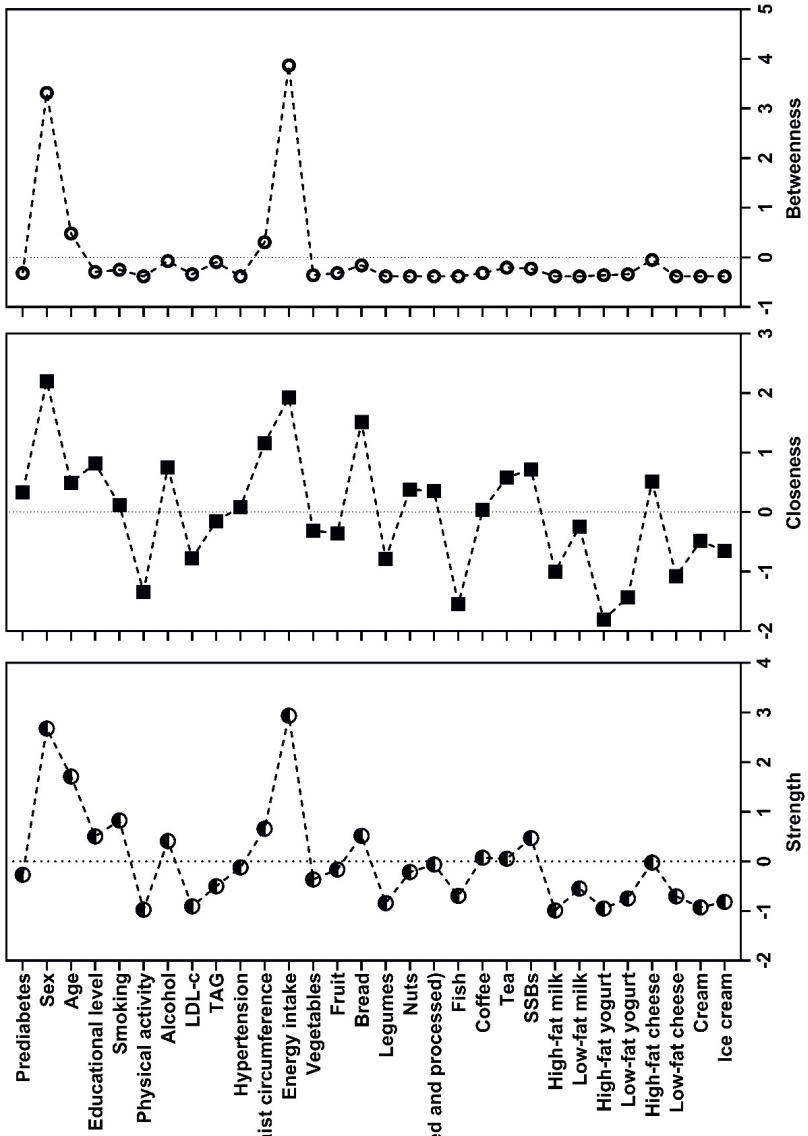


Figure 5. Standardized centrality of each variable in the network as indicated by the strength (i.e., the sum of the absolute connection weights of a node, with higher values indicating stronger connectivity of a node), closeness (i.e., inverse sum of the distance between one node and all other nodes) and betweenness (i.e., number of times the node is in the shortest paths between other nodes). Abbreviations: LDL-C, low-density lipoprotein cholesterol; SSB, sugar-sweetened beverages; TAG, triglycerides.

Discussion

In this study, total dairy, fermented dairy, and cheese, irrespective of fat content, cream, and ice cream intake were not associated with the risk of prediabetes. In regression analysis, plain and low-fat milk intake were associated with a higher risk of prediabetes in the top compared with the bottom quartile. The network of interrelations between dairy, risk factors, and dietary intake showed that low-fat milk connected with energy-dense foods, including bread, meat, and high-fat cheese. High-fat yogurt intake was associated with a lower risk of prediabetes; however, this association was no longer significant after further adjustment for intake of other food groups. High-fat yogurt had no clear role in the networks, but low-fat yogurt was linked to healthy food groups and physical activity in network analysis, although associations in regression analyses were neutral.

Findings in context

We found that a higher intake of low-fat milk was associated with higher incidence of prediabetes. This was in line with the Dutch Rotterdam Study ($HR_{serving/day}$ 1.07; 95%CI 1.01-1.13) [12] and the Australian Diabetes, Obesity and Lifestyle (AusDiab) study (RR highest compared with lowest tertile 1.15 95%CI 0.97-1.35; P_{trend} = 0.04) [13]. However, no associations between milk intake and prediabetes were found in the Hoorn Studies [11] and in cross-sectional studies with prediabetes or glucose measurements as outcomes [7-9, 69, 70]. A low centrality and predictability of low-fat milk intake in the networks may indicate that significant associations in the regression analysis could be due to intake levels coinciding with influential risk factors, such as energy intake and waist circumference. In contrast to our current findings, a protective nonlinear association of low-fat milk was found in the US Framingham Offspring Study (*n* = 1867, 12 years follow-up, HR ≥ 14 compared with 0 servings/wk 0.84; 95%CI 0.62, 1.12, P_{trend} = 0.49) [10]. In the United States, milk consumers might be healthier than those in Europe, where milk is consumed by a wider range of populations [71]. So far, evidence for a causal relation between milk and glucose metabolism from randomized controlled trials is lacking [72, 73]. Furthermore, Mendelian randomization studies with T2D as the outcome showed no association [74-76].

Exploratory subgroup analyses demonstrated that high-fat yogurt intake was associated with a lower risk of prediabetes only in participants with a high educational level. This could be due to the presence of multiple risk factors among people with lower educational levels, resulting in higher T2D risk regardless of dairy intake [46]. Furthermore, populations with lower education may be more likely to misreport, resulting in the attenuation of associations [77]. An inverse association for high-fat yogurt intake and risk of prediabetes was also found in the Rotterdam Study ($RR_{serving/day}$ 0.67; 95%CI 0.51, 0.89) [12] but not the Hoorn Studies [11]. In contrast, low-fat yogurt intake was not associated with risk of prediabetes in this study and in previous cohorts [11, 12]. Several meta-analyses showed that total yogurt intake is associated with lower

T2D risk (RRs ranging from 0.74–0.86); however, no distinction is made between low-fat and high-fat yogurt [78]. A cluster of low-fat yogurt, fruit, vegetables, fish, legumes, and physical activity was identified in our study, which also connected to educational level and inversely to SSBs. This is in line with prior studies showing that low-fat yogurt intake relates to overall healthier behaviours [14-17]. Two prior studies using network modelling found that milk and dairy intake connected positively to fruit intake, but they did not assess specific dairy types [22, 23]. This clustering of low-fat yogurt intake with healthy behaviours could also reflect reporting bias of socially desirable behaviours. Further studies are required to establish whether previously reported associations are due to residual confounding or if causal mechanisms link yogurt to diabetes.

We found that participants with a lower intake of high-fat dairy types and a higher intake of low-fat dairy types were characterized by a higher baseline diabetes risk and more often had a desire to lose weight. Awareness of risk could result from a family history of the disease, being overweight or obese, public health campaigns, or opportunistic screening by general practitioners. This awareness, or desire to lose weight, might result in better adherence to dietary guidelines recommending low-fat dairy types to limit saturated fat and caloric intake [24]. Thus, reverse causation may explain the inverse associations with high-fat dairy types with prediabetes, and positive associations with low-fat dairy.

The network figure showed a clustering of clinical markers, co-occurring within the metabolic syndrome, explaining the minor changes in effect estimates of regression models with additional adjustment for clinical markers. Waist circumference and age connected this prediabetes cluster to sociodemographic and health risk factors, which were in turn connected to food group and dairy intake. This highlights the role of obesity as a gateway between behaviour and clinical outcomes. In the Rotterdam Study, mainly weight development mediated associations of dairy intake and prediabetes [12]. Age is a crucial confounder in these associations, especially relevant in cohorts with wide age range, such as the Lifelines Cohort.

Strengths and limitations

A major strength of the current study is the large sample size and relatively young population (mean age 45.5 ± 12.3 y), thus including more participants susceptible to developing prediabetes compared with earlier studies with middle-aged samples [10-12, 25]. Complementary to the regression analysis, the network analysis helped to visualize the system of interrelated risk factors and behaviours in this population [19, 20]. With the networks, we could highlight the bridging role of obesity and certain behavioural patterns linked to dairy, potentially interesting for further research.

The results should be interpreted considering several limitations. First, the FFQ was administered at baseline only, and dietary changes during follow-up are plausible.

However, the mean follow-up period was 4 y, and prior studies showed that dietary patterns were relatively stable over follow-up periods of ≤5 years [79, 80]. Second, misclassification of the exposure is possible as the FFQ is self-report and prone to measurement errors. Because the FFQ is suited for ranking of participants, associations were presented with the dairy exposure both continuously and categorized. Nevertheless, artificial categorization might result in bias and less power compared with continuous analysis [81], advocating the need to improve the accuracy of dietary assessment. In the network structure, weaker connections for health risk factors and food groups compared with clinical markers may reflect a greater extent of measurement uncertainty. Third, extensive phenotyping has been done in the Lifelines Cohort, enabling correction for many known confounders in regression and network analysis; nevertheless, residual confounding can never be ruled out. Lastly, the Lifelines Cohort is representative of the general population of the Northern Netherlands [82], but generalizability to other populations might be limited. Thus, reproducibility in other samples is advocated, especially in non-White and non-Western regions.

Conclusion

In conclusion, in this large, population-based prospective cohort of Dutch adults, low-fat milk intake was associated with higher risk of prediabetes. Inconsistent associations between low-fat and high-fat dairy types might partly be attributed to reverse causation. Furthermore, network analysis showed clustering of risk factors and behaviours in relation to prediabetes, with low influence of dairy intake. Future research should account for this complex system of interacting risk factors. Nevertheless, well-designed randomized controlled trials are needed to fully elucidate the health effects of dairy, personalized on relevant risk modifiers.

List of supplementary materials chapter 5

Supplemental table 1. Missing values of covariates in participants before imputation (*n* = 74,132).

Supplemental figure 1. Distribution of dairy food intake in the diet (weight percentage, %) (*n* = 74,132).

Supplemental table 2. Baseline characteristics of participants in the Lifelines cohort across different population subgroups (n total = 74,132).

Supplemental table 3. Associations of dairy product types and prediabetes risk in the Lifelines cohort, stratified by sex, age, baseline waist circumference and educational level (*n* = 74,132).

Supplemental table 4. Sensitivity analyses of associations of dairy product types and prediabetes in the Lifelines cohort.

Supplemental table 5. Diabetes risk at baseline based on PROCAM diabetes risk score and desire to lose weight according to top and bottom quartiles of total dairy and dairy type intake.

Supplemental table 6. Cross-sectional associations between diabetes risk and dairy intake at baseline (*n* = 72,615).

Supplemental table 7. Cross-sectional associations between desire to lose weight and dairy intake at baseline (*n* = 74,001).

Supplemental table 8. Associations of dairy product types and prediabetes in the Lifelines cohort, additionally adjusted for 'desire to lose weight' (*n* = 74,001).

Supplemental figure 2. Clusters of variables (represented by the different colours) that strongly connect in the network.

Supplemental figure 3. Network structure of low-fat yogurt, food groups and energy intake, lifestyle risk factors, socio-demographic characteristics, clinical markers and prediabetes of the study population with complete data for variables in the model (*n* = 67,206).

Supplemental figure 4. Plot of bootstrapped sampling variation around the edge-weights reflecting accuracy of the edge-weights, a measure of the stability of the network.

Supplemental figure 5. Average correlation coefficients between centrality indices of the networks within the case-dropped bootstrapped samples and the original sample reflecting the stability of the centrality indices.

Supplemental figure 6. Network structure of dairy, food groups and energy intake, lifestyle risk factors, socio-demographic characteristics, clinical markers and prediabetes of the study population with complete data for variables in the model with a nonparanormal transformation applied to non-normally distributed continuous variables (*n* = 67,206).

Scan this QR code to download the supplementary materials.

References

1. World Health Organization (WHO), Definition and diagnosis of diabetes mellitus and intermediate hyperglycaemia: report of a WHO/IDF consultation. 2006.
2. Gerstein, H.C., et al., Annual incidence and relative risk of diabetes in people with various categories of dysglycemia: a systematic overview and meta-analysis of prospective studies. Diabetes Res Clin Pract, 2007. **78**(3): p. 305-12.
3. Yeboah, J., et al., Impaired fasting glucose and the risk of incident diabetes mellitus and cardiovascular events in an adult population: MESA (Multi-Ethnic Study of Atherosclerosis). Journal of the American College of Cardiology, 2011. **58**(2): p. 140-146.
4. Mutie, P.M., et al., An investigation of causal relationships between prediabetes and vascular complications. Nat Commun, 2020. **11**(1): p. 4592.
5. Cai, X., et al., Association between prediabetes and risk of all cause mortality and cardiovascular disease: updated meta-analysis. Bmj, 2020. **370**: p. m2297.
6. Soedamah-Muthu, S.S. and J. de Goede, Dairy Consumption and Cardiometabolic Diseases: Systematic Review and Updated Meta-Analyses of Prospective Cohort Studies. Curr Nutr Rep, 2018. **7**(4): p. 171-182.
7. Eussen, S.J., et al., Consumption of dairy foods in relation to impaired glucose metabolism and type 2 diabetes mellitus: the Maastricht Study. Br J Nutr, 2016. **115**(8): p. 1453-61.
8. Brouwer-Brolsma, E.M., et al., Dairy product consumption is associated with pre-diabetes and newly diagnosed type 2 diabetes in the Lifelines Cohort Study. Br J Nutr, 2018. **119**(4): p. 442-455.
9. Breuninger, T.A., et al., Differential associations between diet and prediabetes or diabetes in the KORA FF4 study. J Nutr Sci, 2018. **7**: p. e34.
10. Hruby, A., et al., Associations of Dairy Intake with Incident Prediabetes or Diabetes in Middle-Aged Adults Vary by Both Dairy Type and Glycemic Status. J Nutr, 2017. **147**(9): p. 1764-1775.
11. Slurink, I.A.L., et al., Dairy product consumption and incident prediabetes in Dutch middle-aged adults: the Hoorn Studies prospective cohort. Eur J Nutr, 2022. **61**(1): p. 183-196.
12. Slurink, I.A.L., et al., Dairy Product Consumption in Relation to Incident Prediabetes and Longitudinal Insulin Resistance in the Rotterdam Study. Nutrients, 2022. **14**(3).
13. Slurink, I.A., et al., Dairy Product Consumption and Incident Prediabetes in the Australian Diabetes, Obesity, and Lifestyle Study With 12 Years of Follow-Up. J Nutr, 2023. **153**(6): p. 1742-1752.
14. Sun, Q. and F. Hu, Yogurt consumption and health outcomes: a review of epidemiologic evidence. Functional Food Reviews, 2013. **5**(2): p. 52-61.
15. Martinez-Gonzalez, M.A., et al., Yogurt consumption, weight change and risk of overweight/obesity: the SUN cohort study. Nutr Metab Cardiovasc Dis, 2014. **24**(11): p. 1189-96.
16. Wang, H., et al., Yogurt consumption is associated with better diet quality and metabolic profile in American men and women. Nutr Res, 2013. **33**(1): p. 18-26.
17. Possa, G., J.E. Corrente, and M. Fisberg, Yogurt consumption is associated with a better lifestyle in Brazilian population. BMC Nutr, 2017. **3**: p. 29.
18. Dalege, J., et al., Toward a formalized account of attitudes: The Causal Attitude Network (CAN) model. Psychol Rev, 2016. **123**(1): p. 2-22.
19. Iqbal, K., et al., Gaussian graphical models identified food intake networks and risk of type 2 diabetes, CVD, and cancer in the EPIC-Potsdam study. European Journal of Nutrition, 2019. **58**(4): p. 1673-1686.
20. Jahanmiri, R., et al., Saturated fats network identified using Gaussian graphical models is associated with metabolic syndrome in a sample of Iranian adults. Diabetology & Metabolic Syndrome, 2022. **14**(1): p. 1-11.
21. Hoang, T., J. Lee, and J. Kim, Differences in Dietary Patterns Identified by the Gaussian Graphical Model in Korean Adults With and Without a Self-Reported Cancer Diagnosis. J Acad Nutr Diet, 2021. **121**(8): p. 1484-1496.e3.

22. Schwedhelm, C., et al., Meal and habitual dietary networks identified through semiparametric Gaussian copula graphical models in a German adult population. PLoS One, 2018. **13**(8): p. e0202936.
23. Hoang, T., J. Lee, and J. Kim, Network Analysis of Demographics, Dietary Intake, and Comorbidity Interactions. Nutrients, 2021. **13**(10): p. 3563.
24. Comerford, K.B., et al., Global Review of Dairy Recommendations in Food-Based Dietary Guidelines. Front Nutr, 2021. **8**: p. 671999.
25. Scholtens, S., et al., Cohort Profile: LifeLines, a three-generation cohort study and biobank. Int J Epidemiol, 2015. **44**(4): p. 1172-80.
26. Rhee, J.J., et al., Comparison of methods to account for implausible reporting of energy intake in epidemiologic studies. American journal of epidemiology, 2015. **181**(4): p. 225-233.
27. Willett, W.C., Nutritional Epidemiology. 1998, New York: Oxford University Press.
28. International Expert Committee report on the role of the A1C assay in the diagnosis of diabetes. Diabetes Care, 2009. **32**(7): p. 1327-34.
29. Brouwer-Brolsma, E.M., et al., Development and external validation of the 'Flower-FFQ': a FFQ designed for the Lifelines Cohort Study. Public Health Nutr, 2022. **25**(2): p. 225-236.
30. Siebelink, E., A. Geelen, and J.H.M. de Vries, Self-reported energy intake by FFQ compared with actual energy intake to maintain body weight in 516 adults. British Journal of Nutrition, 2011. **106**(2): p. 274-281.
31. The Dutch National Institute for Public Health and the Environment (RIVM), NEVO-Table. Dutch Food Composition Database 2011., Dutch Nutrition Centre, Editor. 2011: Den Haag.
32. Vinke, P.C., et al., Development of the food-based Lifelines Diet Score (LLDS) and its application in 129,369 Lifelines participants. European Journal of Clinical Nutrition, 2018. **72**(8): p. 1111-1119.
33. Dutch National Food Consumption Survey 2007-2010 | Part 2 Total Foods [Internet]. 2010. p. 1157–3368. Available from: *http://www.rivm.nl/Documenten_en_publicaties/Wetenschappelijk/Tabellen_grafieken/Leefstijl_Voeding/VCP/Basis_2011/VCP_2007_2010_Deel_2_Voedingsmiddelen_NEVO_codes/Download/VCP_2007_2010_Deel_2_Voedingsmiddelen_NEVO_codes.or*
34. Jansen, H., et al., Determinants of HbA1c in nondiabetic Dutch adults: genetic loci and clinical and lifestyle parameters, and their interactions in the Lifelines Cohort Study. J Intern Med, 2013. **273**(3): p. 283-93.
35. American Diabetes Association Professional Practice Committee, 2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes—2022. Diabetes Care, 2021. **45**(Supplement_1): p. S17-S38.
36. van Herpt, T.T., et al., Lifetime risk to progress from pre-diabetes to type 2 diabetes among women and men: comparison between American Diabetes Association and World Health Organization diagnostic criteria. BMJ Open Diabetes Research and Care, 2020. **8**(2): p. e001529.
37. Vistisen, D., et al., Risk of cardiovascular disease and death in individuals with prediabetes defined by different criteria: the Whitehall II study. Diabetes Care, 2018. **41**(4): p. 899-906.
38. Byambasukh, O., Physical activity and cardiometabolic health: Focus on domain-specific associations of physical activity over the life course. 2020.
39. Wendel-Vos, G.C., et al., Reproducibility and relative validity of the short questionnaire to assess health-enhancing physical activity. J Clin Epidemiol, 2003. **56**(12): p. 1163-9.
40. von Eckardstein, A., H. Schulte, and G. Assmann, Risk for Diabetes Mellitus in Middle-Aged Caucasian Male Participants of the PROCAM Study: Implications for the Definition of Impaired Fasting Glucose by the American Diabetes Association. The Journal of Clinical Endocrinology & Metabolism, 2000. **85**(9): p. 3101-3108.
41. Knol, M.J., et al., Overestimation of risk ratios by odds ratios in trials and cohort studies: alternatives to logistic regression. CMAJ, 2012. **184**(8): p. 895-9.
42. Zou, G., A modified poisson regression approach to prospective studies with binary data. American journal of epidemiology, 2004. **159**(7): p. 702-706.
43. Zou, G., A modified poisson regression approach to prospective studies with binary data. Am J Epidemiol, 2004. **159**(7): p. 702-6.
44. Dutch Health Council (Gezondheidsraad), Dutch dietary guidelines 2015 (Richtlijnen goede voeding 2015). Publication nr. 2015/24. ISBN 978-94-6281-089-1. The Hague. 2015.

45. Schwingshackl, L., et al., Food groups and risk of type 2 diabetes mellitus: a systematic review and meta-analysis of prospective studies. Eur J Epidemiol, 2017. **32**(5): p. 363-375.
46. Vinke, P.C., et al., Socio-economic disparities in the association of diet quality and type 2 diabetes incidence in the Dutch Lifelines cohort. EClinicalMedicine, 2020. **19**: p. 100252.
47. Willett, W.C., G.R. Howe, and L.H. Kushi, Adjustment for total energy intake in epidemiologic studies. The American journal of clinical nutrition, 1997. **65**(4): p. 1220S-1228S.
48. Haslbeck, J.M.B. and L.J. Waldorp, mgm: Estimating Time-Varying Mixed Graphical Models in High-Dimensional Data. Journal of Statistical Software, 2020. **93**(8): p. 1 - 46.
49. Epskamp, S., et al., qgraph: Network Visualizations of Relationships in Psychometric Data. Journal of Statistical Software, 2012. **48**(4): p. 1 - 18.
50. Epskamp, S., D. Borsboom, and E.I. Fried, Estimating psychological networks and their accuracy: A tutorial paper. Behavior Research Methods, 2018. **50**(1): p. 195-212.
51. Baba, K., R. Shibata, and M. Sibuya, Partial Correlation and Conditional Correlation as Measures of Conditional Independence. Australian & New Zealand Journal of Statistics, 2004. **46**(4): p. 657-664.
52. Lauritzen, S.L., Graphical models. Vol. 17. 1996: Clarendon Press.
53. Greenland, S., Quantifying biases in causal models: classical confounding vs collider-stratification bias. Epidemiology, 2003. **14**(3): p. 300-306.
54. Tibshirani, R., Regression shrinkage and selection via the lasso. 2011.
55. Foygel, R. and M. Drton, Extended Bayesian information criteria for Gaussian graphical models. Advances in neural information processing systems, 2010. **23**.
56. Epskamp, S. and E.I. Fried, A tutorial on regularized partial correlation networks. Psychological methods, 2018. **23**(4): p. 617.
57. Borsboom, D. and A.O. Cramer, Network analysis: an integrative approach to the structure of psychopathology. Annu Rev Clin Psychol, 2013. **9**: p. 91-121.
58. Haslbeck, J. and L.J. Waldorp, How well do network models predict observations? On the importance of predictability in network models. Behavior research methods, 2018. **50**(2): p. 853-861.
59. Csardi, G. and T. Nepusz, The igraph software package for complex network research. InterJournal, complex systems, 2006. **1695**(5): p. 1-9.
60. Boccaletti, S., et al., Complex networks: Structure and dynamics. Physics Reports, 2006. **424**(4): p. 175-308.
61. Briganti, G., et al., Network analysis of empathy items from the interpersonal reactivity index in 1973 young adults. Psychiatry Res, 2018. **265**: p. 87-92.
62. Bringmann, L.F., et al., What do centrality measures measure in psychological networks? Journal of abnormal psychology, 2019. **128**(8): p. 892.
63. Epskamp, S., Network Psychometrics (Doctoral dissertation). , ed. A.U.o. Amsterdam. 2017.
64. Rodebaugh, T.L., et al., Does centrality in a cross-sectional network suggest intervention targets for social anxiety disorder? Journal of consulting and clinical psychology, 2018. **86**(10): p. 831.
65. Epskamp, S., Psychometric network models from time-series and panel data. Psychometrika, 2020. **85**(1): p. 206-231.
66. Liu, H., J. Lafferty, and L. Wasserman, The nonparanormal: Semiparametric estimation of high dimensional undirected graphs. Journal of Machine Learning Research, 2009. **10**(10).
67. Liu, H., et al., High-dimensional semiparametric Gaussian copula graphical models. 2012.
68. Jiang, H., et al., huge: High-Dimensional Undirected Graph Estimation. R package version 1.3.5. *https://CRAN.R-project.org/package=huge*. 2021.
69. Struijk, E.A., et al., Dairy product intake in relation to glucose regulation indices and risk of type 2 diabetes. Nutr Metab Cardiovasc Dis, 2013. **23**(9): p. 822-8.
70. Drehmer, M., 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): p. 775-82.

71. U.S. Department of Health and Human Services and U.S. Department of Agriculture, 2015 – 2020 Dietary Guidelines for Americans. 8th Edition (2015). Available at https:// health. gov/ our- work/ food- nutri tion/ previ ous- dieta ry- guide lines/ 2015
72. Guo, J., et al., The Impact of Dairy Products in the Development of Type 2 Diabetes: Where Does the Evidence Stand in 2019? Adv Nutr, 2019. **10**(6): p. 1066-1075.
73. Eelderink, C., et al., The effect of high compared with low dairy consumption on glucose metabolism, insulin sensitivity, and metabolic flexibility in overweight adults: a randomized crossover trial. The American journal of clinical nutrition, 2019. **109**(6): p. 1555-1568.
74. Vimaleswaran, K.S., et al., Evidence for a causal association between milk intake and cardiometabolic disease outcomes using a two-sample Mendelian Randomization analysis in up to 1,904,220 individuals. International Journal of Obesity, 2021. **45**(8): p. 1751-1762.
75. Vissers, L.E.T., et al., Dairy Product Intake and Risk of Type 2 Diabetes in EPIC-InterAct: A Mendelian Randomization Study. Diabetes Care, 2019. **42**(4): p. 568-575.
76. Bergholdt, H.K., B.G. Nordestgaard, and C. Ellervik, Milk intake is not associated with low risk of diabetes or overweight-obesity: a Mendelian randomization study in 97,811 Danish individuals. The American journal of clinical nutrition, 2015. **102**(2): p. 487-496.
77. Calvert, C., et al., Using cross-check questions to address the problem of mis-reporting of specific food groups on food frequency questionnaires. European journal of clinical nutrition, 1997. **51**(10): p. 708-712.
78. Alvarez-Bueno, C., et al., Effects of Milk and Dairy Product Consumption on Type 2 Diabetes: Overview of Systematic Reviews and Meta-Analyses. Adv Nutr, 2019. **10**(suppl_2): p. S154-S163.
79. Jankovic, N., et al., Stability of dietary patterns assessed with reduced rank regression; the Zutphen Elderly Study. Nutrition journal, 2014. **13**(1): p. 1-9.
80. Nagel, G., et al., Long-term reproducibility of a food-frequency questionnaire and dietary changes in the European Prospective Investigation into Cancer and Nutrition (EPIC)-Heidelberg cohort. Br J Nutr, 2007. **98**(1): p. 194-200.
81. DeCoster, J., M. Gallucci, and A.-M.R. Iselin, Best practices for using median splits, artificial categorization, and their continuous alternatives. Journal of experimental psychopathology, 2011. **2**(2): p. 197-209.
82. Klijs, B., et al., Representativeness of the LifeLines Cohort Study. PLOS ONE, 2015. **10**(9): p. e0137203.

Chapter 6

Dairy intake and risk of prediabetes and type 2 diabetes in the Fenland study

Manuscript based on this chapter:

Isabel A.L Slurink, Nina Kupper, Tom Smeets, and Sabita S. Soedamah-Muthu. Dairy consumption and risk of prediabetes and type 2 diabetes in the Fenland study. Clinical Nutrition, Volume 43, Issue 11, 69 - 79.

Abstract

Background

Limited observational evidence suggests that a higher intake of high-fat dairy may be associated with lower prediabetes risk, while opposite associations have been observed for low-fat milk intake. This study aimed to examine associations between baseline and changes in dairy consumption, risk of prediabetes, and glycaemic status.

Methods

7,521 participants from the prospective UK Fenland study were included (mean age 48.7 ± 2.0 years, 51.9% female). Dairy intake was measured using self-reported food frequency questionnaires. Associations with prediabetes risk and glycaemic status were analysed using Poisson regression models adjusted for social demographics, health behaviours, family history of diabetes and food group intake.

Results

At a mean follow-up of 6.7 ± 2.0 years, 290 participants developed prediabetes (4.3%). Most dairy products were not significantly associated with prediabetes risk. A higher baseline intake of high-fat dairy ($RR_{serving/day}$ 1.20, 95%CI 1.03-1.39) and high-fat milk ($RR_{serving/day}$ 1.22, 1.01-1.47) were associated with higher prediabetes risk. Conversely, low-fat milk was associated with lower prediabetes risk ($RR_{serving/day}$ 0.86, 0.75-0.98). In the analyses evaluating dietary changes over time, increases in high-fat milk were inversely associated with risk of progressing from normoglycaemia to prediabetes or type 2 diabetes ($RR_{servings/day}$ 0.86, 95%CI 0.75-0.99).

Conclusions

This population-based study showed that most dairy products are not associated with prediabetes risk or progression in glycaemic status. Positive associations of high-fat dairy, high-fat milk, and the inverse association of low-fat milk with prediabetes risk found were inconsistent with prior literature and suggestive of the need for future research on environmental, behavioural, and biological factors that explain the available evidence.

Introduction

Prediabetes is defined as the intermediate stage between normoglycaemia and diabetes, including impaired fasting glucose and/or impaired glucose tolerance [1]. People in this early risk stage of type 2 diabetes (T2D) already display metabolic disturbances and are at high risk for microvascular and macrovascular complications [2-5]. The prevalence of prediabetes and T2D is increasing worldwide, [6, 7], especially a high prevalence is observed in people with obesity and of older age [8, 9]. Early-stage prevention is needed because a considerable proportion of people with prediabetes develop incident T2D (9-84%, depending on follow-up duration and prediabetes definition) [10] and cardiovascular diseases [5]. Nevertheless, people with prediabetes may revert to normoglycaemia [10, 11]. Improving dietary and physical activity behaviours is the recommended approach to prevent and manage prediabetes, and effectiveness has been shown in randomized controlled trials [12, 13].

Dairy products are widely consumed by many populations. Yet, dietary clinical trial studies show heterogeneous health effects of dairy, plausibly due to differences in study populations and variations in consumed dairy types [14, 15]. Few prospective cohort studies have presented conflicting results regarding the association between dairy intake and incident prediabetes [16-20]. Some beneficial associations were observed for high-fat dairy types and prediabetes risk [16-18, 20], while unfavourable associations for low-fat milk intake were found [18-20]. To our knowledge, no study has evaluated the association of changes in dairy consumption with prediabetes risk, while analysing changes in a dietary exposure may help characterise a causal association better than analysing a dietary exposure at a single time point [21]. Overall, additional studies are needed on this topic, especially analyses in cohort studies with repeated assessments of dairy intake. Therefore, this study aimed to examine associations between baseline and changes in dairy consumption (quantity and type, e.g. milk, yogurt, and cheese), risks of prediabetes, and change in glycaemic status (normoglycaemia, prediabetes and T2D) in the UK prospective Fenland study.

Methods

Study design and population

The Fenland study is an ongoing cohort study of people born between 1950 and 1975 in Cambridgeshire, UK, aiming to investigate how environmental and genetic factors contribute to metabolic disorders. Baseline measurements (Phase 1) were conducted between 2005 and 2015 (n = 12,434, response rate: 27%). Eligible participants were recruited by their general practice to attend one of the clinic sites at Ely, Wisbech, or Cambridge, UK, where assessment took place including self-reported questionnaires, clinical measurements, and blood tests for biochemical measurements. Exclusion criteria were a known history of diabetes, psychotic or terminal illness, inability to walk

unaided and pregnancy or lactation. Each participant underwent a single follow-up measurement (Phase 2) between 2014 and 2020 (*n* = 7,795, follow-up rate: 62.7%). The study was approved by the Cambridge Local Ethics Committee. All participants gave written informed consent.

For the current analysis, we excluded participants without the Phase 2 assessment (*n* = 4,608), pregnant women (*n* = 9), and participants with missing dietary data at baseline (*n* = 5), energy intakes at baseline outside the sex-specific ranges (<500 and ≥3500 kcal/day in women and <800 and ≥4000 kcal/day in men) considered as implausible [22] (*n* = 115), T2D at baseline (*n* = 116), missing data for glycaemic outcomes at baseline (*n* = 25), or missing glycaemic outcomes at follow-up (*n* = 34), resulting in 7,521 participants (**Figure 1**). This subset was used for descriptive analyses. For analysis with prediabetes risk as outcome, we further excluded participants with prediabetes (*n* = 819) at baseline, and T2D at follow-up (*n* = 63), resulting in 6,639 participants. For the analyses where changes in dairy intake were modelled as exposure variables, we excluded participants with missing dietary data (*n* = 22) or implausible energy intake at follow-up (*n* = 89), resulting in 7,410 participants.

Ascertainment of prediabetes and type 2 diabetes

At baseline and follow-up clinic visits, fasting blood samples were drawn. Furthermore, a standard oral-glucose-tolerance test (OGTT) using a 75-g glucose drink was administered to obtain the two-hour plasma glucose (2hPG). Fasting plasma glucose (FPG) concentrations were assessed using the hexokinase method. HbA1c was measured with ion-exchange high-performance liquid chromatography. At baseline and follow-up, prediabetes was defined based on FPG levels between 110 and 125 mg/dL (6.1 and 6.9 mmol/L), 2hPG levels between 140 and 199 mg/dL (7.8 and 11.0 mmol/L), or HbA1c levels between 42 and 48 mmol/mol (6.0 and 6.5%) as defined by the World Health Organization and the International Expert Committee (WHO-IEC) [1, 23]. T2D at baseline was defined as self-reported usage of oral antidiabetic agents, FPG levels ≥7.0 mmol/L, 2hPG levels ≥11.1 mmol/L, or HbA1c levels ≥48 mmol/mol. New-onset T2D during follow-up was ascertained if a participant met any of the following criteria: ICD-10 codes, physician diagnosis of T2D, self-reported diagnosis, self-reported usage of oral antidiabetic agents, FPG ≥7.0 mmol/L, 2-h glucose ≥11.1 mmol/L or HbA1c ≥48 mmol/mol. We additionally defined prediabetes based on three different definitions as many studies do not have all three glycaemic measures available when defining prediabetes. Studies utilize an FPG-only based definition, or FPG in combination with either 2hPG or HbA1c, as the OGTT measurement is costly and burdensome. We also based the participant selection on the assumption that only data for certain glycaemic measures was available. This affected the number of participants with prevalent prediabetes and T2D at baseline, T2D at follow-up, and the number of participants with missing data for glycaemic markers at baseline and follow-up (**Table 1**).

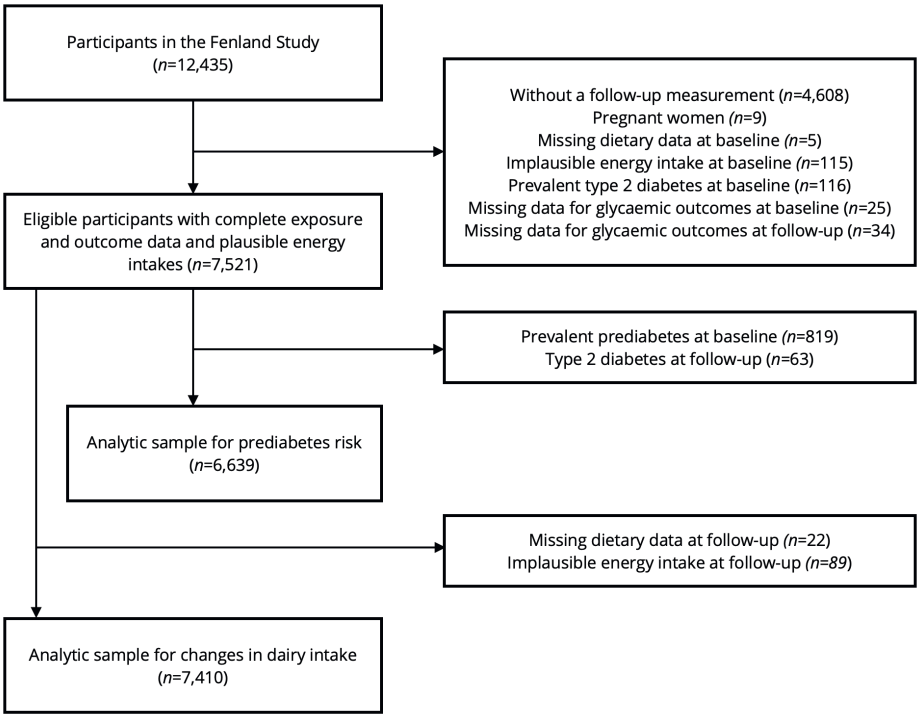


Figure 1. Flow-chart for inclusion of participants for the present analysis of the Fenland study.

Table 1. Inclusion of participants and prediabetes incidence based on different prediabetes outcome definitions in the Fenland study.

	Glycaemic markers used to define prediabetes			
	FPG, 2hPG & HbA1c	FPG only	FPG & 2hPG	FPG & HbA1c
Prevalent type 2 diabetes at baseline	116	51	81	99
Prevalent prediabetes at baseline	819	107	436	566
Type 2 diabetes at follow-up	63	106	80	75
Additional missing values in glycaemic markers	-	68	17	14
Analytical sample for prediabetes risk	6639	7305	7024	6888
Prediabetes incidence	290 (4.4%)	106 (1.5%)	254 (3.6%)	192 (2.8%)

Abbreviations: 2hPG, 2-hour Plasma Glucose; FPG, Fasting Plasma Glucose; HbA1c, Glycated Haemoglobin.

Dietary assessment

Habitual dietary data over the previous year were measured at baseline and follow-up using a 130-item semi-quantitative food frequency questionnaire (FFQ) [24]. The in-house software was used to estimate the daily food and nutrient intakes, in gram/day, from the FFQs, using food composition data from McCance and Widdowson's Composition of foods (5th Edition) and its supplements [25]. Spearman correlation coefficients between the FFQ and 16-day weighed records were previously reported as 0.55 and 0.56 for the estimated intakes of total fat and saturated fat, respectively, indicating moderate validity [26, 27]. A previous study reported that the correlation coefficients between the FFQ and 7-day food diaries were 0.56 for milk, 0.57 for yogurt and 0.33 for cheese [28].

The FFQ included standard questions on different dairy types, a free-text question asking types of milk consumed regularly, and a question asking milk consumption with milk with tea, coffee, and cereals. Dairy foods were assessed in servings/day and categorized as milk (any types of cow's milk), yogurt, cheese, cream, and ice cream, of which serving sizes followed the previous dairy analysis in the Fenland study [29] (**Table 2**). Each dairy category was further divided into low-fat (liquid products <2%, cheese <20%) and high-fat (liquid products ≥2%, cheese ≥20%). At baseline and follow-up, 7.2% and 21.5% of participants consumed milk for whom we could not determine the fat content. Therefore, for those participants, intakes of non-specific milk were divided equally into high-fat and low-fat milk intakes. Butter was not included in the dairy group in line with previous research [30, 31] and analysed as a separate food group (33).

Data on total energy (kcal/day), alcohol intake (g/day), the intake of macronutrients, added sugar, calcium, sodium, and food groups (fruits, vegetables, whole grains, refined grains, potatoes, legumes, nuts, tea, coffee, red meat, processed meat, fish, and sugar-sweetened beverages (SSBs)) were also obtained from the FFQ. The intakes of macronutrients (fat, saturated fatty acids (SFAs), protein and carbohydrates) were expressed as percentage from total energy (en%). Furthermore, the Dietary Approaches to Stop Hypertension (DASH) adherence score and the Mediterranean diet score (MDS) were calculated based on the FFQ as measures of diet quality [32-34] (details in Supplementary Materials).

Dietary quality

The Dietary Approaches to Stop Hypertension (DASH) was adapted from Fung *et al.* and consisted of eight dietary components (grains/grain products, vegetables, fruits, low-fat/fat-free dairy, red and processed meat, nuts/seeds/dry beans, dietary sodium, and foods high in added sugar) [32]. The DASH score was calculated by summing the quantile scores of each energy adjusted dietary component, ranging between 8 (least healthy) and 40 (most healthy). The tertiles Mediterranean diet score (MDS) was designed to reflect a degree of dietary adherence to the Mediterranean dietary pyramid proposed by the Mediterranean Diet Foundation for both Mediterranean and non-

Mediterranean countries [33, 34]. The MDS reflected consumption levels of fifteen food groups: vegetables, legumes, and fish as healthy food groups; red meat, processed meat, potato, and sweets as unhealthy food groups, and fruits, cereals, nuts, eggs, dairy, white meat, and alcoholic beverages as food groups for which the pyramid recommended a moderate intake. A continuous score from 0 to 1 was assigned to each of the components according to the recommended consumption level, and a total score could range from 0 to 15 points.

Table 2. Food items included in total dairy and dairy types and consumption in the Fenland study (*n* = 7,521).

Dairy type	Definition	Consumers	Baseline	Follow-up	Change
		%	Servings/day	Servings/day	Servings/day
Primary exposures					
Milk	High-fat milk; low-fat milk	92.4%	1.44 ± 0.87	1.25 ± 0.85	-0.19 ± 0.82
High-fat	Full cream, silver; Channel Islands, gold, non-specific milk (50%) of intake	16.0%	0.19 ± 0.53	0.25 ± 0.52	-0.02 ± 0.5[1]
Low-fat	Semi-skimmed, red/white; dried milk, non-specific milk (50%) of intake	85.1%	1.25 ± 0.90	1.00 ± 0.82	-0.18 ± 0.88[1]
Yogurt	High-fat yogurt; low-fat yogurt	78.4%	0.42 ± 0.49	0.46 ± 0.54	0.04 ± 0.56
High-fat	Full fat or Greek yogurt	34.7%	0.08 ± 0.19	0.15 ± 0.30	0.07 ± 0.31
Low-fat	Low-fat yogurt, fromage frais	70.1%	0.34 ± 0.46	0.31 ± 0.46	-0.03 ± 0.50
Cheese	High-fat cheese; low-fat cheese	98.8%	0.55 ± 0.45	0.54 ± 0.48	-0.01 ± 0.51
High-fat	Cheese, e.g. Cheddar, Brie, Edam	98.4%	0.44 ± 0.36	0.44 ± 0.37	0.00 ± 0.38
Low-fat	Cottage cheese, low-fat soft cheese	38.2%	0.11 ± 0.27	0.11 ± 0.29	-0.01 ± 0.34
Secondary exposures					
Total dairy products	High-fat dairy; low-fat dairy	99.7%	2.64 ± 1.20	2.48 ± 1.24	-0.16 ± 1.20
High-fat	High-fat milk; high-fat yogurt; high-fat cheese; total cream; ice cream; milk puddings, e.g. rice, custard, trifle; dairy desserts	99.4%	0.93 ± 0.79	1.05 ± 0.82	0.06 ± 0.81[1]
Low-fat	Low-fat milk; low-fat yogurt; low-fat cheese	94.6%	1.71 ± 1.10	1.43 ± 1.03	-0.21 ± 1.07[1]
Fermented dairy	High-fat fermented dairy; low-fat fermented dairy	99.2%	0.97 ± 0.71	1.00 ± 0.78	0.03 ± 0.79
High-fat	High-fat yogurt; high-fat cheese	98.6%	0.52 ± 0.43	0.58 ± 0.49	0.07 ± 0.50
Low-fat	Low-fat yogurt; low-fat cheese	76.7%	0.46 ± 0.58	0.42 ± 0.59	-0.04 ± 0.63
Cream	Double or clotted cream, single or sour cream	47.1%	0.09 ± 0.19	0.09 ± 0.21	0.00 ± 0.23
Ice cream	Ice cream, choc ices	64.2%	0.05 ± 0.08	0.06 ± 0.10	0.00 ± 0.10
Butter	Butter on bread or vegetables	60.4%	0.49 ± 0.86	0.57 ± 0.88	0.07 ± 0.90

Values are mean ± SD. Excluding participants who consumed non-specific milk at baseline or follow-up (*n* = 2,162, 29%) to avoid classification errors in change estimates. Consumers were defined as consuming >0 servings/day of a specific dairy type. Serving sizes were one average glass of milk; 200g, one pot of yogurt; 125g, a medium serving of cheese; 40g, one tablespoon for single cream; 15g, one tablespoon for double cream; 30g one average scoop/tub for ice cream; 150g, one teaspoon for butter, 10g [29].

Covariates

Sociodemographic and health factors

Data on covariates related to sociodemographic factors, health behaviours and medication use were collected at baseline and follow-up using a general health questionnaire. Information was obtained on sex, age (years), education, ethnicity, occupation, income, marital status, smoking status, family history of diabetes, and self-reported medication use. Categorisation of those covariates were operated as previously conducted [29]. Education was defined as low (no formal qualifications, primary school: School Leaving Certificate, Certificate of Secondary Education, or Ordinary Level), intermediate (high school: City & Guilds qualifications, apprenticeship, matriculation, trade, or Advanced Level) or high (higher vocational, college, or university education) and age on completing formal education (years). Ethnic origin was classified into two categories (White and non-White including Black Caribbean, other Black, Indian, Pakistani, Bangladesh, Chinese, others). Occupation was assessed as paid or voluntarily full-time or part-time working (yes or no including housekeeping, retired, unemployed, on sick leave or studying). Total household income over the 12 previous months was categorised as <£20,000, £20,000-40,000 and >£40,000 per year [35]. Marital status was defined as single, married or widowed/separated. Smoking status was defined as current, former, and never. Pack-years among smokers were calculated as a product of self-reported duration and smoking amounts. Physical activity was objectively measured over 7 days using a combined heart rate and movement sensor (Actiheart, CamNtech) [36] and individually calibrated with a treadmill test to derive physical activity energy expenditure (PAEE) (kJ/kg/d) [37]. Questions related to family history of diabetes (parent or sibling) were included (yes or no). In consenting participants weighting ≤140 kg, full body dual-energy X-ray absorptiometry (DXA) scans (GE Lunar Prodigy Advanced, GE Medical Systems) and enCORE software version 14-16 (GE Healthcare) were used to derive fat mass measurements across body regions [38]. This method was validated against the gold-standard 4-compartment method [39].

Anthropometric and clinical measurements

Data on anthropometric measures were obtained from clinical measurements at baseline and follow-up according to a standardized protocol [40, 41]. 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-stretchable tape. Body fat percentage was determined by dual-energy X-ray absorptiometry [38]. Systolic and diastolic blood pressure were measured thrice using an Accutorr sphygmomanometer after the participant had been rested for five minutes [41] and expressed in mmHg. The mean of the three measurements was used for analysis. Hypertension at baseline or follow-up was defined as systolic blood pressure ≥140 mm Hg or diastolic blood pressure ≥90 mm Hg [42] or use of anti-hypertensive medication or self-reported high blood pressure.

Blood lipid markers and biomarkers

Plasma total cholesterol, high-density lipoprotein cholesterol (HDL-C) and triacylglycerol (TAG) concentrations (mmol/L) in fasting plasma samples were measured using standard enzymatic methods [43]. Low-density lipoprotein cholesterol (LDL-C) concentrations were calculated using Friedewald's formula [44]. Dyslipidaemia at baseline or follow-up was defined as plasma total cholesterol concentration above 190 mg/dl (~4.9 mmol/L) [45] or use of lipid lowering medication. The plasma phospholipids fatty acids pentadecanoic acid (C15:0), heptadecanoic acid (C17:0), and *trans*-palmitoleic acid (*t*C16:1n7) have been used as biomarkers for dairy fat intake and found negatively associated with T2D in cohort studies [46]. Meta-analyses of observational studies show that myristic acid (C14:0) correlates with dairy intake, [47] and higher concentrations are linked to higher T2D risk [48]. The methods used to estimate the quantities of these biomarkers were described in detail previously [49].

Statistical analyses

Baseline characteristics were described across intake range categories of milk, yogurt, and cheese. For the analysis of associations of dairy intakes with prediabetes risk at follow-up, Poisson regression with robust variance was used to estimate the relative risk (RR) with 95% confidence interval (CI) [50]. For the analysis of change in glycaemic status, the outcome was treated as an ordinal variable representing three levels of glycaemic status as normoglycaemic, prediabetes, and T2D, and evaluated with Poisson regression. Follow-up periods varying by participant were modelled as an offset variable. Dairy intakes were modelled as categorical variables based on non-consumers and consumer categories (two or three groups depending on the intake distribution), a continuous scale (servings/day) at baseline, and changes in dairy intake between the baseline and the follow-up assessment (i.e., the absolute difference, subtracting the baseline intake from the follow-up intake). A non-linear association was assessed with a second-order polynomial or restricted cubic spline function compared to the linear model assessed by likelihood ratio tests.

Confounders were selected and modelled sequentially to present RR estimates with different sets of covariates adjusted for, based on existing evidence, biological plausibility, and statistical efficacy [51, 52]. We first adjusted for age, sex, study site (Cambridge, Ely, and Wisbech), and total energy intake. We considered the following covariates to adjust for: educational level, age at finishing education, ethnicity, marital status, occupation, income, smoking, alcohol intake (restricted cubic spline terms), PAEE, family history of T2D, hypertension, dyslipidaemia, waist circumference, food groups found to be associated with T2D in other cohorts (fruits, vegetables, whole grains, refined grains, potatoes, legumes, nuts, tea, coffee, red meat, processed meat, fatty fish, and sugar-sweetened beverages) [53], the tertiles Mediterranean Diet Score (tMDS) adherence, the Dietary Approaches to Stop Hypertension (DASH) adherence, and intake of all other dairy foods when evaluating each dairy subtype. For changes in dairy intake,

models included the baseline value of the outcome, baseline intake of the dairy type and changes in covariates if applicable. Potential effect modification by age, sex, waist circumference and educational level were explored in the most adjusted models and stratified associations were presented in case of a significant interaction ($p < 0.05$).

Multiple secondary analyses were performed to evaluate relevant hypotheses or to assess the robustness of the findings. We evaluated dairy consumption after energy adjustment with a residual technique [22]. We repeated main analyses using different prediabetes definitions (**Table 1**). Linear regression models were used to analyse changes in dairy intake with parallel changes in continuous variables of glycaemic markers (FPG, 2hPG, and HbA1c) between the baseline and follow-up measurement. In analyses for changes in high-fat and low-fat milk and dairy, we excluded participants who consumed non-specific milk with unknown fat content at baseline or follow-up (*n* = 2,162, 29%) to avoid classification errors in change estimates. Furthermore, we used linear mixed models with each dairy exposure, follow-up time and an interaction between the dairy exposure and follow-up time as a fixed effect. The main term represents the overall association between dairy consumption and the glycaemic outcome over the follow-up period. The interaction term reflects how the association of dairy consumption with glycaemic markers changes over time. Additionally, the models included a random intercept for participants and a random slope for the follow-up time. The results were expressed as beta coefficients and 95% confidence intervals (β, 95% CI). To assess the impact of dairy fat biomarkers on the associations between dairy intake and prediabetes risk that were significant in main analysis, we calculated the percentage change in regression coefficients before and after introducing dairy fat biomarkers into the models [54, 55]. This percentage change was expressed as $100 \times [(\beta_{\text{ref model}} - \beta_{\text{ref model + dairy fat biomarkers}})/\beta_{\text{ref model}}]$. We considered the percentage of attenuation of the association between each dairy type and prediabetes as the degree of mediation due to the dairy fat biomarkers. Confidence intervals for the percentage attenuation were estimated by bootstrapping with 2000 iterations. Lastly, we applied network models using the statistical analysis approach described in Chapter 6.

Missing values for covariates were imputed using multiple imputation (*n* = 10) (**Supplemental Table 1**) [56]. All analyses were conducted using R, version 4.1.2 (R Foundation for Statistical Computing, Vienna, Austria). P-values (two-sided) <0.05 were considered statistically significant.

Results

Descriptive characteristics

In this study sample of 7,521 participants, most participants consumed dairy (99.7%) contributing to a median 14.0 weight% of the total diet (**Table 2; Supplemental Figure 1**). Dairy intake mainly consisted of low-fat milk (mean ± SD 1.25±0.90 servings/day), high-fat

cheese (0.44±0.36 servings/day) and low-fat yogurt (0.34±0.46 servings/day). Low-fat milk intake decreased slightly over follow-up on average (-0.18±0.88 servings/day). The mean age was 48.7±2.0 years, 51.9% was female, and 9.4% were current smokers. The mean waist circumference was 90.0±12.9 cm, the mean BMI was 26.4±4.4 kg/m^2 and 17.6% were obese BMI ≥30 kg/m^2). The demographic and health characteristics across high and low intake ranges of milk, yogurt and cheese are presented in **Table 3**, showing noteworthy patterns. For example, compared to non-consumers, dairy consumers were more likely to report a White ethnic background. High milk consumers were more likely to be men and with lower educational level, while high yogurt and cheese consumers were more likely to be women and with higher education level. Compared to low yogurt consumers, high yogurt consumers were less often smokers, with a higher body fat percentage, but lower prevalences of hypertension and dyslipidaemia, and higher diet quality. Among high cheese consumers the prevalence of hypertension and dyslipidaemia was also lower compared to low cheese consumers. Furthermore, diet quality was slightly higher, however also intake of meat and SSBs was higher. Descriptive characteristics by dairy consumption, by sex, and by included versus excluded from the current analyses are presented in **Supplemental Table 2.** Weak correlations between dairy types, blood lipids, inflammatory markers and dairy fat biomarkers were found (**Supplemental Table 3**). The highest correlations were found between high-fat fermented dairy, high-fat cheese, butter, and dairy fat biomarker C15:0 (range 0.22-0.24).

Table 3. Baseline characteristics of participants in the Fenland study in the total population and across bottom and top one third (T3) of each intake levels of milk, yogurt, and cheese (*n* = 7,521).

	Total	Milk intake		Yogurt intake		Cheese intake	
		Zero	T3	Zero	T3	Zero	T3
N participants	7,521	2,869	932	2,584	1,420	1,880	1,885
Total dairy intake (servings/day)		0.6 ± 0.3	3.1 ± 0.3	0.0 ± 0.0	1.2 ± 0.6	0.1 ± 0.1	1.1 ± 0.5
Range		0.0-0.7	2.9-3.7	0.0-0.1	0.87-6.12	0.0-0.2	0.7-7.0
Follow-up time (years)	6.7 ± 2.0	6.6 ± 2.0	6.8 ± 2.1	6.8 ± 2.0	6.6 ± 2.0	6.8 ± 2.1	6.7 ± 2.0
Age at baseline (years)	48.7 ± 7.4	48.7 ± 7.4	48.3 ± 7.4	48.5 ± 7.5	49.5 ± 7.3	49.3 ± 7.4	48.1 ± 7.5
Sex, female (%)	51.9	56.4	34.4	41.4	64.6	49.1	56.8
Educational level (%)							
Intermediate	45.4	44.4	48.3	47.4	45.3	50.9	40.2
High	38.3	40.0	32.2	33.1	40.7	30.0	47.4
Age at finishing education (years)	19.3 ± 4.4	19.5 ± 4.7	18.9 ± 4.1	18.9 ± 4.4	19.5 ± 4.3	18.8 ± 4.5	20.1 ± 4.7
Ethnic origin, non-White (%)	2.3	3.5	1.4	2.9	1.8	3.9	1.4
Working (%)	89.0	89.8	88.8	88.7	89.3	89.7	88.2
Household income (%)							
£20,000-40,000	33.8	32.8	36.8	36.1	33.6	32.8	33.9
>£40,000	55.3	55.9	51.3	51.4	55.8	53.7	56.2
Marital status (%)							
Married	82.7	81.3	82.4	81.6	82.3	81.1	81.8
Widowed/separated	9.1	9.0	9.6	9.1	9.6	10.1	9.0
Smoking status (%)							
Former	33.3	33.8	31.3	32.2	31.7	33.7	32.3
Current	9.4	10.0	12.7	13.6	5.9	9.0	9.3

Table 3. Baseline characteristics of participants in the Fenland study in the total population and across bottom and top one third (T3) of each intake levels of milk, yogurt, and cheese (*n* = 7,521) (continued).

	Total	Milk intake		Yogurt intake		Cheese intake	
		Zero	T3	Zero	T3	Zero	T3
Pack-years among smokers	4,024 [1,653-8,273]	3,681 [1,612-7,853]	4,675 [1,812-9,526]	4,743 [2,238-9,423]	3,232 [1,375-7,163]	4,401 [2,104-9,017]	3,707 [1,456-7,453]
Alcohol intake (g/day)	6.4 [1.7; 12.5]	6.4 [1.7; 12.5]	5.8 [0.9; 11.1]	6.2 [1.6; 13.3]	5.8 [1.6; 10.9]	5.4 [0.9; 11.1]	6.7 [1.7; 13.6]
Physical activity (kJ/kg/d)	54.6 ± 21.7	53.0 ± 20.7	59.4 ± 24.4	54.8 ± 22.3	54.7 ± 21.8	52.9 ± 22.1	56.0 ± 21.5
Family history of diabetes (%)	23.4	23.9	20.6	24.7	24.4	24.2	23.2
BMI (kg/m2)	26.4 ± 4.4	26.3 ± 4.5	26.8 ± 4.2	26.7 ± 4.6	26.1 ± 4.4	26.7 ± 4.6	26.0 ± 4.3
Body fat (%)	32.6 ± 8.4	33.0 ± 8.4	30.9 ± 8.3	31.7 ± 8.4	33.7 ± 8.5	32.7 ± 8.4	32.6 ± 8.4
Waist circumference (cm)	90.0 ± 12.9	89.3 ± 13.1	92.6 ± 12.5	91.6 ± 13.3	88.0 ± 12.5	90.7 ± 12.9	88.8 ± 12.8
Hypertension (%)	36.7	36.5	36.8	40.2	34.6	39.6	33.5
Dyslipidaemia (%)	9.0	8.7	9.4	10.2	7.6	10.5	7.3
Dietary intake at baseline							
Energy intake (kcal/day)	1,948 ± 577	1,792 ± 556	2,238 ± 620	1,864 ± 578	2,073 ± 557	1,698 ± 511	2,197 ± 605
DASH score	24.3 ± 4.5	23.4 ± 4.5	24.9 ± 4.4	22.8 ± 4.5	25.9 ± 4.3	24.2 ± 4.6	24.8 ± 4.6
tMDS score	7.6 ± 1.5	7.6 ± 1.5	7.3 ± 1.5	7.2 ± 1.5	8.0 ± 1.4	7.4 ± 1.5	8.0 ± 1.4
Fruit (g/day)	243 ± 178	236 ± 179	242 ± 170	191 ± 153	324 ± 213	223 ± 178	274 ± 190
Vegetables (g/day)	273 ± 137	273 ± 138	269 ± 132	250 ± 133	309 ± 164	257 ± 139	298 ± 131
Wholegrains (g/day)	84.9 ± 81.6	78.8 ± 81.2	90.5 ± 86.2	69.6 ± 73.4	100.5 ± 86.7	78.5 ± 82.8	98.8 ± 86.9
Refined grains (g/day)	124 ± 73	115 ± 72	135 ± 75	121 ± 77	124 ± 70	106 ± 73	140 ± 76
Meat (red and processed) (g/day)	63.9 ± 42.1	59.6 ± 41.0	72.4 ± 48.4	67.7 ± 45.3	59.2 ± 40.6	62.1 ± 40.5	61.8 ± 47.1
Coffee (g/day)	289 ± 308	265 ± 298	344 ± 347	287 ± 321	294 ± 303	270 ± 312	309 ± 310
Tea (g/day)	493 ± 366	400 ± 350	598 ± 395	499 ± 382	497 ± 364	493 ± 370	479 ± 366

Table 3. Baseline characteristics of participants in the Fenland study in the total population and across bottom and top one third (T3) of each intake levels of milk, yogurt, and cheese (*n* = 7,521) (continued).

	Total	Milk intake		Yogurt intake		Cheese intake	
		Zero	T3	Zero	T3	Zero	T3
SSB (g/day)	102.9 ± 115.0	106.1 ± 129.9	106.8 ± 109.6	103.7 ± 128.6	104.2 ± 109.0	95.1 ± 118.3	112.5 ± 122.2
Total fat (en%)	31.9 ± 5.4	32.1 ± 5.7	32.0 ± 5.3	32.5 ± 5.7	31.1 ± 5.3	30.2 ± 5.6	33.3 ± 5.3
SFAs (en%)	11.8 ± 2.9	11.4 ± 2.9	12.5 ± 3.1	12.0 ± 3.1	11.6 ± 2.8	10.6 ± 2.8	12.8 ± 2.9
Protein (en%)	18.2 ± 3.5	17.9 ± 3.7	19.0 ± 3.3	17.9 ± 3.5	18.8 ± 3.5	18.9 ± 3.8	17.8 ± 3.3
Carbohydrate (en%)	50.2 ± 6.9	49.8 ± 7.6	50.2 ± 6.2	49.3 ± 7.4	51.5 ± 6.6	51.2 ± 7.5	49.5 ± 6.8
Intrinsic and milk sugars (g/day)	59.2 ± 34.8	55.8 ± 36.0	67.6 ± 38.1	58.8 ± 37.1	61.3 ± 33.0	52.4 ± 33.5	65.2 ± 36.8
Calcium (mg/day)	1,034 ± 336	800 ± 253	1,473 ± 271	919 ± 321	1,205 ± 335	882 ± 308	1,209 ± 344
Sodium (mg/day)	2,007 ± 712	1,842 ± 697	2,320 ± 765	1,896 ± 698	2,162 ± 717	1,710 ± 626	2,347 ± 752

Values are mean ± SD for continuous variables with a normal distribution, or median [IQR] for continuous variables with a skewed distribution, percentages for categorical variables, based on unimputed data.
Abbreviations: BMI, Body Mass Index; DASH, Dietary Approaches to Stop Hypertension; en%, percentage of total energy intake; SFA, Saturated Fatty acids; SSB, Sugar Sweetened Beverages; SD, Standard Deviation; tMDS, tertiles Mediterranean diet score.

Baseline dairy intake and prediabetes risk

At a mean follow-up of 6.7 ± 2.0 years, 290 incident cases of prediabetes were identified among 6,639 participants (4.4%). Total dairy, low-fat dairy, fermented dairy, low-fat fermented dairy, milk, yogurt and low-fat yogurt, cream, ice cream and butter were not associated with risk of prediabetes. After adjusting for potential confounders at baseline, a higher intake of high-fat dairy ($RR_{serving/day}$ 1.20, 95%CI 1.03-1.39) and high-fat milk ($RR_{serving/day}$ 1.22, 95%CI 1.01-1.47) were associated with higher prediabetes risk (**Figure 2**). Associations of high-fat fermented dairy, high-fat yogurt, and total, high-fat, and low-fat cheese were similar ($RR_{serving/day}$ ranging from 1.22-1.28). On the contrary, low-fat milk was associated with lower prediabetes risk ($RR_{serving/day}$ 0.86, 0.75-0.98). Non-linear associations were not evident from categorical analyses (**Table 4**) and in models with fractional polynomial or spline terms. In secondary analyses, additionally adjusting for BMI, marital status, income, diet quality instead of food groups, or intake of other dairy types did not change any of the parameter estimates (**Supplemental Table 4**).

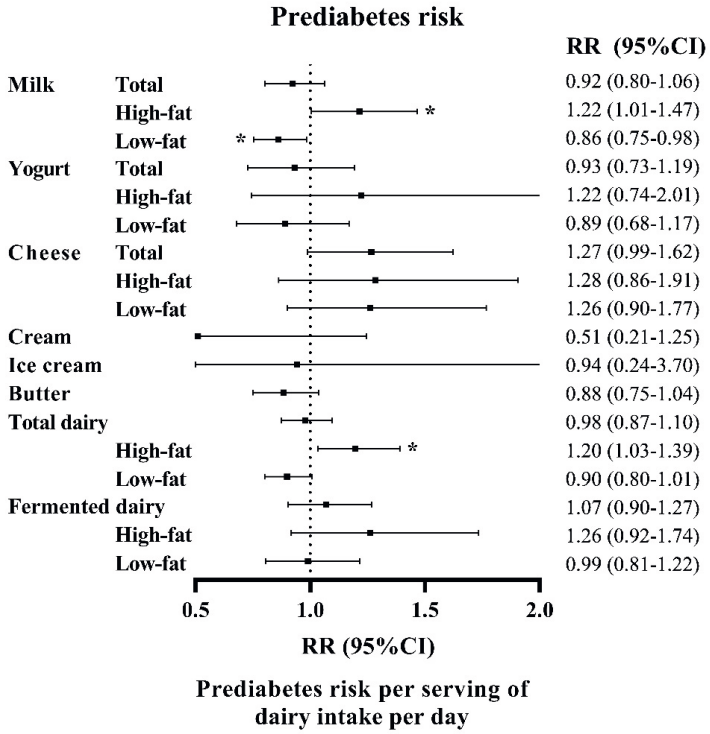


Figure 2. Associations of dairy intake by serving/day and prediabetes risk (*n* = 6,639). Bars represent continuous relative risks (RR) with 95% confidence intervals (CI) adjusted for age, sex, study site, energy intake, educational level, age at completion of education, ethnic origin, alcohol use, smoking behaviour, physical activity, family history, intake of fruit, vegetables, whole grains, refined grains, potatoes, legumes, nuts, red and processed meat, fatty fish, coffee, tea, sugar-sweetened beverages, hypertension, dyslipidaemia, and waist circumference. * P-value <0.05.

Significant interactions were observed for baseline intake of low-fat dairy and age, and total dairy, total and low-fat fermented dairy, and total yogurt with educational level (**Supplemental Table 5**). In participants aged <50 years, low-fat dairy was associated with a lower risk of prediabetes ($RR_{serving/day}$ 0.80, 0.67-0.97), but not in participants aged ≥50 years ($RR_{serving/day}$ 0.97, 0.84-1.12). Total dairy, fermented dairy, low-fat fermented dairy, and yogurt were associated with lower prediabetes risk in participants with a high educational level (RRs ranging from 0.59-0.78), however, the associations were not evident among those with intermediate ($RR_{serving/day}$ ranging from 1.00-1.02) and low educational levels (1.11-1.42).

Table 4. Associations of dairy intake and prediabetes risk in the Fenland study (*n* = 6,639).

	Relative risk (95% CI) across intake range categories[1]					Continuous[2]
	Q1	Q2	Q3	Q4	P_{trend}	RR (95%CI)
Total dairy						
n cases/n total	92/1,666	58/1,673	71/1,631	69/1,669		
Median, servings/d	1.32	2.17	2.91	3.98		
Model 1	1 (ref)	0.64 (0.46-0.88)	0.82 (0.60-1.10)	0.79 (0.57-1.09)	0.31	0.97 (0.87-1.07)
Model 2	1 (ref)	0.65 (0.47-0.90)	0.87 (0.64-1.18)	0.83 (0.60-1.15)	0.49	0.99 (0.89-1.09)
Model 3	1 (ref)	0.65 (0.47-0.90)	0.87 (0.64-1.19)	0.81 (0.57-1.16)	0.48	0.98 (0.88-1.10)
Model 4	1 (ref)	0.66 (0.48-0.92)	0.87 (0.64-1.20)	0.82 (0.57-1.17)	0.48	0.98 (0.87-1.10)
High-fat dairy						
n cases/n total	81/1,687	63/1,645	73/1,658	73/1,649		
Median, servings/d	0.24	0.57	0.92	1.67		
Model 1	1 (ref)	0.86 (0.62-1.20)	1.06 (0.76-1.48)	1.16 (0.82-1.65)	0.22	1.14 (0.99-1.32)
Model 2	1 (ref)	0.91 (0.66-1.27)	1.15 (0.83-1.60)	1.32 (0.94-1.87)	0.05	1.19 (1.04-1.37)*
Model 3	1 (ref)	0.89 (0.64-1.23)	1.10 (0.79-1.54)	1.26 (0.89-1.79)	0.10	1.17 (1.01-1.36)*
Model 4	1 (ref)	0.89 (0.64-1.23)	1.09 (0.78-1.52)	1.32 (0.93-1.88)	0.06	1.20 (1.03-1.39)*
Low-fat dairy						
n cases/n total	85/1,666	66/1,620	68/1,680	71/1,673		
Median, servings/d	0.54	1.27	1.90	3.00		
Model 1	1 (ref)	0.77 (0.56-1.06)	0.78 (0.58-1.07)	0.81 (0.59-1.11)	0.25	0.91 (0.81-1.01)
Model 2	1 (ref)	0.78 (0.57-1.06)	0.81 (0.59-1.10)	0.81 (0.59-1.12)	0.28	0.91 (0.81-1.01)
Model 3	1 (ref)	0.79 (0.58-1.08)	0.83 (0.61-1.14)	0.84 (0.60-1.17)	0.39	0.91 (0.81-1.02)
Model 4	1 (ref)	0.78 (0.57-1.07)	0.81 (0.59-1.11)	0.81 (0.58-1.13)	0.29	0.90 (0.80-1.01)
Total fermented dairy						
n cases/n total	76/1,612	75/1,674	63/1,661	76/1,692		
Median, servings/d	0.27	0.65	1.04	1.64		
Model 1	1 (ref)	1.04 (0.76-1.43)	0.91 (0.65-1.26)	1.11 (0.80-1.54)	0.66	1.02 (0.86-1.20)
Model 2	1 (ref)	1.11 (0.81-1.51)	0.97 (0.70-1.36)	1.22 (0.87-1.71)	0.35	1.05 (0.90-1.23)
Model 3	1 (ref)	1.14 (0.83-1.56)	0.99 (0.71-1.38)	1.30 (0.91-1.85)	0.23	1.07 (0.91-1.27)
Model 4	1 (ref)	1.12 (0.81-1.53)	0.96 (0.69-1.35)	1.27 (0.89-1.81)	0.28	1.07 (0.90-1.27)

Table 4. Associations of dairy intake and prediabetes risk in the Fenland study (*n* = 6,639) (continued).

	Relative risk (95% CI) across intake range categories[1]					Continuous[2]
	Q1	Q2	Q3	Q4	P_{trend}	RR (95%CI)
High-fat fermented dairy						
n cases/n total	86/1,669	69/1,628	66/1,664	69/1,678		
Median, servings/d	0.13	0.37	0.52	0.93		
Model 1	1 (ref)	0.87 (0.64-1.19)	0.91 (0.65-1.26)	1.02 (0.73-1.44)	0.81	1.14 (0.82-1.60)
Model 2	1 (ref)	0.90 (0.66-1.22)	0.99 (0.71-1.39)	1.16 (0.82-1.65)	0.34	1.28 (0.95-1.74)
Model 3	1 (ref)	0.87 (0.64-1.19)	0.98 (0.70-1.36)	1.13 (0.79-1.60)	0.43	1.26 (0.92-1.73)
Model 4	1 (ref)	0.89 (0.65-1.21)	1.00 (0.71-1.39)	1.13 (0.79-1.60)	0.43	1.26 (0.92-1.74)
Low-fat fermented dairy						
n cases/n total	63/1,535	96/1,877	58/1,555	73/1,672		
Median, servings/d	0.00	0.14	0.43	1.01		
Model 1	1 (ref)	1.30 (0.96-1.78)	0.93 (0.66-1.32)	1.06 (0.76-1.48)	0.59	0.97 (0.79-1.19)
Model 2	1 (ref)	1.38 (1.01-1.89)	1.01 (0.71-1.44)	1.10 (0.79-1.55)	0.68	0.96 (0.79-1.17)
Model 3	1 (ref)	1.43 (1.04-1.95)	1.05 (0.74-1.50)	1.18 (0.83-1.68)	0.97	1.00 (0.82-1.22)
Model 4	1 (ref)	1.41 (1.03-1.93)	1.03 (0.72-1.47)	1.15 (0.81-1.63)	0.85	0.99 (0.81-1.22)
Milk						
n cases/n total	126/2,543	89/2,303	41/979	34/814		
Median, servings/d	0.73	1.47	2.20	2.93		
Model 1	1 (ref)	0.76 (0.58-0.99)	0.82 (0.59-1.16)	0.84 (0.58-1.22)	0.23	0.94 (0.82-1.07)
Model 2	1 (ref)	0.76 (0.58-1.00)	0.83 (0.59-1.17)	0.85 (0.59-1.25)	0.26	0.94 (0.83-1.08)
Model 3	1 (ref)	0.75 (0.58-0.98)	0.81 (0.57-1.16)	0.81 (0.55-1.20)	0.20	0.93 (0.80-1.07)
Model 4	1 (ref)	0.76 (0.58-0.99)	0.81 (0.57-1.16)	0.80 (0.54-1.18)	0.18	0.92 (0.80-1.06)
High-fat milk						
n cases/n total	241/5,608	20/546	29/485			
Median, servings/d	0.00	0.73	1.47			
Model 1	1 (ref)	0.87 (0.55-1.35)	1.41 (0.97-2.06)		0.20	1.19 (0.99-1.44)
Model 2	1 (ref)	0.87 (0.55-1.36)	1.50 (1.03-2.19)		0.13	1.22 (1.01-1.47)*
Model 3	1 (ref)	0.87 (0.55-1.36)	1.40 (0.96-2.05)		0.22	1.18 (0.97-1.43)
Model 4	1 (ref)	0.88 (0.56-1.39)	1.50 (1.02-2.20)		0.12	1.22 (1.01-1.47)*

Table 4. Associations of dairy intake and prediabetes risk in the Fenland study (*n* = 6,639) (continued).

	Relative risk (95% CI) across intake range categories[1]					Continuous[2]
	Q1	Q2	Q3	Q4	P_{trend}	RR (95%CI)
Low-fat milk						
n cases/n total	54/1,193	101/1,965	77/1,977	58/1,504		
Median, servings/d	0.00	0.73	1.47	2.20		
Model 1	1 (ref)	1.12 (0.81-1.54)	0.83 (0.59-1.16)	0.83 (0.58-1.20)	0.08	0.88 (0.77-1.00)
Model 2	1 (ref)	1.11 (0.81-1.53)	0.83 (0.59-1.17)	0.83 (0.57-1.19)	0.08	0.88 (0.77-1.00)
Model 3	1 (ref)	1.12 (0.82-1.54)	0.84 (0.60-1.19)	0.82 (0.56-1.20)	0.09	0.87 (0.76-1.00)
Model 4	1 (ref)	1.10 (0.80-1.52)	0.82 (0.58-1.16)	0.79 (0.54-1.15)	0.06	0.86 (0.75-0.98)*
Total yogurt						
n cases/n total	121/2,251	37/1,093	77/2,020	55/1,275		
Median, servings/d	0.00	0.14	0.43	1.01		
Model 1	1 (ref)	0.68 (0.48-0.98)	0.72 (0.55-0.95)	0.83 (0.60-1.14)	0.33	0.90 (0.69-1.16)
Model 2	1 (ref)	0.76 (0.53-1.10)	0.78 (0.58-1.03)	0.87 (0.63-1.19)	0.43	0.91 (0.71-1.16)
Model 3	1 (ref)	0.79 (0.55-1.13)	0.79 (0.59-1.05)	0.92 (0.65-1.30)	0.66	0.93 (0.73-1.19)
Model 4	1 (ref)	0.80 (0.56-1.15)	0.79 (0.59-1.05)	0.92 (0.65-1.30)	0.66	0.93 (0.73-1.19)
High-fat yogurt						
n cases/n total	203/4,307	42/1,131	17/533	28/668		
Median, servings/d	0.00	0.07	0.14	0.43		
Model 1	1 (ref)	0.87 (0.62-1.20)	0.81 (0.49-1.32)	1.07 (0.72-1.58)	0.88	1.02 (0.56-1.89)
Model 2	1 (ref)	0.97 (0.70-1.35)	0.88 (0.54-1.43)	1.19 (0.80-1.78)	0.46	1.20 (0.69-2.08)
Model 3	1 (ref)	1.01 (0.72-1.40)	0.90 (0.55-1.47)	1.23 (0.82-1.85)	0.37	1.19 (0.71-2.00)
Model 4	1 (ref)	1.05 (0.75-1.46)	0.91 (0.55-1.48)	1.28 (0.85-1.92)	0.29	1.22 (0.74-2.01)
Low-fat yogurt						
n cases/n total	90/1,973	86/1,777	54/1,476	60/1,413		
Median, servings/d	0.00	0.07	0.43	1.01		
Model 1	1 (ref)	1.10 (0.82-1.47)	0.80 (0.58-1.12)	0.90 (0.65-1.25)	0.24	0.88 (0.67-1.17)
Model 2	1 (ref)	1.18 (0.88-1.57)	0.86 (0.62-1.21)	0.92 (0.66-1.28)	0.25	0.88 (0.67-1.14)
Model 3	1 (ref)	1.20 (0.90-1.61)	0.88 (0.63-1.24)	0.96 (0.68-1.36)	0.38	0.90 (0.69-1.17)
Model 4	1 (ref)	1.19 (0.89-1.59)	0.87 (0.62-1.22)	0.95 (0.67-1.34)	0.34	0.89 (0.68-1.17)

Table 4. Associations of dairy intake and prediabetes risk in the Fenland study (*n* = 6,639) (continued).

	Relative risk (95% CI) across intake range categories[1]					Continuous[2]
	Q1	Q2	Q3	Q4	P_{trend}	RR (95%CI)
Cheese						
n cases/n total	83/1,644	68/1,633	61/1,681	78/1,681		
Median, servings/d	0.138355	0.380865	0.55464	0.9785425		
Model 1	1 (ref)	0.87 (0.63-1.20)	0.82 (0.59-1.15)	1.12 (0.81-1.55)	0.41	1.18 (0.92-1.53)
Model 2	1 (ref)	0.91 (0.66-1.25)	0.89 (0.64-1.25)	1.24 (0.89-1.73)	0.17	1.26 (0.99-1.59)
Model 3	1 (ref)	0.89 (0.65-1.23)	0.88 (0.63-1.23)	1.25 (0.89-1.74)	0.16	1.28 (1.00-1.63)
Model 4	1 (ref)	0.91 (0.66-1.25)	0.88 (0.63-1.23)	1.23 (0.88-1.71)	0.20	1.27 (0.99-1.62)
High-fat cheese						
n cases/n total	87/1,647	63/1,657	72/1,650	68/1,685		
Median, servings/d	0.12	0.28	0.47	0.81		
Model 1	1 (ref)	0.78 (0.56-1.07)	0.96 (0.70-1.33)	0.96 (0.68-1.36)	0.87	1.19 (0.79-1.80)
Model 2	1 (ref)	0.81 (0.59-1.12)	1.05 (0.76-1.45)	1.07 (0.75-1.51)	0.47	1.33 (0.92-1.91)
Model 3	1 (ref)	0.80 (0.58-1.10)	1.02 (0.74-1.40)	1.02 (0.72-1.45)	0.62	1.29 (0.87-1.92)
Model 4	1 (ref)	0.80 (0.58-1.10)	1.03 (0.75-1.42)	1.02 (0.71-1.45)	0.63	1.28 (0.86-1.91)
Low-fat cheese						
n cases/n total	180/4,109	62/1,365	18/562	30/603		
Median, servings/d	0.00	0.11	0.21	0.65		
Model 1	1 (ref)	1.08 (0.81-1.44)	0.75 (0.47-1.22)	1.17 (0.80-1.71)	0.58	1.17 (0.84-1.63)
Model 2	1 (ref)	1.12 (0.84-1.49)	0.79 (0.49-1.28)	1.18 (0.80-1.72)	0.53	1.19 (0.85-1.66)
Model 3	1 (ref)	1.15 (0.87-1.53)	0.84 (0.52-1.35)	1.30 (0.88-1.93)	0.25	1.28 (0.92-1.78)
Model 4	1 (ref)	1.13 (0.85-1.50)	0.80 (0.49-1.28)	1.26 (0.85-1.87)	0.34	1.26 (0.90-1.77)
Cream						
n cases/n total	166/3,492	66/1,413	33/856	25/878		
Median, servings/d	0.00	0.09	0.14	0.28		
Model 1	1 (ref)	1.02 (0.77-1.35)	0.88 (0.61-1.28)	0.68 (0.44-1.05)	0.09	0.34 (0.12-0.94)*
Model 2	1 (ref)	1.07 (0.81-1.42)	0.96 (0.66-1.40)	0.79 (0.51-1.21)	0.36	0.48 (0.20-1.19)
Model 3	1 (ref)	1.07 (0.81-1.41)	0.95 (0.66-1.38)	0.80 (0.52-1.23)	0.38	0.49 (0.20-1.21)
Model 4	1 (ref)	1.07 (0.81-1.42)	0.97 (0.67-1.41)	0.81 (0.53-1.24)	0.42	0.51 (0.21-1.25)

Table 4. Associations of dairy intake and prediabetes risk in the Fenland study (*n* = 6,639) (continued).

	Relative risk (95% CI) across intake range categories[1]					Continuous[2]
	Q1	Q2	Q3	Q4	P_{trend}	RR (95%CI)
Ice cream						
n cases/n total	92/2,378	129/2,640	44/1,032	25/589		
Median, servings/d	0.00	0.04	0.08	0.23		
Model 1	1 (ref)	1.28 (0.98-1.67)	1.19 (0.83-1.70)	1.10 (0.70-1.71)	0.77	1.31 (0.34-5.01)
Model 2	1 (ref)	1.29 (0.99-1.68)	1.18 (0.82-1.70)	1.11 (0.71-1.73)	0.76	1.29 (0.34-4.88)
Model 3	1 (ref)	1.27 (0.97-1.66)	1.14 (0.79-1.65)	1.06 (0.67-1.66)	0.95	1.06 (0.28-3.98)
Model 4	1 (ref)	1.22 (0.93-1.59)	1.08 (0.75-1.57)	1.01 (0.64-1.58)	0.87	0.94 (0.24-3.70)
Butter						
n cases/n total	131/2,604	54/1,332	40/950	65/1,753		
Median, servings/d	0.00	0.07	0.43	1.00		
Model 1	1 (ref)	0.88 (0.65-1.21)	0.94 (0.66-1.33)	0.85 (0.62-1.15)	0.36	0.88 (0.74-1.04)
Model 2	1 (ref)	0.95 (0.69-1.30)	1.03 (0.72-1.46)	0.94 (0.69-1.27)	0.75	0.92 (0.78-1.08)
Model 3	1 (ref)	0.95 (0.70-1.30)	1.02 (0.72-1.45)	0.89 (0.66-1.21)	0.53	0.89 (0.76-1.05)
Model 4	1 (ref)	0.94 (0.69-1.29)	1.03 (0.72-1.46)	0.88 (0.65-1.20)	0.48	0.88 (0.75-1.04)

[1] Relative risks (95CIs) were estimated across four categories split by quartile values (Q1 to Q4) or non-consumers + tertile or median categories with the lowest category as the reference, adjusted for covariates as follows: Model 1 included age, sex, study site and energy intake. Model 2 additionally adjusted for educational level, age at completion of education, ethnic origin, alcohol use, smoking behaviour, physical activity, and family history of diabetes. Model 3 additionally adjusted for dietary intakes (fruits, vegetables, whole grains, refined grains, potatoes, legumes, nuts, red and processed meat, fatty fish, coffee, tea, and sugar-sweetened beverages). Model 4 additionally adjusted for hypertension, dyslipidaemia, and waist circumference. Linear trend across intake range categories was assessed by including median values of each category as a continuous variable in the model.
[2] Relative risks per 1 serving/day (see definition in Table 2) were estimated. *P-value <0.05.
Abbreviations: CI, Confidence Interval; Q, Quartile.

Changes in dairy intake and glycaemic status

Of 7,410 participants, among those with normoglycaemia, 287 developed prediabetes (3.9%) and 61 T2D (0.8%), and among those with prediabetes, 99 developed T2D (1.3%). Of 803 participants with prediabetes at baseline, 470 (58.5%, 6.3% of total) regressed to normoglycaemia. Changes in intakes of high-fat milk were inversely associated with progressing from normoglycaemia to prediabetes or T2D ($RR_{serving/day}$ 0.86, 0.75-0.99) (**Figure 3**). None of the other dairy types were associated with glycaemic status, but changes in intakes of butter were positively associated with progression in glycaemic status ($RR_{serving/day}$ 1.11, 1.01-1.22). Additionally adjusting for BMI, marital status, income (results not shown), diet quality instead of food groups, or intake of other dairy types did not change parameter estimates (**Supplemental table 6**). In the non-linear association analyses (**Supplemental figure 2**), changes in high-fat fermented dairy consumption showed a U-shaped association (P=0.04 for the curvature), where a higher risk was observed when the intake either decreased or increased over the follow-up. For high-fat milk (P=0.03 for the curvature), risk was lowest at a modest increase of ~1 serving/day with a monotonic negative association of the dietary change with the prediabetes risk.

Significant interactions were observed for changes in intake of low-fat fermented dairy and low-fat yogurt with age, and total dairy, high-fat dairy, fermented dairy, high-fat fermented dairy, low-fat fermented dairy, yogurt, high-fat yogurt, and ice cream with waist circumference (**Supplemental Table 7**). In participants with a high waist circumference, changes in fermented dairy ($RR_{serving/day}$ 1.11, 1.01-1.22), low-fat fermented dairy ($RR_{serving/day}$ 1.16, 1.01-1.31) and yogurt ($RR_{serving/day}$ 1.14, 1.02-1.29) were positively associated with a higher risk of progressing from normoglycaemia to prediabetes or T2D. Other stratified analyses were not significant.

Secondary analysis

Associations with the energy-adjusted dairy exposure were similar (**Supplemental Tables 4 and 6**). The prediabetes incidence was 1.5% based on FPG only, 3.6% based on FPG and 2hPG, and 2.8% based on FPG and HbA1c, compared to 4.4% based on the WHO-IEC definition (**Table 1**). The impact of using different prediabetes definition on the effect estimates differed by dairy type (**Supplemental table 8**). For most dairy types, the effect estimates using the FPG & 2hPG definition most closely resembled those in main analysis using the WHO-IEC definition. Furthermore, associations of dairy types with prediabetes based on FPG only were slightly stronger compared to those using the WHO-IEC definition. The association between higher intake of high-fat dairy and higher prediabetes risk in main analysis was only significant, and slightly stronger, when using the FPG & HbA1c definition ($RR_{serving/day}$ 1.27, 95%CI 1.07-1.51), but not with the other outcome definitions. Higher intake of low-fat dairy was associated with lower prediabetes risk based on the three alternative definitions ($RR_{serving/day}$ ranging from 0.75-0.88), with strongest associations when using the FPG only definition ($RR_{serving/day}$ 0.75, 95%CI 0.61-0.94, $RR_{Q4vs.Q1}$ 0.53, 95%CI 0.30-0.93, P_{trend} = 0.04). For milk, all four outcome definitions

resulted in similar estimates, ranging from 1.17-1.25 for high-fat milk and 0.82-0.86 for low-fat milk. Using the FPG only definition resulted in a significant association for yogurt ($RR_{serving/day}$ 0.45, 0.25-0.80, $RR_{Q4vs.Q1}$ 0.44, 95%CI 0.21-0.95, P_{trend} = 0.02) and low-fat yogurt ($RR_{serving/day}$ 0.42, 95%CI 0.23-0.75; $RR_{Q4vs.Q1}$ 0.51, 95%CI 0.25-1.02, P_{trend} = 0.01). Higher low-fat yogurt was also significantly associated with prediabetes based on the FPG & HbA1c definition ($RR_{serving/day}$ 0.62, 95%CI 0.41-0.95; $RR_{Q4vs.Q1}$ 0.66, 0.42-1.03, P_{trend} = 0.01). The associations for cheese ($RR_{serving/day}$ 1.40, 95%CI 1.13-1.75; $RR_{Q4vs.Q1}$ 1.41, 95%CI 0.99-2.00, P_{trend} = 0.04), high-fat cheese ($RR_{serving/day}$ 1.46, 95%CI 1.04-2.05) and cream ($RR_{serving/day}$ 0.31, 95%CI 0.09-1.00) were stronger using the FPG & 2hPG definition compared to main analysis.

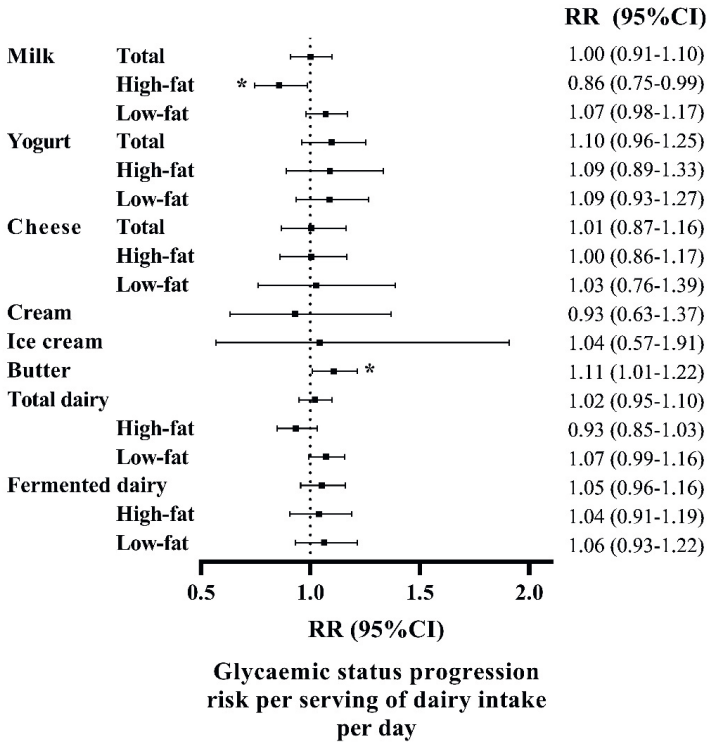


Figure 3. Associations of changes in dairy intake from baseline to follow-up in servings/day with glycaemic status (*n* = 7,410). Glycaemic status was operationalized as an ordinal variable with three levels: normoglycaemia, prediabetes and T2D, and thus relative risks (RR) represents the average risk of prediabetes from normoglycaemia and of T2D from prediabetes. RR with 95% confidence intervals (CI) adjusted for the baseline value of the outcome, age, sex, study site, energy intake, educational level, age at completion of education, ethnic origin, baseline and changes in alcohol use, smoking behaviour, baseline and changes in physical activity, family history, baseline intake of the dairy type, baseline and changes in dietary intakes (fruit, vegetables, whole grains, refined grains, potatoes, legumes, nuts, red and processed meat, fatty fish, coffee, tea, and sugar-sweetened beverages), hypertension, dyslipidaemia and baseline and changes in waist circumference. * P-value <0.05.

Parallel change analysis

Changes in low-fat dairy and low-fat milk were positively associated with changes in FPG ($\beta_{serving/day}$=0.02, 0.00-0.04, and 0.03, 0.01-0.05, respectively) and 2hPG (0.04, 0.00-0.08 and 0.06, 0.01-0.11, respectively) (**Supplemental Figure 3, Supplemental Table 6**). Changes in high-fat milk were inversely associated with changes in FPG ($\beta_{serving/day}$=-0.03, -0.05 to 0.00). Upon exclusion of participants who consumed non-specific milk with unknown fat content at baseline or follow-up (*n* = 2,162, 29%), these associations with FPG and 2hPG were no longer evident. Other associations of changes in dairy intake during follow-up and changes in FPG, OGT and HbA1c were not significant.

Linear mixed models

Using mixed models, a similar positive association for low-fat milk ($\beta_{serving/day*time}$ 0.003 mmol/L, 95%CI 0.0003, 0.005) was found (**Supplemental Table 9**). Furthermore, total dairy, milk, and butter were positively associated with FPG over follow-up, although effect sizes were modest ($\beta_{serving/day*time}$ 0.002-0.003 mmol/L). In contrast to the parallel change analysis, low-fat dairy and high-fat milk were not associated with FPG. Associations for 2hPG were different in comparison to the parallel change analysis. Fermented dairy ($\beta_{serving/day*time}$ 0.01 mmol/L, 95%CI 0.001, 0.02), high-fat fermented dairy ($\beta_{serving/day*time}$ 0.01 mmol/L, 95%CI 0.001, 0.03) and butter were positively associated with 2hPG over follow-up, but low-fat dairy and low-fat milk were not. Additionally, low-fat dairy ($\beta_{serving/day*time}$ 0.002 mmol/L, 95%CI 0.0005, 0.003), low-fat milk ($\beta_{serving/day*time}$ 0.002 mmol/L, 95%CI 0.001, 0.003), and ice cream ($\beta_{serving/day*time}$ 0.02 mmol/L, 95%CI 0.003, 0.03) were positively associated with HbA1c over follow-up. Inverse associations were found for high-fat dairy ($\beta_{serving/day*time}$ -0.002 mmol/L, 95%CI -0.003, -0.0001), and high-fat milk ($\beta_{serving/day*time}$ -0.004 mmol/L, 95%CI -0.006, -0.001), and HbA1c over follow-up.

Dairy fat biomarkers

Additional adjustment for C14:0 attenuated the associations of high-fat dairy (% attenuation: 16.2%, 95%CI 4.2-95.5%) and high-fat milk (8.5%, not significant) with prediabetes risk (**Supplemental Table 10**). Associations were slightly stronger when adjusting for C15:0, C17:0, *t*C16:1n7, and all dairy fat biomarkers together (attenuation ranging from -5.5% to -15.9%), but bootstrapped confidence intervals did not indicate significant differences. The dairy fat biomarkers had no relative contribution to inverse associations of low-fat milk and prediabetes.

Network models

The network of conditional independent relationships of food groups and dairy types showed several foods that are frequently consumed together within an eating pattern, as indicated by positive (blue) edges (i.e., conditional independent relationships), or not, as indicated by negative (orange) edges (**Supplemental figure 4**). High-fat dairy types were connected by positive edges, except for high-fat milk and high-fat cheese which were negatively connected. High-fat milk and high-fat yogurt were negatively connected

with their low-fat counterpart. Low-fat cheese was only positively connected with low-fat yogurt, which were also positively connected with fruit, vegetables, fish, legumes, and whole grains. Meat, potatoes, refined grains, SSBs, high-fat cheese and vegetables were positively connected and somewhat connected with high-fat milk, low-fat milk, and ice-cream. In contrast, negative connections between beverages including tea, coffee, and SSBs were observed.

The complete network structure showed that prediabetes was only weakly connected to waist circumference and TAG (**Supplemental figure 5**). Waist circumference, LDL-C, hypertension, and TAG clustered, bridging prediabetes and sociodemographic characteristics and health factors, further connecting with the food groups. Physical activity clustered strongly with waist circumference, age, and sex. Food groups with high energy density clustered around energy intake, including whole- and refined grains, meat, nuts and SSBs, as well as high-fat and low-fat milk, high-fat cheese, and ice cream. Also, alcohol intake strongly correlated with energy intake. A link between milk intake and ethnicity was observed. Furthermore, a cluster of fruit, vegetables, fish, legumes, high-fat yogurt, low-fat yogurt, and low-fat cheese was observed. Of continuous nodes, energy intake (mean explained variance 0.80 as indicated by the blue rings in the figures), refined grains (0.54), age at finishing education (0.43), waist circumference (0.41), meat (0.38), and vegetables (0.36) had the highest predictability. The individual dairy types and fish had the lowest predictability. Most categorical nodes, including for example prediabetes, ethnicity, and family history of diabetes, had a high accuracy due to their marginal probability. For sex and education, the network provided some additional information.

In this network, energy intake, sex, age, waist circumference, refined grains, and meat, showed the highest centrality (**Supplemental figure 6**). Also, alcohol had high strength and closeness, while vegetables had high betweenness. Of dairy types, high-fat cheese had the highest centrality. Prediabetes, LDL-c, high-fat yogurt, and low-fat cheese showed low centrality compared to other nodes. None of the dairy types were directly connected with prediabetes or the clinical markers, only high-fat cheese was weakly connected to hypertension. Fruit and fish connected to waist circumference, and meat and coffee to TAG.

Results on clustering, edge–weight accuracy and the semiparametric copula network can be found in **Supplemental figure 7-9**.

Discussion

In this large prospective cohort with 6.7 years of follow-up, most dairy types showed no significant association with prediabetes risk. A higher baseline intake of high-fat dairy and high-fat milk were associated with higher prediabetes risk, after adjustment for potential confounders and background diet. Specific high-fat dairy types and cheese

exhibited similar, but imprecise associations. In contrast, higher low-fat milk intake was associated with lower prediabetes risk. Myristic acid (C14:0) of plasma phospholipids contributed to elucidating the positive associations between high-fat dairy and high-fat milk intakes and prediabetes risk. Total dairy, fermented dairy, low-fat fermented dairy, milk, yogurt and low-fat yogurt, cream, ice cream and butter showed little association with prediabetes risk. Changes in high-fat milk during follow-up were inversely associated with the risk of progressing to prediabetes or T2D.

Findings in context

Our finding of a positive association between high-fat dairy and high-fat milk with prediabetes risk was not in line with neutral or inverse associations reported in earlier observational cohort studies [16-20]. In the Australian AusDiab study (*n* = *4*,891), inverse associations for high-fat dairy, high-fat milk and total cheese with prediabetes risk were found [20]. Similarly, the United States Framingham Offspring Cohort (*n* = 1,867) showed inverse associations for total dairy, while the Dutch Rotterdam Study (*n* = 6,053) showed an inverse association for high-fat milk and high-fat yogurt [18]. Moreover, the observed inverse association of low-fat milk in our study partly contradicts the results of prior observational studies [17-19]. Some associations were heterogeneous by age and educational level. Low-fat dairy was only associated with lower risk in people younger than 50 years. Furthermore, total dairy, fermented dairy, low-fat fermented dairy, and yogurt were only associated with lower prediabetes risk in participants with a high educational level. Younger adults were slightly higher educated (41.2% versus 35.2%) and this subgroup might have been more likely to adapt healthier behaviours including the selection of low-fat dairy foods and yogurt to adhere to dietary guidelines, which could explain lower prediabetes risk in this subgroup. In the Lifelines study, high-fat yogurt intake was associated with a lower risk of prediabetes only in participants with a high educational level [19]. These disparities with prior literature highlight the complexity of the relationship between dairy and prediabetes. They reflect the influence of confounding of population characteristics, geographical location, dairy consumption habits and methodological differences (e.g., prediabetes definition) and their interactions in these associations.

We found that plasma phospholipid fatty acid myristic acid (C14:0), a correlate of dairy intake, [47] and linked to higher T2D risk [48], partly explained the positive association between high-fat dairy intake and prediabetes risk. Additionally, the dairy fat biomarkers C15:0, C17:0, *t*C16:1n7 behaved as a positive confounder for the association between high-fat dairy intake and prediabetes risk, although their relative contribution was small. Our results are in line with previously reported opposing associations for C14:0 and C17:0 with T2D [46, 57], strengthening the robustness of our findings based on self-reported dairy intake. Nevertheless, it remains uncertain whether these dairy fat biomarkers truly reflect intake of dairy, other ruminant foods, or if they are influenced by confounding health behaviours [58]. Furthermore, C14:0 is derived from both exogeneous and

endogenous synthesis, although de novo lipogenesis contributes only minimally to C14:0 concentrations. Further research is needed to identify a biologically sound explanation for these findings, accounting for behavioural factors correlated with low-fat or high-fat dairy consumption and other non-fat dairy constituents, such as protein, vitamin D and calcium [59].

The lack of associations in changes in high-fat dairy and low-fat milk, combined with inverse associations for high-fat milk with glycaemic status shows that baseline dairy intake and changes of dairy intake signify distinct exposure statuses. Participants who initially consumed higher amounts of dairy were more likely to reduce their intake during follow-up (regression to the mean), and a negative correlation between baseline and changes in dairy intake might result in contradicting estimates. Changes in intakes of total and low-fat fermented dairy, and yogurt were only positively associated with risk of progressing to prediabetes or T2D in participants with a high waist circumference, indicating heterogeneity of associations in different subgroups compared to associations with baseline dairy intake and prediabetes. Diabetes risk awareness, or desire to lose weight, especially among those with a high waist circumference, may drive people to switch from high-fat to low-fat dairy intake to adhere to dietary guidelines in an attempt to mitigate their risk [19]. Thus, consumption of total and low-fat fermented dairy, and yogurt may be associated with metabolic changes or behaviours that contribute to an elevated risk of progressing to prediabetes or T2D. In this cohort, reverse causation might have been more present in the context of dietary change patterns.

In our secondary analysis, we found that changes in low-fat dairy and low-fat milk were positively associated with changes in FPG and 2hPG during follow-up, while high-fat milk was inversely associated with FPG changes. Other associations of changes in different dairy subtypes with parallel changes in FPG, 2hGT and HbA1c were not significant. In the UK EPIC-Norfolk Study *(n* = 15,612, mean follow up 3.7 ± 0.7 years), changes in high-fat milk intake were positively associated with changes in HbA1c (0.52, 95%CI 0.06 - 0.97 mmol/L), but other glycaemic markers were not assessed [28]. Possible explanations for this inconsistency may be the larger increase in HbA1c in the EPIC-Norfolk study compared to our study, and opposite changes in milk consumption patterns. Thus, although parallel-change analyses in observational studies may be more consistent with evidence from randomized controlled trials as they capture dynamic changes over time [21], sufficient change in exposure and outcome is needed to be able to detect meaningful associations. Randomized controlled trials on the effects of milk on glycaemic parameters have shown null effects, but evidence is limited [60]. Most interventions have used a mixture of dairy types, limiting the differentiation of effects by fat content. A meta-analysis of 34 studies showed that high dairy diets (3.1 servings/day) were associated with increased FPG compared with low dairy diets (0.5 serving/day), but not in RCTs with sufficient duration (≥24 weeks), with energy restriction and FPG as primary outcome, and the evidence was graded low [61]. A small inverse association was found

for elevated dairy intake with HbA1c in four studies with sufficient duration (≥24 weeks) (*n* = 512; MD: −0.09%; 95% CI: −0.16%, −0.03%; p = 0.005, I^2 = 0%), but not with energy restriction and HbA1c as primary outcome. No evidence for causal effects was found for dairy and glycaemic markers using Mendelian randomization analysis [62].

Strengths and limitations

This study has multiple strengths including the large sample size, inclusion of important covariates including objectively measured physical activity, and sensitivity analyses showing robustness of results. We compared two exposures, dairy subtypes at baseline and changes over time, enabling comparison of different analytical strategies within a single cohort which has not been done before. The results should be interpreted considering limitations. First, although the FFQ was examined for its validity in a similar sample [26-28], and we adjusted for BMI that might cause dietary misreporting [63], the diet was self-reported and therefore prone to misreporting. Furthermore, the FFQ included the main dairy groups but lacked detailed information on nutrient contents of dairy types. Therefore, assumptions about the fat content of milk were made, and it is unknown how differences in sugar content of dairy types may have influenced our results. Second, low consumption for some dairy types, especially high-fat dairy types, limited the assessment of (non-linear) associations at high intake levels. Third, although we adjusted for many confounders, residual confounding cannot be ruled out due to the observational nature of our study. In particular, unmeasured conditions and behaviours linked to consuming high-fat and low-fat dairy products and changing each consumption might confound our findings substantially. Fourth, we performed multiple tests of correlated dietary intakes on correlated outcomes which might inflate type I error rate. Lastly, the response rate was 27% at baseline and the sample was of largely white European origin with higher dairy consumption compared to the general UK population [64], limiting external generalisability to other ethnic groups with different consumption patterns.

Conclusion

In this prospective analysis of a large cohort, most dairy types were not associated with prediabetes and progression in glycaemic status. Inconsistent with prior observational studies, high-fat dairy and high-fat milk were positively associated with prediabetes risk, and low-fat milk was inversely associated with prediabetes risk. Furthermore, high baseline high-fat milk intake may relate to decreased consumption of high-fat milk intake during follow-up resulting in the inverse association between changes in high-fat milk intake and the risk of progression from normoglycaemia to prediabetes or diabetes. Our results underscore the complexity of these associations and support the need for well-designed trials and observational studies to elucidate the potential effects of dairy products in the prevention of prediabetes.

List of supplementary materials chapter 6

Supplemental table 1. Missing values of covariates in participants before imputation (*n* = 7,521).

Supplemental figure 1. Distribution of dairy food intake in the diet (weight percentage, %) (*n* = 7,521).

Supplemental table 2. Baseline characteristics of participants in the Fenland study across different population subgroups (*n* =7,521).

Supplemental table 3. Pearson correlation coefficients of dairy types with blood markers (*n* = 7,521).

Supplemental table 4. Sensitivity analyses of associations of dairy product types and incident prediabetes in the Fenland study (*n* = 6,639).

Supplemental table 5. Associations of dairy intake and incident prediabetes in the Fenland study, stratified by age and educational level (*n* = 6,639).

Supplemental table 6. Sensitivity analyses of associations of increases in dairy product type intake and glycaemic status and markers in the Fenland study (*n* = 7,410).

Supplemental figure 2. Non-linear association between changes in high-fat fermented dairy and high-fat milk intake and glycaemic status in the Fenland study.

Supplemental table 7. Associations of changes in dairy product type intake and glycaemic status in the Fenland study, stratified by age or waist circumference (*n* = 7,410).

Supplemental figure 3. Associations of changes in dairy intake from baseline to follow-up in servings/day with changes in glycaemic markers (FPG, 2hPG, HbA1c) (*n* = 7,410).

Supplemental table 8. Associations of dairy intake and prediabetes risk in the Fenland study using different prediabetes definitions (*n* = 6,639).

Supplemental table 9. Longitudinal associations of repeated measures of dairy intake and glycaemic outcomes at baseline and follow-up in the Fenland study.

Supplemental table 10. Contribution of selected dairy fat biomarkers to the association of dairy type intake and prediabetes risk in participants with complete data of dairy fat biomarkers (*n* = 6,418).

Supplemental figure 4. Network structure of dairy intake and food groups of the study population with complete data for variables in the model (*n* = 6,639).

Supplemental figure 5. Network structure of dairy intake, food groups and energy intake, health factors, sociodemographic characteristics, clinical markers, and prediabetes of the study population with complete data for variables in the model (*n* = 6,162).

Supplemental figure 6. Standardized centrality of each variable in the network as indicated by the strength, closeness and betweenness.

Supplemental figure 7. Clusters of variables that strongly connect in the network.

Supplemental figure 8. Plot of bootstrapped sampling variation around the edge-weights reflecting accuracy of the edge-weights, a measure of the stability of the network.

Supplemental figure 8. Plot of bootstrapped sampling variation around the edge-weights reflecting accuracy of the edge-weights, a measure of the stability of the network.

Supplemental figure 9. Network structure of dairy types, food groups and energy intake, health factors, socio-demographic characteristics, clinical markers and prediabetes of the study population with complete data for variables in the model with a nonparanormal transformation applied to non-normally distributed continuous variables (*n* = 6,162).

Scan this QR code to download the supplementary materials.

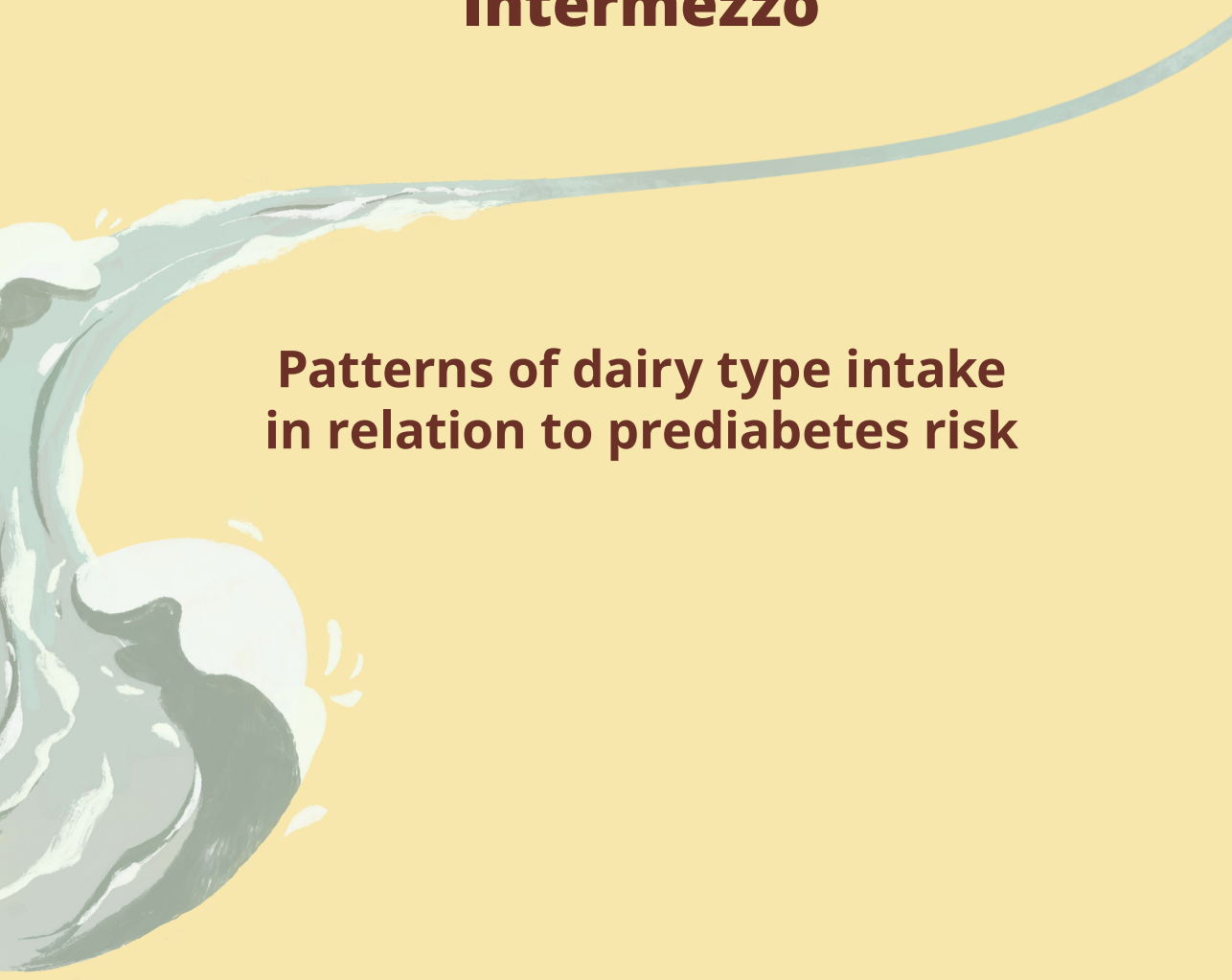
References

1. World Health Organization (WHO), *Definition and diagnosis of diabetes mellitus and intermediate hyperglycaemia: report of a WHO/IDF consultation.* 2006.
2. Gerstein, H.C., et al., *Annual incidence and relative risk of diabetes in people with various categories of dysglycemia: a systematic overview and meta-analysis of prospective studies.* Diabetes Res Clin Pract, 2007. **78**(3): p. 305-12.
3. Yeboah, J., et al., *Impaired fasting glucose and the risk of incident diabetes mellitus and cardiovascular events in an adult population: MESA (Multi-Ethnic Study of Atherosclerosis).* Journal of the American College of Cardiology, 2011. **58**(2): p. 140-146.
4. Mutie, P.M., et al., *An investigation of causal relationships between prediabetes and vascular complications.* Nat Commun, 2020. **11**(1): p. 4592.
5. Cai, X., et al., *Association between prediabetes and risk of all cause mortality and cardiovascular disease: updated meta-analysis.* Bmj, 2020. **370**: p. m2297.
6. Danaei, G., et al., *National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2·7 million participants.* Lancet, 2011. **378**(9785): p. 31-40.
7. *Effects of diabetes definition on global surveillance of diabetes prevalence and diagnosis: a pooled analysis of 96 population-based studies with 331,288 participants.* Lancet Diabetes Endocrinol, 2015. **3**(8): p. 624-37.
8. Liu, C., et al., *Trends in Self-reported Prediabetes and Metformin Use in the USA: NHANES 2005-2014.* J Gen Intern Med, 2020. **35**(1): p. 95-101.
9. Khan, M.A.B., et al., *Epidemiology of type 2 diabetes–global burden of disease and forecasted trends.* Journal of epidemiology and global health, 2020. **10**(1): p. 107.
10. Richter, B., et al., *Development of type 2 diabetes mellitus in people with intermediate hyperglycaemia.* Cochrane Database of Systematic Reviews, 2018(10).
11. Perreault, L., et al., *Effect of regression from prediabetes to normal glucose regulation on long-term reduction in diabetes risk: results from the Diabetes Prevention Program Outcomes Study.* The Lancet, 2012. **379**(9833): p. 2243-2251.
12. Glechner, A., et al., *Effects of lifestyle changes on adults with prediabetes: A systematic review and meta-analysis.* Primary Care Diabetes, 2018. **12**(5): p. 393-408.
13. Galaviz, K.I., et al., *Interventions for Reversing Prediabetes: A Systematic Review and Meta-Analysis.* American Journal of Preventive Medicine, 2022. **62**(4): p. 614-625.
14. Yu, E. and F.B. Hu, *Dairy Products, Dairy Fatty Acids, and the Prevention of Cardiometabolic Disease: a Review of Recent Evidence.* Curr Atheroscler Rep, 2018. **20**(5): p. 24.
15. Weaver, C.M., *Dairy matrix: is the whole greater than the sum of the parts?* Nutr Rev, 2021. **79**(Suppl 2): p. 4-15.
16. Hruby, A., et al., *Associations of Dairy Intake with Incident Prediabetes or Diabetes in Middle-Aged Adults Vary by Both Dairy Type and Glycemic Status.* J Nutr, 2017. **147**(9): p. 1764-1775.
17. Slurink, I.A.L., et al., *Dairy product consumption and incident prediabetes in Dutch middle-aged adults: the Hoorn Studies prospective cohort.* Eur J Nutr, 2022. **61**(1): p. 183-196.
18. Slurink, I.A.L., et al., *Dairy Product Consumption in Relation to Incident Prediabetes and Longitudinal Insulin Resistance in the Rotterdam Study.* Nutrients, 2022. **14**(3).
19. Slurink, I.A.L., et al., *Dairy consumption and incident prediabetes: prospective associations and network models in the large population-based Lifelines study.* The American Journal of Clinical Nutrition.
20. Slurink, I.A., et al., *Dairy Product Consumption and Incident Prediabetes in the Australian Diabetes, Obesity, and Lifestyle Study With 12 Years of Follow-Up.* J Nutr, 2023. **153**(6): p. 1742-1752.
21. Smith, J.D., et al., *A Comparison of Different Methods for Evaluating Diet, Physical Activity, and Long-Term Weight Gain in 3 Prospective Cohort Studies.* J Nutr, 2015. **145**(11): p. 2527-34.
22. Willett, W.C., G.R. Howe, and L.H. Kushi, *Adjustment for total energy intake in epidemiologic studies.* The American journal of clinical nutrition, 1997. **65**(4): p. 1220S-1228S.

23. *International Expert Committee report on the role of the A1C assay in the diagnosis of diabetes.* Diabetes Care, 2009. **32**(7): p. 1327-34.
24. Bingham, S.A., et al., *Nutritional methods in the European prospective investigation of cancer in Norfolk.* Public health nutrition, 2001. **4**(3): p. 847-858.
25. Mulligan, A.A., 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): p. e004503.
26. Bingham, S.A., et al., *Validation of dietary assessment methods in the UK arm of EPIC using weighed records, and 24-hour urinary nitrogen and potassium and serum vitamin C and carotenoids as biomarkers.* International journal of epidemiology, 1997. **26**(suppl_1): p. S137.
27. Harding, A.-H., et al., *Fat consumption and HbA1c levels: the EPIC-Norfolk study.* Diabetes care, 2001. **24**(11): p. 1911-1916.
28. Trichia, E., et al., *The associations of longitudinal changes in consumption of total and types of dairy products and markers of metabolic risk and adiposity: findings from the European Investigation into Cancer and Nutrition (EPIC)-Norfolk study, United Kingdom.* Am J Clin Nutr, 2020. **111**(5): p. 1018-1026.
29. Trichia, E., et al., *Associations of types of dairy consumption with adiposity: cross-sectional findings from over 12 000 adults in the Fenland study, UK.* British Journal of Nutrition, 2019. **122**(8): p. 928-935.
30. de Oliveira Otto, M.C., et al., *Dietary intake of saturated fat by food source and incident cardiovascular disease: the Multi-Ethnic Study of Atherosclerosis.* The American journal of clinical nutrition, 2012. **96**(2): p. 397-404.
31. Vissers, L.E., et al., *Fatty acids from dairy and meat and their association with risk of coronary heart disease.* European journal of nutrition, 2019. **58**(7): p. 2639-2647.
32. Monsivais, P., et al., *Greater accordance with the Dietary Approaches to Stop Hypertension dietary pattern is associated with lower diet-related greenhouse gas production but higher dietary costs in the United Kingdom.* The American journal of clinical nutrition, 2015. **102**(1): p. 138-145.
33. Bach-Faig, A., et al., *Mediterranean diet pyramid today. Science and cultural updates.* Public health nutrition, 2011. **14**(12A): p. 2274-2284.
34. Tong, T.Y.N., et al., *Prospective association of the Mediterranean diet with cardiovascular disease incidence and mortality and its population impact in a non-Mediterranean population: the EPIC-Norfolk study.* BMC Medicine, 2016. **14**(1): p. 135.
35. Maguire, E.R. and P. Monsivais, *Socio-economic dietary inequalities in UK adults: an updated picture of key food groups and nutrients from national surveillance data.* British Journal of Nutrition, 2015. **113**(1): p. 181-189.
36. Brage, S., et al., *Reliability and validity of the combined heart rate and movement sensor Actiheart.* European journal of clinical nutrition, 2005. **59**(4): p. 561-570.
37. Brage, S., et al., *Hierarchy of individual calibration levels for heart rate and accelerometry to measure physical activity.* Journal of Applied Physiology, 2007. **103**(2): p. 682-692.
38. Clifton, E.A., et al., *Associations between body mass index-related genetic variants and adult body composition: the Fenland cohort study.* International journal of obesity, 2017. **41**(4): p. 613-619.
39. Watson, L.P., M.C. Venables, and P.R. Murgatroyd, *An investigation into the differences in bone density and body composition measurements between 2 GE lunar densitometers and their comparison to a 4-component model.* Journal of Clinical Densitometry, 2017. **20**(4): p. 498-506.
40. O'Connor, L., et al., *Intakes and sources of dietary sugars and their association with metabolic and inflammatory markers.* Clinical Nutrition, 2018. **37**(4): p. 1313-1322.
41. De Lucia Rolfe, E., et al., *Association between birth weight and visceral fat in adults.* The American Journal of Clinical Nutrition, 2010. **92**(2): p. 347-352.
42. Organization., W.H., *Hypertension. 2021. Available from: https://www.who.int/news-room/fact-sheets/detail/hypertension.*
43. Tong, T.Y., et al., *A combination of metabolites predicts adherence to the Mediterranean diet pattern and its associations with insulin sensitivity and lipid homeostasis in the general population: the Fenland study, United Kingdom.* The Journal of nutrition, 2020. **150**(3): p. 568-578.
44. Friedewald, W.T., R.I. Levy, and D.S. Fredrickson, *Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge.* Clinical chemistry, 1972. **18**(6): p. 499-502.

45. Mach, F., et al., *2019 ESC/EAS Guidelines for the management of dyslipidaemias: lipid modification to reduce cardiovascular risk: the Task Force for the management of dyslipidaemias of the European Society of Cardiology (ESC) and European Atherosclerosis Society (EAS).* European heart journal, 2020. **41**(1): p. 111-188.
46. Imamura, F., 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. **15**(10): p. e1002670.
47. Pranger, I.G., et al., *Fatty acids as biomarkers of total dairy and dairy fat intakes: a systematic review and meta-analysis.* Nutrition reviews, 2019. **77**(1): p. 46-63.
48. Huang, L., et al., *Circulating saturated fatty acids and incident type 2 diabetes: a systematic review and meta-analysis.* Nutrients, 2019. **11**(5): p. 998.
49. Wang, L.Y., et al., *Development and validation of a robust automated analysis of plasma phospholipid fatty acids for metabolic phenotyping of large epidemiological studies.* Genome Med, 2013. **5**(4): p. 39.
50. Knol, M.J., et al., *Overestimation of risk ratios by odds ratios in trials and cohort studies: alternatives to logistic regression.* CMAJ, 2012. **184**(8): p. 895-9.
51. Greenland, S. and N. Pearce, *Statistical foundations for model-based adjustments.* Annual review of public health, 2015. **36**: p. 89-108.
52. VanderWeele, T.J., *Principles of confounder selection.* European journal of epidemiology, 2019. **34**: p. 211-219.
53. Schwingshackl, L., et al., *Food groups and risk of type 2 diabetes mellitus: a systematic review and meta-analysis of prospective studies.* Eur J Epidemiol, 2017. **32**(5): p. 363-375.
54. Stringhini, S., et al., *Contribution of modifiable risk factors to social inequalities in type 2 diabetes: prospective Whitehall II cohort study.* Bmj, 2012. **345**.
55. Richiardi, L., R. Bellocco, and D. Zugna, *Mediation analysis in epidemiology: methods, interpretation and bias.* International journal of epidemiology, 2013. **42**(5): p. 1511-1519.
56. Sterne, J.A., et al., *Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls.* Bmj, 2009. **338**.
57. Forouhi, N.G., 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.* The Lancet Diabetes & Endocrinology, 2014. **2**(10): p. 810-818.
58. Drouin-Chartier, J.P., et al., *Dairy consumption, plasma metabolites, and risk of type 2 diabetes.* Am J Clin Nutr, 2021. **114**(1): p. 163-174.
59. Mozaffarian, D. and J.H.Y. Wu, *Flavonoids, Dairy Foods, and Cardiovascular and Metabolic Health: A Review of Emerging Biologic Pathways.* Circ Res, 2018. **122**(2): p. 369-384.
60. Kiesswetter, E., et al., *Effects of Dairy Intake on Markers of Cardiometabolic Health in Adults: A Systematic Review with Network Meta-Analysis.* Advances in Nutrition, 2023. **14**(3): p. 438-450.
61. O'Connor, S., et al., *Increased Dairy Product Intake Modifies Plasma Glucose Concentrations and Glycated Hemoglobin: A Systematic Review and Meta-Analysis of Randomized Controlled Trials.* Advances in Nutrition, 2019. **10**(2): p. 262-279.
62. *Dairy Intake and Body Composition and Cardiometabolic Traits among Adults: Mendelian Randomization Analysis of 182041 Individuals from 18 Studies.* Clin Chem, 2019. **65**(6): p. 751-760.
63. Trijsburg, L., 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 nutrition, 2017. **20**(4): p. 598-607.
64. Page P, Roberts C, Steer T, Mablethorpe N, Cox L, Meadows S, et al. National diet and nutrition survey results from years 7 and 8 (combined) of the rolling programme (2014/2015 to 2015/2016). 2018.
65. World Health Organization (WHO), *Waist circumference and waist-hip ratio: report of a WHO expert consultation. Geneva, World Health Organization (WHO), 2008*. 2006.
66. Fortunato, S., *Community detection in graphs.* Physics reports, 2010. **486**(3-5): p. 75-174.
67. Yang, Z., R. Algesheimer, and C.J. Tessone, *A Comparative Analysis of Community Detection Algorithms on Artificial Networks.* Scientific Reports, 2016. **6**(1): p. 30750.

Intermezzo



Patterns of dairy type intake in relation to prediabetes risk

Introduction

The umbrella term 'total dairy intake' includes many food items, each with different nutritional values and associated eating behaviours. Analysing patterns of dairy food consumption provides more insight into which different dairy types are combined within an individual's diet or if there are distinct patterns in their consumption [1], and the cumulative and interactive associations of dairy foods on prediabetes risk. Therefore, we aimed to describe dairy consumption patterns and relate these to prediabetes risk.

Methods

Using data from participants in the Lifelines study (*n* = 74,132, **Chapter 5**) and Fenland study (*n* = 7,521, **Chapter 6**), principal component analysis (PCA) was conducted using the prcomp function in R to identify consumption patterns based on intake of eight dairy types (high-fat milk, low-fat milk, high-fat yogurt, low-fat yogurt, high-fat cheese, low-fat cheese, cream, and ice cream). In each cohort, we retained six patterns based on a cumulative explained variance of >0.80% (**Table 1**). The orthogonally rotated standardized pattern scores were divided into quartiles. Baseline characteristics for the top quartile of these pattern scores were shown. Poisson regression with robust variance was used to estimate the relative risk (RR) with 95% confidence interval (CI) [50] for associations of the pattern scores with prediabetes risk, with the pattern scores modelled as categorical variables based on quartiles, and on a continuous scale per standard deviation (SD) increase.

Results

Dairy consumption patterns

Table 1 shows the factor loading of each of the 8 identified principal components with their corresponding factor loadings. Four patterns emerged within both cohorts. First, a *'high-fat dairy'* pattern was observed, positively associated with high-fat milk, high-fat yogurt, and high-fat cheese intakes in the Lifelines study, and with high-fat milk intake in the Fenland study, explaining 19.5% and 19.7% of the total variability in dairy intake, respectively. Second, a pattern characterized by high intake of *'ice cream*' was identified in both cohorts, explaining 9.8% and 11.9% of the variance, respectively. Third, a *'yogurt'* pattern was evident in both cohorts, negatively associated with cheese intake in the Lifelines study (9.0% explained variance) and with low-fat cheese and cream in the Fenland study (9.1% explained variance). Fourth, a '*high-fat yogurt*' pattern (9.6% and 9.7% explained variance, respectively) exhibited negative associations with high-fat milk intake in the Lifelines cohort and with low-fat yogurt and cream intake in the Fenland cohort.

Table 1. Factor loadings (correlation coefficients) between the dairy types and consumption patterns of dairy types derived from principal component analysis[1].

	Lifelines study							
Dairy types	High-fat dairy pattern	Low-fat milk and (ice) cream pattern	Low-fat cheese pattern	Milk avoidance pattern	Low-fat yogurt pattern	Ice cream pattern	High-fat yogurt pattern	Yogurt pattern
High-fat milk	**0.41**	-0.26	**0.42**	-0.07	0.33	-0.06	**-0.64**	-0.25
Low-fat milk	-0.13	**0.41**	-0.24	**-0.79**	0.32	0.09	-0.17	0.00
High-fat yogurt	**0.47**	-0.18	0.28	-0.23	0.31	0.22	**0.57**	0.38
Low-fat yogurt	-0.30	0.34	0.12	**0.47**	**0.68**	-0.03	-0.06	0.29
High-fat cheese	**0.43**	0.19	**-0.46**	0.24	0.29	0.07	0.25	**-0.60**
Low-fat cheese	**-0.43**	0.02	**0.56**	-0.13	0.08	0.04	0.37	**-0.59**
Cream	0.31	**0.52**	0.25	-0.04	-0.18	**-0.73**	0.11	0.05
Ice cream	0.21	**0.56**	0.29	0.14	-0.33	**0.64**	-0.15	0.01
Eigenvalues[2]	1.56	1.23	1.08	0.94	0.91	0.79	0.77	0.72
Explained variance[3]	19.5%	15.4%	13.5%	11.8%	11.4%	9.8%	9.6%	9.0%
Cumulative explained variance[3]	19.5%	34.9%	48.3%	60.1%	71.5%	81.4%	91.0%	100%

Table 1. Factor loadings (correlation coefficients) between the dairy types and consumption patterns of dairy types derived from principal component analysis[1] (continued).

	Fenland study							
Dairy types	High-fat dairy pattern	Dairy diversity pattern	Low-fat avoidance pattern	Ice cream pattern	High-fat cheese pattern	High-fat yogurt pattern	Yogurt pattern	Milk pattern
High-fat milk	**0.63**	-0.03	-0.26	0.16	0.11	-0.11	0.09	**0.69**
Low-fat milk	**-0.62**	0.11	0.27	-0.17	0.02	0.01	-0.03	**0.71**
High-fat yogurt	0.19	**0.48**	0.03	**-0.41**	-0.38	**0.42**	**0.49**	0.01
Low-fat yogurt	-0.33	0.28	**-0.52**	0.13	0.11	**-0.50**	**0.51**	-0.07
High-fat cheese	0.10	**0.49**	0.24	0.04	**0.81**	0.15	0.0	-0.1
Low-fat cheese	-0.15	0.34	**-0.63**	0.05	-0.07	0.36	**-0.56**	0.04
Cream	0.21	**0.48**	0.23	-0.22	-0.26	**-0.62**	**-0.41**	-0.03
Ice cream	-0.04	0.30	0.28	**0.84**	**-**0.31	0.11	0.06	0.02
Eigenvalues[2]	1.58	1.51	1.11	0.95	0.81	0.78	0.72	0.53
Explained variance[3]	19.7%	18.9%	13.9%	11.9%	10.2%	9.7%	9.1%	6.6%
Cumulative explained variance[3]	19.7%	38.6%	52.5%	64.4%	74.6%	84.3%	93.4%	100%

[1] Factor loading ≥(-)0.40 are presented in bold and the magnitude of factor loadings is represented using different shades of grey, with darker shades indicating stronger relationships. A positive factor loading indicates that as the factor increases, the observed variable also increases. A negative factor loading indicates that as the factor increases, the observed variable decreases.

[2] Eigenvalues: the amount of variance explained by each pattern, indicating the significance of each factor in explaining the overall variance.

[3] Explained variance: the proportion of total variance in the original dataset that is explained by each pattern. Higher values indicate that the pattern captures a larger proportion of the variability in the dairy types.

Furthermore, each cohort revealed unique patterns. In the Lifelines study, a *'low-fat milk and (ice)cream*' pattern (15.4%) was identified as positively associated with low-fat milk, cream, and ice cream intakes. Also, a *'low-fat cheese'* pattern (13.5%) was identified as positively associated with low-fat cheese and high-fat milk intake but negatively correlated with high-fat cheese intake. Two patterns with positive associations with low-fat yogurt intake were found, a specific '*low-fat yogurt*' pattern (11.4%), and a '*milk avoidance*' pattern (13.5%) with a strong negative association with low-fat milk intake.

In the Fenland study, a *'dairy diversity'* pattern (18.9%) showed moderate associations with all dairy types except milk. A *'low-fat avoidance'* pattern (13.9%) exhibited moderate associations with all dairy types but negative associations with low-fat yogurt and low-fat cheese intake. A *'high-fat cheese'* pattern (10.2%) predominantly associated with high-fat cheese intake, was identified. Lastly, a *'milk'* pattern (6.6%) showed positive associations with both high-fat and low-fat milk intake.

There was some variation in baseline characteristics of individuals in each of the highest quartiles of the pattern scores (**Table 2a and b**). In the Lifelines study, participants in the highest quartile of the '*high-fat dairy'* pattern were younger, were more often current smoker, and had a lower diet quality and higher en% fat compared to participants in the highest quartiles of the other patterns as shown in **Table 2a**. In the Fenland study, the highest quartile of the '*high-fat dairy'* pattern did not exhibit notable differences compared to participants in the highest quartiles of the other patterns. Furthermore, participants in the highest quartile of the '*milk avoidance'* pattern in the Lifelines were more often female, with lower energy intake compared to the other top quartiles, and in the top '*low-fat yogurt'* quartile, physical activity and diet quality were slightly higher. In the Fenland study, participants in the highest quartile of the '*dairy diversity'* pattern were more often female, had a higher education, were less often current smokers, and had a higher diet quality compared to participants in the highest quartiles of the other patterns. Furthermore, participants in the highest quartile of the '*ice cream*' pattern in the Fenland study were less often highly educated and had a higher BMI, but this was not the case in the Lifelines study.

The main results of the associations analyses between the dietary patterns and prediabetes risk were as follows. In the Lifelines study, the '*high-fat dairy'* pattern was associated with lower prediabetes risk in model 3 (RR per 1 SD 0.96, 95%CI 0.93-0.99) but this association attenuated after further adjustment for cardiometabolic markers in model 4 (RR per 1 SD 0.99, 95%CI 0.96-1.02) (**Table 3**). Furthermore, the '*low-fat yogurt'* pattern was associated with higher prediabetes risk in model 4 (RR per 1 SD 1.05, 95%CI 1.01-1.09). In the Fenland study, none of the dairy consumption patterns were significantly associated with prediabetes risk, with estimates ranging from 0.96 (RR model 4, per 1 SD, 95%CI 0.86-0.1.07) for the '*low-fat avoidance'* pattern to 1.11 (RR model 4, per 1 SD, 95%CI 0.97-1.27) for the '*high-fat yogurt'* pattern (**Table 4**).

Table 2a. Baseline characteristics of participants of the Lifelines Cohort in the total cohort and according to the highest quartiles of the dairy food consumption patterns (*n* = 74,132).

	Highest quartile of dairy food consumption pattern					
	High-fat dairy	Low-fat milk and (ice)cream	Low-fat cheese	Milk avoidance	Low-fat yogurt	Ice cream
Total dairy intake, (servings/d)	3.5 ± 1.8	3.6 ± 1.6	3.1 ± 1.6	2.9 ± 1.6	3.7 ± 1.7	3.2 ± 1.7
Range	0.2-20.5	0.27-19.2	0.05-19.6	0.07-19.3	0.40-20.5	0.04-20.5
Follow-up time	4.1 ± 1.1	4.1 ± 1.1	4.1 ± 1.1	4.1 ± 1.0	4.2 ± 1.1	4.1 ± 1.1
Sex, female (%)	52.1	55.6	57.5	65.3	62.1	58.5
Age at baseline (y)	46.3 ± 11.6	45.4 ± 11.8	46.4 ± 12.6	47.1 ± 12.2	46.2 ± 11.7	45.8 ± 12.3
Educational level (%)						
Low	26.4	25.6	30.8	27.4	27.1	27.1
Intermediate	39.7	40.8	40.4	39.7	40.6	39.7
High	33.8	33.6	28.8	32.8	32.3	33.2
Smoking (%)						
Never	48.2	52.4	49.4	46.1	49.2	48.1
Former	33.1	34.3	34.5	37.6	35.1	34.8
Current	18.8	13.3	16.1	16.4	15.7	17.1
Pack years	13.6 [6.5, 21.8]	11.4 [5.1, 19.5]	12.3 [5.7, 20.8]	13.3 [6.6, 21.8]	12.0 [5.8, 20.0]	12.0 [5.5, 20.1]
Alcohol (g/d)	7.96 ± 9.12	7.11 ± 8.00	6.69 ± 8.12	7.60 ± 9.00	7.02 ± 8.26	7.47 ± 8.80
Physical activity (min/wk)	260 [100, 555]	270 [120, 570]	270 [120, 580]	255 [110, 540]	280 [120, 600]	270 [120, 560]
Family history of diabetes (%)						
No	29.5	27.2	28.5	29.4	27.3	28.9
Yes	8.1	7.6	8.2	8.3	7.7	8.2
Unknown/missing	62.4	65.3	63.4	62.2	64.9	63.0
BMI (kg/m2)	25.0 ± 3.7	25.8 ± 3.9	25.6 ± 3.8	25.6 ± 3.9	25.6 ± 3.9	25.6 ± 3.9
Waist circumference (cm)	88.5 ± 11.4	89.7 ± 11.4	89.2 ± 11.4	88.6 ± 11.5	89.0 ± 11.4	89.2 ± 11.5
Total cholesterol (mmol/L)	5.2 ± 1.0	5.1 ± 1.0	5.2 ± 1.0	5.2 ± 1.0	5.1 ± 1.0	5.1 ± 1.0
LDL cholesterol (mmol/L)	3.3 ± 0.9	3.3 ± 0.9	3.3 ± 0.9	3.3 ± 0.9	3.3 ± 0.9	3.3 ± 0.9
HDL cholesterol (mmol/L)	1.5 ± 0.4	1.5 ± 0.4	1.5 ± 0.4	1.6 ± 0.4	1.5 ± 0.4	1.5 ± 0.4
TAG (mmol/L)	0.9 [0.7, 1.3]	0.9 [0.7, 1.3]	1.0 [0.7, 1.3]	0.9 [0.7, 1.3]	0.9 [0.7, 1.3]	1.0 [0.7, 1.3]
Hypertension (%)	22.0	22.8	24.8	23.9	23.7	23.9

Table 2a. Baseline characteristics of participants of the Lifelines Cohort in the total cohort and according to the highest quartiles of the dairy food consumption patterns (*n* = 74,132) (continued).

	Highest quartile of dairy food consumption pattern					
	High-fat dairy	Low-fat milk and (ice)cream	Low-fat cheese	Milk avoidance	Low-fat yogurt	Ice cream
Dietary intake						
Energy intake (kcal/d)	2278 ± 600	2260 ± 570	2109 ± 611	1989 ± 572	2119 ± 591	2059 ± 563
Diet quality[1]	23.3 ± 5.7	24.3 ± 5.6	24.3 ± 6.1	24.3 ± 6.3	25.5 ± 5.8	24.7 ± 5.9
Fruit (g/d)	131 ± 110	142 ± 105	147 ± 113	148 ± 116	154 ± 112	141 ± 109
Vegetables (g/day)	107 ± 59.3	107 ± 56.9	106 ± 58.4	108 ± 60.0	110 ± 58.3	104 ± 57.4
Bread (g/day)	121 ± 63.2	117 ± 58.7	116 ± 62.0	109 ± 60.4	118.2 ± 60.1	114.7 ± 59.2
Legumes (g/day)	10.8 ± 16.1	9.91 ± 14.5	10.0 ± 15.4	9.2 ± 15.0	10.2 ± 15.2	9.7 ± 14.9
Nuts (g/day)	14.4 ± 15.7	14.1 ± 14.7	12.8 ± 14.7	12.8 ± 14.9	12.8 ± 14.5	12.8 ± 14.4
Meat (red and processed) (g/day)	69.6 ± 34.9	70.3 ± 32.0	67.2 ± 33.3	64.6 ± 33.6	67.4 ± 32.2	66.6 ± 32.4
Fish (g/day)	12.5 ± 12.4	12.9 ± 11.9	12.8 ± 12.5	13.0 ± 13.4	12.5 ± 12.5	12.6 ± 12.4
Coffee (g/day)	448 ± 276	437 ± 269	431 ± 273	427 ± 277	442 ± 265	426 ± 267
Tea (g/day)	248 ± 243	250 ± 235	258 ± 245	281 ± 264	272 ± 248	254 ± 244
SSBs (g/day)	162 ± 180	148 ± 164	142 ± 175	131 ± 170	129 ± 155	138 ± 166
Total fat (en%)	36.4 ± 4.6	35.2 ± 4.3	34.4 ± 4.8	34.6 ± 5.1	34.4 ± 4.6	34.5 ± 4.7
Protein (en%)	14.7 ± 2.0	15.1 ± 2.0	15.3 ± 2.4	15.3 ± 2.5	15.8 ± 2.3	15.4 ± 2.2
Carbohydrates (en%)	44.8 ± 5.6	45.8 ± 5.0	46.4 ± 5.4	45.6 ± 6.1	45.8 ± 5.3	45.9 ± 5.6
Added sugar (g/day)	60.7 ± 34.9	58.8 ± 31.5	55.3 ± 34.1	50.1 ± 32.4	51.0 ± 31.3	51.3 ± 31.1
Calcium (mg/day)	1081 ± 368	1119 ± 350	1040 ± 367	955 ± 37.9	1147 ± 368	1024 ± 347

Values are mean ± SD for continuous variables with a normal distribution, or median [IQR] for continuous variables with a skewed distribution, percentages for categorical variables, based on unimputed data.

[1] Diet quality was measured using the Lifelines Diet Score (LLDS), reflecting adherence to the 2015 Dutch Dietary Guidelines for prevention of chronic disease, with higher scores indicating better adherence [30]. The LLDS consists of 12 food groups, and possible scores range from 0 to 48.

Abbreviations: BMI, body mass index; en%, percentage of total energy intake; HDL, high-density lipoprotein; LDL, low-density lipoprotein; SD, standard deviation; SSB, sugar-sweetened beverage; TAG, triglyceride.

Table 2b. Baseline characteristics of participants in the Fenland study according to the highest quartiles of the dairy food consumption patterns in the Fenland study (*n* = 7,521).

	Highest quartile of dairy food consumption pattern					
	High-fat dairy	Dairy diversity	Low-fat avoidance	Ice cream	High-fat cheese	High-fat yogurt
Total dairy intake (servings/d)	2.3 ± 1.3	3.4 ± 1.2	2.6 ± 1.1	2.7 ± 1.2	2.8 ± 1.2	2.6 ± 1.2
Range	0-11.6	0.9-11.6	0.4-9.3	0.3-9.4	0.4-9.9	0.2-10.8
Follow-up time (years)	6.9 ± 1.9	6.6 ± 2.0	6.7 ± 2.1	7.0 ± 2.0	6.9 ± 2.0	6.6 ± 2.1
Age at baseline (years)	48.5 ± 7.3	48.5 ± 7.5	48.3 ± 7.4	48.6 ± 7.4	48.0 ± 7.4	48.5 ± 7.5
Sex, female (%)	51.1	58.0	39.6	47.7	49.9	48.1
Educational level (%)						
Intermediate	40.1	36.3	42.9	50.9	45.9	43.2
High	44.9	54.3	41.0	30.1	36.3	40.7
Age at finishing education (years)	18 [16, 22]	21 [17, 23]	18 [16, 22]	17 [16, 21]	18 [16, 21]	18 [16, 22]
Ethnic origin, non-White (%)	3.1	1.8	2.0	2.4	1.6	3.4
Working (%)	12.6	11.5	11.1	10.8	10.9	11.6
Household income (%)	12.6	8.5	8.5	12.1	12.0	11.5
£20,000-40,000	32.0	29.8	32.3	38.0	35.5	35.6
>£40,000	55.4	61.6	59.2	49.9	52.4	52.9
Marital status (%)						
Married	79.3	84.9	85.5	82.0	81.2	880.0
Widowed/separated	9.2	7.5	7.2	8.7	7.9	8.7
Smoking status (%)						
Former	33.4	31.6	31.9	34.1	33.7	33.5
Current	10.8	7.5	10.4	8.8	11.3	11.1
Pack-years among smokers	3,745 [1,389; 8,728]	2,931 [1,188; 6,547]	4,402 [1,924; 7,903]	3,912 [1,562; 8,045]	4,383 [2,322; 8,529]	3,698 [1,655; 7,854]
Alcohol intake (g/day)	6.4 [1.6, 13.2]	7.5 [2.8, 13.6]	7.5 [2.1, 14.5]	5.8 [0.9, 10.9]	6.4 [1.7, 13.4]	6.2 [1.7, 12.5]
Physical activity (kJ/kg/day)	54.8 ± 21.8	55.8 ± 21.6	55.9 ± 22.0	55.0 ± 21.8	55.5 ± 21.8	54.4 ± 21.5
Family history of diabetes (%)	22.4	20.7	23.0	24.3	23.6	24.0
BMI (kg/m2)	25.8 ± 4.3	25.8 ± 4.2	26.4 ± 4.3	27.0 ± 4.7	26.3 ± 4.4	26.3 ± 4.4
Body fat (%)	31.6 ± 8.4	32.6 ± 8.2	31.2 ± 8.0	32.8 ± 8.6	32.3 ± 8.4	32.1 ± 8.1

Table 2b. Baseline characteristics of participants in the Fenland study according to the highest quartiles of the dairy food consumption patterns in the Fenland study (*n* = 7,521) (continued).

	Highest quartile of dairy food consumption pattern					
	High-fat dairy	Dairy diversity	Low-fat avoidance	Ice cream	High-fat cheese	High-fat yogurt
Waist circumference (cm)	88.8 ± 12.9	88.3 ± 12.4	91.3 ± 12.7	91.8 ± 13.4	90.0 ± 13.1	90.2 ± 13.0
Hypertension (%)	34.6	32.8	37.2	38.2	36	38.4
Dyslipidaemia (%)	8.7	8.2	10.0	9.8	8.2	9.2
Dietary intake at baseline						
Energy intake (kcal/day)	1,960 ± 613	2,229 ± 561	2,057 ± 586	2,076 ± 605	2,020 ± 581	1,976 ± 584
DASH score	22.9 ± 4.7	24.8 ± 4.3	23.0 ± 4.3	23.6 ± 4.8	24.0 ± 4.5	24.2 ± 4.7
tMDS score	7.7 ± 1.6	8.1 ± 1.4	7.4 ± 1.5	7.4 ± 1.5	7.6 ± 1.5	7.7 ± 1.5
Fruit (g/day)	235 ± 184	276 ± 179	209 ± 157	258 ± 189	232 ± 164	238 ± 183
Vegetables (g/day)	274 ± 139	300 ± 128	251 ± 121	271 ± 134	271 ± 136	272 ± 142
Wholegrains (g/day)	84.3 ± 84.7	96.8 ± 86.6	72.6 ± 72.0	81.4 ± 83.2	84.6 ± 82.8	86.0 ± 86.7
Refined grains (g/day)	125 ± 76	144 ± 72	138 ± 74	128 ± 75	129 ± 73	127 ± 72
Meat (red and processed) (g/day)	60.9 ± 42.8	64.4 ± 41.5	70.6 ± 44.8	69.1 ± 43.8	62.8 ± 44.7	61.4 ± 46.3
Coffee (g/day)	283 ± 309	319 ± 299	313 ± 318	275 ± 313	296 ± 316	293 ± 317
Tea (g/day)	448 ± 371	494 ± 346	517 ± 372	462 ± 374	499 ± 376	487 ± 377
SSB (g/day)	110 ± 126	111 ± 107	110 ± 120	117 ± 129	103 ± 114	109 ± 128
Total fat (en%)	33.7 ± 5.7	33.8 ± 4.8	33.4 ± 5.1	32.3 ± 5.1	32.6 ± 5.2	32.5 ± 5.2
SFAs (en%)	12.8 ± 3.3	13.2 ± 2.6	12.9 ± 2.7	12.0 ± 2.8	12.3 ± 2.8	12.1 ± 2.8
Protein (en%)	17.5 ± 3.5	17.6 ± 3.0	17.4 ± 3.0	17.8 ± 3.5	18.0 ± 3.2	17.9 ± 3.4
Carbohydrate (en%)	48.9 ± 7.6	49.1 ± 6.3	49.1 ± 6.5	50.9 ± 6.4	49.6 ± 6.7	49.9 ± 6.9
Intrinsic and milk sugars (g/day)	57.6 ± 26.2	70.0 ± 24.9	57.9 ± 22.7	64.9 ± 26.7	61.9 ± 24.0	61.0 ± 26.3
Calcium (mg/day)	949 ± 361	1,204 ± 334	1,041 ± 327	1,051 ± 344	1,113 ± 340	1,041 ± 351
Sodium (mg/day)	1,962 ± 739	2,296 ± 714	2,066 ± 708	2,122 ± 737	2,155 ± 745	2,044 ± 746

Values are mean ± SD for continuous variables with a normal distribution, or median [IQR] for continuous variables with a skewed distribution, percentages for categorical variables, based on unimputed data.
Abbreviations: BMI, Body Mass Index; DASH, Dietary Approaches to Stop Hypertension; en%, percentage of total energy intake; SFA, Saturated Fatty acids; SSB, Sugar Sweetened Beverages; SD, Standard Deviation; tMDS, tertiles Mediterranean diet score.

Table 3. Associations of dairy food consumption pattern scores and prediabetes risk in the Lifelines study (*n* = 74,132).

	Relative risk (95% CI) across intake range categories[1]					Continuous[2]
	Q1	Q2	Q3	Q4	P_{trend}	RR (95%CI)
High-fat dairy pattern						
n cases/n total	738/17,977	651/18,467	651/18,595	706/19,092		
Model 1[1]	1 (ref)	0.94 (0.84-1.04)	0.90 (0.81-1.00)	0.90 (0.82-1.00)	0.05	0.96 (0.94-0.99)*
Model 2	1 (ref)	0.95 (0.86-1.05)	0.92 (0.83-1.02)	0.92 (0.83-1.02)	0.10	0.97 (0.94-1.00)*
Model 3	1 (ref)	0.93 (0.84-1.04)	0.90 (0.81-1.00)	0.90 (0.81-1.00)	0.05*	0.96 (0.93-0.99)*
Model 4	1 (ref)	0.95 (0.85-1.05)	0.94 (0.85-1.05)	0.98 (0.88-1.09)	0.78	0.99 (0.96-1.02)
Low-fat milk and (ice)cream						
n cases/n total	700/17,299	664/18,229	677/18,980	705/19,623		
Model 1	1 (ref)	0.91 (0.82-1.01)	0.91 (0.82-1.00)	0.92 (0.83-1.02)	0.12	0.99 (0.96-1.02)
Model 2	1 (ref)	0.94 (0.85-1.05)	0.96 (0.87-1.07)	0.99 (0.89-1.10)	0.91	1.01 (0.98-1.05)
Model 3	1 (ref)	0.94 (0.85-1.05)	0.95 (0.86-1.06)	0.97 (0.87-1.08)	0.66	1.01 (0.97-1.04)
Model 4	1 (ref)	0.92 (0.83-1.02)	0.93 (0.84-1.04)	0.93 (0.84-1.04)	0.24	0.99 (0.96-1.03)
Low-fat cheese						
n cases/n total	664/18,104	679/18,693	665/18,643	738/18,692		
Model 1	1 (ref)	1.04 (0.94-1.16)	0.99 (0.89-1.10)	1.03 (0.93-1.14)	0.81	0.99 (0.96-1.03)
Model 2	1 (ref)	1.07 (0.97-1.19)	1.02 (0.92-1.14)	1.03 (0.93-1.14)	0.77	0.99 (0.96-1.03)
Model 3	1 (ref)	1.07 (0.96-1.19)	1.02 (0.92-1.13)	1.03 (0.93-1.15)	0.75	0.99 (0.96-1.03)
Model 4	1 (ref)	1.06 (0.95-1.17)	1.00 (0.90-1.11)	1.02 (0.92-1.13)	0.91	0.99 (0.96-1.03)
Milk avoidance						
n cases/n total	687/18187	655/18460	704/18589	700/18895		
Model 1	1 (ref)	0.94 (0.85-1.04)	0.97 (0.88-1.08)	0.88 (0.79-0.97)	0.02*	0.96 (0.93-1.00)
Model 2	1 (ref)	0.95 (0.85-1.05)	0.98 (0.88-1.08)	0.90 (0.81-0.99)	0.06	0.97 (0.94-1.01)
Model 3	1 (ref)	0.95 (0.85-1.05)	0.98 (0.89-1.09)	0.90 (0.81-1.00)	0.07	0.98 (0.94-1.01)
Model 4	1 (ref)	0.93 (0.84-1.03)	0.98 (0.88-1.08)	0.90 (0.81-1.00)	0.10	0.98 (0.94-1.01)

Table 3. Associations of dairy food consumption pattern scores and prediabetes risk in the Lifelines study (*n* = 74,132) (continued).

	Relative risk (95% CI) across intake range categories[1]					Continuous[2]
	Q1	Q2	Q3	Q4	P_{trend}	RR (95%CI)
Low-fat yogurt						
n cases/n total	644/18,154	673/18,241	696/18,745	733/18,991		
Model 1	1 (ref)	1.05 (0.95-1.17)	1.07 (0.96-1.18)	1.07 (0.97-1.19)	0.20	1.03 (1.00-1.08)
Model 2	1 (ref)	1.05 (0.94-1.16)	1.08 (0.97-1.20)	1.08 (0.98-1.20)	0.11	1.04 (1.00-1.08)*
Model 3	1 (ref)	1.05 (0.95-1.17)	1.09 (0.98-1.21)	1.11 (1.00-1.23)	0.05*	1.05 (1.01-1.09)*
Model 4	1 (ref)	1.05 (0.95-1.17)	1.08 (0.98-1.20)	1.10 (0.99-1.22)	0.07	1.05 (1.01-1.09)*
Ice cream						
n cases/n total	661/18,702	664/18,309	714/18,545	707/18,575		
Model 1	1 (ref)	1.00 (0.90-1.11)	1.04 (0.93-1.15)	1.04 (0.93-1.15)	0.44	1.01 (0.97-1.05)
Model 2	1 (ref)	1.01 (0.91-1.12)	1.06 (0.95-1.17)	1.07 (0.96-1.18)	0.17	1.02 (0.98-1.07)
Model 3	1 (ref)	1.02 (0.92-1.13)	1.07 (0.96-1.18)	1.08 (0.97-1.20)	0.12	1.03 (0.99-1.07)
Model 4	1 (ref)	1.00 (0.90-1.11)	1.04 (0.94-1.15)	1.05 (0.95-1.16)	0.30	1.01 (0.97-1.06)

[1] Relative risks (95CIs) were estimated across four categories split by quartile values (Q1 to Q4) or non-consumers + tertile or median categories with the lowest category as the reference, adjusted for covariates as follows: Model 1 included age, sex, energy intake and follow-up duration. Model 2 additionally adjusted for educational level, alcohol use, smoking behaviour, physical activity level and family history of diabetes. Model 3 additionally adjusted for food groups associated with type 2 diabetes, including fruit, vegetables, bread, legumes, nuts, red and processed meat, coffee, tea, and sugar-sweetened beverages. Model 4 additionally adjusted for TAGs, LDL cholesterol, waist circumference, and hypertension. Linear trend across intake range categories was assessed by including median values of each category as a continuous variable in the model.

[2] Relative risks per 1 SD in PCA score were estimated.

[3] Median intake per intake range category in servings/d. * P=0.01 to 0.05, ** P<0.01. Abbreviations: CI, Confidence Interval; Q, Quartile.

Table 4. Associations of dairy food consumption pattern scores and prediabetes risk in the Fenland study (*n* = 6,639).

	Relative risk (95% CI) across intake range categories[1]					Continuous[2]
	Q1	Q2	Q3	Q4	P_{trend}	RR (95%CI)
High-fat dairy pattern						
n cases/n total	73/1668	79/1667	68/1659	70/1644		
Model 1[1]	1 (ref)	1.15 (0.84-1.57)	1.03 (0.74-1.42)	1.06 (0.76-1.46)	0.95	1.05 (0.96-1.15)
Model 2	1 (ref)	1.19 (0.87-1.63)	1.08 (0.78-1.49)	1.12 (0.81-1.55)	0.70	1.07 (0.97-1.17)
Model 3	1 (ref)	1.18 (0.86-1.61)	1.06 (0.77-1.47)	1.09 (0.79-1.51)	0.81	1.05 (0.96-1.15)
Model 4	1 (ref)	1.20 (0.88-1.63)	1.07 (0.77-1.49)	1.13 (0.82-1.56)	0.66	1.07 (0.98-1.17)
Dairy diversity pattern						
n cases/n total	78/1638	79/1657	80/1674	53/1669		
Model 1	1 (ref)	1.05 (0.77-1.42)	1.11 (0.81-1.52)	0.80 (0.54-1.17)	0.34	0.96 (0.86-1.06)
Model 2	1 (ref)	1.11 (0.81-1.52)	1.24 (0.90-1.70)	0.96 (0.65-1.43)	0.90	1.02 (0.91-1.13)
Model 3	1 (ref)	1.12 (0.82-1.54)	1.26 (0.91-1.74)	1.00 (0.67-1.49)	0.77	1.03 (0.92-1.15)
Model 4	1 (ref)	1.12 (0.82-1.53)	1.25 (0.90-1.72)	0.99 (0.66-1.48)	0.79	1.02 (0.91-1.14)
Low-fat avoidance pattern						
n cases/n total	77/1644	74/1660	68/1679	71/1655		
Model 1	1 (ref)	0.95 (0.69-1.29)	0.88 (0.64-1.21)	0.96 (0.70-1.33)	0.70	0.99 (0.89-1.10)
Model 2	1 (ref)	0.94 (0.69-1.28)	0.89 (0.65-1.23)	0.96 (0.70-1.33)	0.74	0.99 (0.89-1.10)
Model 3	1 (ref)	0.91 (0.67-1.25)	0.85 (0.61-1.17)	0.91 (0.65-1.26)	0.48	0.96 (0.86-1.07)
Model 4	1 (ref)	0.92 (0.67-1.26)	0.87 (0.63-1.20)	0.90 (0.65-1.26)	0.48	0.96 (0.86-1.07)
Ice cream pattern						
n cases/n total	63/1672	67/1676	75/1669	85/1621		
Model 1	1 (ref)	1.06 (0.76-1.48)	1.17 (0.84-1.63)	1.32 (0.95-1.82)	0.08	1.12 (1.00-1.26)
Model 2	1 (ref)	1.10 (0.78-1.54)	1.15 (0.83-1.60)	1.28 (0.92-1.77)	0.14	1.10 (0.97-1.24)
Model 3	1 (ref)	1.10 (0.78-1.54)	1.13 (0.81-1.58)	1.23 (0.89-1.71)	0.21	1.08 (0.96-1.22)
Model 4	1 (ref)	1.06 (0.76-1.49)	1.07 (0.77-1.50)	1.18 (0.85-1.63)	0.35	1.06 (0.94-1.19)

Table 4. Associations of dairy food consumption pattern scores and prediabetes risk in the Fenland study (*n* = 6,639) (continued).

	Relative risk (95% CI) across intake range categories[1]					Continuous[2]
	Q1	Q2	Q3	Q4	P_{trend}	RR (95%CI)
High-fat cheese pattern						
n cases/n total	76/1646	68/1687	72/1665	74/1640		
Model 1	1 (ref)	0.90 (0.65-1.24)	0.97 (0.71-1.33)	1.03 (0.75-1.41)	0.79	1.02 (0.90-1.16)
Model 2	1 (ref)	0.93 (0.68-1.28)	0.99 (0.72-1.36)	1.04 (0.76-1.43)	0.76	1.03 (0.91-1.17)
Model 3	1 (ref)	0.92 (0.67-1.26)	0.96 (0.70-1.31)	1.00 (0.73-1.37)	0.96	1.01 (0.90-1.15)
Model 4	1 (ref)	0.93 (0.67-1.28)	0.97 (0.71-1.33)	1.01 (0.74-1.39)	0.91	1.02 (0.90-1.15)
High-fat yogurt pattern						
n cases/n total	66/1661	73/1670	69/1674	82/1633		
Model 1	1 (ref)	1.12 (0.81-1.55)	1.04 (0.74-1.44)	1.35 (0.98-1.85)	0.11	1.13 (0.99-1.28)
Model 2	1 (ref)	1.08 (0.78-1.50)	1.00 (0.72-1.39)	1.30 (0.95-1.78)	0.16	1.12 (0.98-1.27)
Model 3	1 (ref)	1.09 (0.79-1.50)	1.01 (0.72-1.41)	1.31 (0.95-1.79)	0.15	1.12 (0.98-1.28)
Model 4	1 (ref)	1.06 (0.77-1.46)	1.00 (0.71-1.39)	1.29 (0.94-1.76)	0.17	1.11 (0.97-1.27)

[1] Relative risks (95CIs) were estimated across four categories split by quartile values (Q1 to Q4) or non-consumers + tertile or median categories with the lowest category as the reference, adjusted for covariates as follows: Model 1 included age, sex, study site and energy intake. Model 2 additionally adjusted for educational level, age at completion of education, ethnic origin, alcohol use, smoking behaviour, physical activity, and family history of diabetes. Model 3 additionally adjusted for dietary intakes (fruits, vegetables, whole grains, refined grains, potatoes, legumes, nuts, red and processed meat, fatty fish, coffee, tea, and sugar-sweetened beverages). Model 4 additionally adjusted for hypertension, dyslipidaemia, and waist circumference. Linear trend across intake range categories was assessed by including median values of each category as a continuous variable in the model.
[2] Relative risks per 1 SD in PCA score were estimated. * P=0.01 to 0.05, **P<0.01.
Abbreviations: CI, Confidence Interval; Q, Quartile.

Discussion

Four patterns of dairy intake observed in the Lifelines and the Fenland study were broadly comparable. In both cohorts, the highest variation in dairy intake was explained by a *'high-fat dairy'* pattern with high intakes of high-fat milk and low intakes of low-fat milk. Only two patterns were significantly associated with prediabetes risk. In the Lifelines study, the '*high-fat dairy*' pattern also strongly correlated with high-fat yogurt intake and was inversely associated with prediabetes risk in models adjusted for sociodemographic, health and dietary factors (RR per SD 0.96, 95%CI 0.93-0.99). This aligns with our findings of a non-significant inverse association for high-fat yogurt intake and prediabetes risk in ($RR_{serving/day}$ 0.80, 95%CI 0.64-1.01) **Chapter 5**. Furthermore, the '*low-fat yogurt*' pattern was positively associated with a risk of prediabetes in fully adjusted models (RR per 1 SD 1.05, 95%CI 1.01-1.09), while no association for higher low-fat yogurt intake and prediabetes risk was found ($RR_{serving/day}$ 1.02, 95%CI 0.90-1.16) in **Chapter 5**. The '*low fat yogurt*' pattern also correlated positively with other dairy types, except for lower ice cream intake. Individuals in the highest quartile of this pattern had slightly higher physical activity and diet quality. As discussed in **Chapter 5**, explanations for this positive association could be residual confounding by health behaviours, and the possibility that these individuals adjusted their lifestyle and dietary habits in accordance with guidelines, potentially due to heightened awareness of cardiometabolic risks. In a cross-sectional study of 1500 Irish adults, a '*reduced fats and yogurts*' was identified, more often female, with better dietary habits, but higher TAG and total cholesterol [2]. They also hypothesized that these individuals might have been following recommendations to consume reduced-fat dairy products due to increased health awareness. Furthermore, a higher carbohydrate intake associated with this cluster might have resulted in the higher TAG and observed cholesterol levels. We did not observe major differences in macronutrient intake across patterns. They also identified a '*whole milk*' cluster with more men, lower diet quality, and higher TNF-α, and a '*butter and cream*' cluster, both not associated with blood lipid profiles.

Our pattern analysis provided insights into the unique combinations of dairy types consumed in both populations, but they did not yield a higher predictive value for prediabetes risk compared to analyses with the intake of each dairy food. The absence of an association for most dairy patterns with prediabetes suggests that mixed intake of various dairy types is not necessarily detrimental for prediabetes, possibly due to the complex interplay of different dairy matrices.

References

1. Hu, F.B., *Dietary pattern analysis: a new direction in nutritional epidemiology.* Current opinion in lipidology, 2002. **13**(1): p. 3-9.
2. Feeney, E.L., et al., *Patterns of dairy food intake, body composition and markers of metabolic health in Ireland: results from the National Adult Nutrition Survey.* Nutrition & Diabetes, 2017. **7**(2): p. e243-e243.

Chapter 7

Systematic review and meta-analysis of dairy intake in relation to prediabetes risk and glycaemic outcomes

Manuscript based on this chapter:

Isabel A.L Slurink, Yakima D. Vogtschmidt, Bo Brummel, Nina Kupper, Tom Smeets, and Sabita S. Soedamah-Muthu. Dairy intake in relation to prediabetes and continuous glycaemic outcomes: a systematic review and dose-response meta-analysis of prospective cohort studies. Current Developments in Nutrition, 2024.

Abstract

Background
Modest inverse associations have been found between dairy intake, particularly yogurt, and type 2 diabetes risk. Investigating associations of dairy intake with early onset of type 2 diabetes offers opportunities for effective prevention of this condition.

Objective
This study aims to investigate the relationships between the intake of different dairy types, prediabetes risk, and continuous glycaemic outcomes.

Methods
Systematic literature searches across multiple databases were performed of studies published up to September 2023. Included were prospective cohort studies in healthy adults that examined the association between dairy intake and prediabetes risk according to diagnostic criteria, or continuous glycaemic markers. A dose-response random-effects meta-analysis was used to derive incremental relative risks (RR) for associations of total dairy, fermented dairy, milk, yogurt, cheese (all total, high-fat, and low-fat), cream, and ice cream with prediabetes risk adjusted for sociodemographic, health and cardiometabolic risk factors, and dietary characteristics.

Results
The meta-analyses encompassed 6,653 prediabetes cases among 95,844 individuals (age range 45.5-65.5 years) including 6 articles describing 9 cohorts. A quadratic inverse association was observed for total dairy intake and prediabetes risk, with the lowest risk at 3.4 servings/day (RR 0.75, 95%CI 0.60-0.93, I^2=18%). Similarly, total, and high-fat cheese exhibited nonlinear inverse associations with prediabetes risk, showing the lowest risk at 2.1 servings/day (0.86, 0.78-0.94, I^2=0%, and 0.90, 0.81-0.99, I^2=12%), but a higher risk at intakes exceeding 4 servings/day. Ice cream intake was linearly associated with prediabetes risk (0.85, 0.73-0.99, I^2=0% at the highest median intake of 0.23 servings/day). Other dairy types showed no statistically significant associations. The systematic review on dairy intake and glycaemic outcomes showed considerable variabilities in design and results.

Discussion
The findings suggest an inverse association between moderate dairy and cheese intake in preventing prediabetes. The potential for reverse causation and residual confounding highlights the need for studies with comprehensive repeated measurements.

Introduction

Prediabetes is a high-risk stage for developing type 2 diabetes (T2D). People in this risk stage already display some insulin resistance and declined pancreatic beta-cell function, resulting in impaired fasting or postprandial glycaemia. The prevalence of prediabetes is increasing at an alarming rate due to the aging of populations, economic developments, and unhealthier lifestyles [1]. Thus, effective preventive strategies for prediabetes are crucially needed. Many dietary guidelines worldwide recommend consuming 2-3 daily servings of dairy, based on systematic reviews showing evidence of a protective association between low-fat dairy and yogurt intake and T2D [2, 3].

The role of different dairy types in relation to prediabetes risk was investigated in detail in prospective cohorts in various Western world countries [4-10]. By focusing on individuals without prediabetes at baseline, these cohort studies offer insights into the risk factors associated with the early onset of T2D. This approach also eliminates a potential source of heterogeneity, as associations of risk factors with T2D may vary depending on the level of glycaemic disturbances at baseline [9]. Findings from these reports, however, show conflicting results [4-9]. Discrepancies in the direction and strength of these associations across different cohorts may be attributed to variations in countries, study population characteristics, variety in dairy products, and corresponding intake ranges.

To comprehensively investigate the association between dairy consumption and glycaemic control, studies incorporating continuous glycaemic markers, including those for insulin sensitivity and resistance, are crucial. These studies may provide insights into distinct mechanistic aspects related to maintenance of normal glycaemic control before the onset of disease. Moreover, continuous glycaemic markers may provide a more sensitive assessment compared to binary outcomes such as prediabetes. However, interpreting related effect estimates across studies becomes challenging due to variations in baseline glycaemic status. Additionally, as compared to risk estimates, effect estimates related to continuous glycaemic markers might be more difficult to translate into actionable public health implications.

To draw robust conclusions about how the type and dosage of dairy products consumed relate to incident prediabetes and glycaemic markers, a systematic review of the literature and meta-analysis of prospective cohort studies is needed. These types of extensive meta-analyses are considered at the top of the hierarchy of evidence [11]. Therefore, these results may further refine current scientific-based food-based dietary guidelines [12]. This study aims to conduct a systematic review and meta-analysis of prospective observational associations between intake of total and different types of dairy products, with different fat contents, and incident prediabetes and glycaemic markers in healthy adult populations.

Methods

The study protocol for this review can be found at https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=431251. This review was performed following the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) and Meta-analysis of Observational Studies in Epidemiology (MOOSE) guidelines [13].

Search strategy

Articles were retrieved from electronic databases including PubMed, Scopus, Web of Science, and the Cochrane Library. The search was confined to studies published from the initiation of research in the field up to April 18, 2024. Searches were performed using key terms in the title/abstract of published studies and with Medical Subject Heading (Mesh) terms, where possible. Authors of relevant abstracts were contacted for potential inclusion of unpublished data. Grey literature was examined by inspecting the first 200 items of a Google Scholar search. The full search strategy can be found in the **Supplemental table 1**.

Study selection

Articles were imported in Endnote and duplicates were removed based on article references. In a step-by-step process, two authors (IS and YDV) independently performed the title screening, abstract screening, and full-text screening according to predefined eligibility criteria in Rayyan [14]. Inclusion criteria were observational studies as design, involving adults (> 18 years) participants with normoglycemia as the study population, dairy food consumption as the main exposure of interest, and prediabetes or glycaemic markers as the main study outcomes. Study designs may include prospective cohort studies, nested case-control studies, case-cohort studies, and observational follow-up studies of RCTs. The articles had to be original research and written in English. For prediabetes (i.e., impaired fasting glucose or impaired glucose tolerance) the following recognized diagnostic criteria were used: fasting plasma glucose (FPG) between 110 and 125 mg/dl [15] or between 100 mg/dl and 125 mg/dl [16]; or haemoglobin A1c (HbA1c) levels between 6.0 and 6.4% [17] or 5.7 and 6.4% [16]; or 2-hour plasma glucose (2hPG) based on an oral glucose tolerance test (OGTT) of ≥140 and <200 mg/dl [15, 16]. Other glycaemic outcomes included fasting (or random) plasma glucose, 2hPG, HbA1c, fasting serum insulin (SI), insulin resistance index (HOMA-IR), insulin sensitivity index (Matsuda index), Stumvoll metabolic clearance rate, Stumvoll insulin sensitivity index, oral glucose insulin sensitivity index (OGIS), Gutt index, QUICKI, glucose-to-insulin ratio, other measures or indices of glucose or insulin sensitivity. We excluded studies conducted in animals, children, pregnant or lactating women, and ill populations (e.g. in patients with diabetes, cardiovascular disease, or cancer). Any disagreement was resolved until consensus was reached (IS and YDV). The reference lists of eligible articles and review articles were checked for additional eligible studies. Of the 159 fully reviewed articles, 18 met the inclusion criteria for a systematic review, of which six met the criteria for meta-analysis (**Figure 1**).

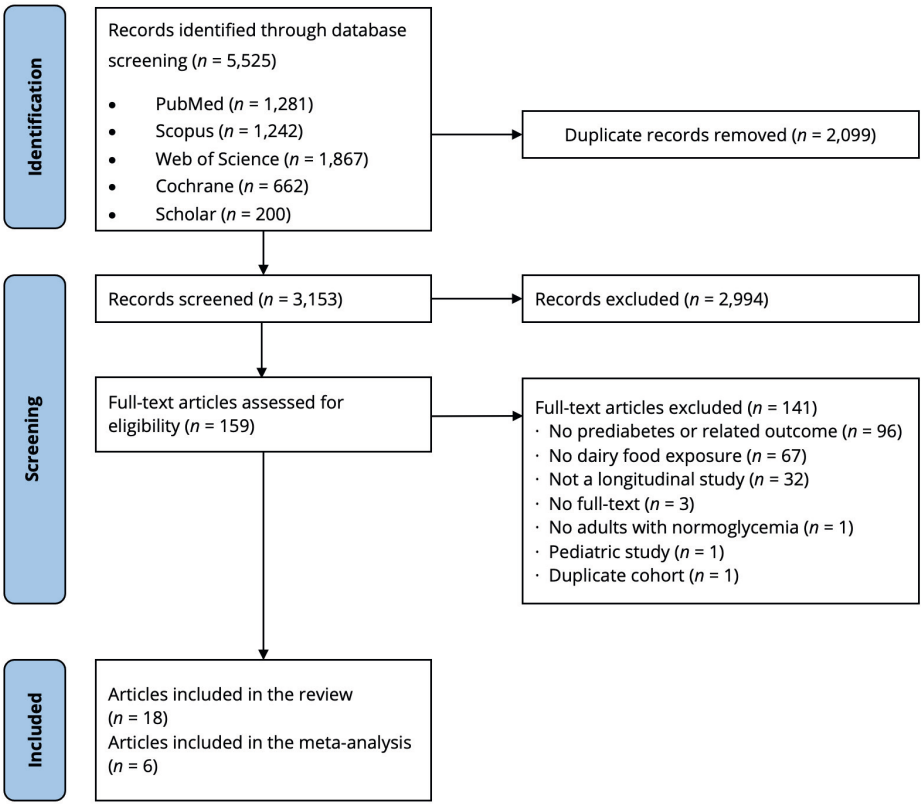


Figure 1. PRISMA flow-chart for the systematic review detailing the database searches, the number of abstracts screened, and the full-text retrieved.

Data extraction

Three authors (IS, YDV and BB) independently extracted data from the full text of eligible articles, according to a predefined protocol. The following data was extracted: bibliographic information; author(s); publication year; journal; title of the article; country; cohort study name; sample size; follow-up; participant characteristics; median or range of intake; number of subjects and prediabetes cases or mean outcome values; and confounder adjusted relative risks (RRs), odds ratios (ORs), hazard ratios (HRs) or beta coefficients (βs) and their corresponding 95%CI or SEs. Effect estimates derived from multiple adjusted models in different studies, where similar confounders were considered, were pooled, allowing for insights into the importance of confounding. For studies not reporting the median of each category, the mean of the lower and the upper limits were extracted. The meta-analyses for composite dairy types (total- and fermented dairy) were conducted in servings/day. Definitions of the different types of dairy in each included article can be found in **Supplemental table 2**. This approach was chosen as all studies uniformly defined a serving size by distinguishing liquid and solid dairy types (e.g., 200 g for liquid dairy foods

and 20 g for solid dairy foods). The intake of different dairy types, presented in servings per day or week, was converted into grams per day using either the reported conversion units in the article or country-specific standard units. For dairy types consisting of food items with different serving sizes, we averaged the serving sizes, which was the case for cheese in one article [9], and cream in two studies [8, 9]. Subsequently, we conducted the meta-analyses, standardizing the measurements to servings/day, enforcing equal water content of each type: 150 g for milk, yogurt, and ice cream; and 20 g for cheese. An exception was made for cream because of low intakes, where we operationalized the serving size as 15 g.

Risk of bias assessment

Three authors (IS, YDV, and BB) independently evaluated the risk of bias for the included studies using the Newcastle-Ottawa Scale (YDV and BB only if the included studies were performed by IS) [18]. This rating scale scores studies from 0 to 9 points on 3 domains: selection of the population, assessment of the outcome, and comparability of the groups. We considered a study to be of high quality if its total score was ≥7 points. Furthermore, the NutriGrade scoring system was used to rate the quality of the meta-evidence of each dairy type and prediabetes (i.e., the confidence in the estimate) [19]. This system included 8 items for meta-analysis of cohort studies, including 1) risk of bias, study quality, and study limitations, 2) precision, 3) heterogeneity, 4) directness, 5) publication bias, 6) funding bias, 7) effect size (based on a meta-analysis comparing highest versus lowest intake category of each study **(Supplemental table 3**), and 8) dose-response. We considered each meta-analysis to be of low (4 -<6 points) moderate (6 - <8 points), or high quality (≥8).

Data synthesis and analysis

Meta-analyses were performed when ≥3 cohorts per dairy type and outcome were available. While this criterion was met for all dairy types regarding prediabetes, it was not fulfilled for the glycaemic markers. Analyses were performed using the R packages dosresmeta, metafor and rms in R version 4.1.2 (R Foundation for Statistical Computing, Vienna, Austria) [20-22].

A two-stage linear random-effects meta-analysis was performed to obtain a single relative risk (RR) of each study expressed per serving size. Forest plots were created displaying the effect size of each study, its precision, and its weight to the summary estimate (**Supplemental figures 1-17**). Contour-enhanced funnel plots were used to investigate potential publication bias and small-study effects were evaluated using the Egger's test [23]. Furthermore, small-study effects were ascertained using a DOI plot to visualize asymmetry and the Luis Furuya–Kanamori index (LFK) index to quantify asymmetry of study effects [24]. The LFK indexes indicated asymmetry as follows: <±1 no asymmetry, ±1-2 minor asymmetry, and >±2 major asymmetry.

Two-stage dose-response random-effects meta-analyses were used to derive incremental dose-response RRs [25]. Potential nonlinear associations were examined using quadratic

and restricted cubic spline models. Likelihood ratio tests and information criteria (Akaike Information Criterion, AIC; and the Bayesian Information Criterion, BIC) were used to determine the most appropriate model fit and knot points. Associations were visualized using spaghetti plots. In spaghetti plots, the pooled RR and 95%CI at each quantity of intake is plotted, as well as cohort-specific RRs with study-specific weights. To assess heterogeneity between studies, the heterogeneity statistic (I^2) was calculated with the Higgins and Thompson method [26]. The Cochrane's Q test was conducted to evaluate if variation in effect estimates is likely due to chance alone. Four models with similar confounder adjustments in each study were compared; model 1 included age, sex, and energy intake, and model 2 additionally adjusted for an indicator of socio-economic status (SES) such as educational level, smoking, alcohol use, physical activity level, family history of diabetes, model 3 additionally adjusted for food groups associated with T2D including fruit, vegetables, bread, legumes, nuts, red and processed meat, coffee, tea, and sugar-sweetened beverages (SSBs), and model 4 additionally adjusted for BMI, waist circumference, and cardiometabolic risk factors (e.g. dyslipidaemia, hypertension). The results of model 4 are presented in the main text.

We performed a sensitivity analysis, excluding one study at a time from the analyses. Furthermore, we repeated the meta-analyses using a fixed-effect model to evaluate the consistency of results assuming a single true effect size across all studies, showing no differences in effect estimates (**Supplemental table 4)**. A meta-regression was performed to explain heterogeneity. Potential moderators included follow-up duration, calendar year of dairy intake assessment, the prediabetes definition used, and the literature quality score. Moderation by age, sex, and BMI was not explored due to limited variation and a restricted number of cohorts. Moderation by geographical location and country-level SES was not feasible as all studies were conducted in Western, high-SES countries. Additionally, the dietary assessment method and quality score were not considered as moderators, as all studies uniformly utilized an FFQ and were graded with similar quality. Each moderator was sequentially added as a covariate in the meta-regression model. The model fit of each bivariate model was compared using a deviance test and the adjusted R^2, indicating the percentage of heterogeneity explained by the covariate.

Results

Dairy and prediabetes

An overview of the characteristics of six articles incorporating nine study populations in prospective cohort designs can be found in **Table 1**. A total of 6,653 prediabetes cases among 95,844 individuals were identified, with the mean age ranging from 45.5 to 65.5 years. The sample sizes ranged from 997 to 74,132, and the follow-up duration ranged from 4.1 to 20.9 years. All articles were based on Western samples; three in the Netherlands, two in the UK [8, 9], and one in Australia [6]. Five articles were published by our research group in collaboration with PIs from the respective cohorts [4-8], and only one additional article

was identified through the systematic review [9]. Based on population-level median intake quantities, total dairy intake ranged from 1.1 to 3.7 servings/day (**Table 2**). Among all dairy types, milk contributed most to the total dairy intake, with median intakes ranging from 0.6 to 1.5 serving/day (96.5 to 220 grams/day) for low-fat milk and 0.03 to 1.3 servings/day (5.2 to 200 grams/day) for high-fat milk. The quality assessment scores of the individual studies ranged from 7 to 9, indicating a high quality of the individual studies (**Supplemental table 5)**. However, the quality of the evidence based on the meta-analyses, as rated by NutriGrade, was low for most dairy types (**Supplemental table 6**). The quality of meta-analyses for total dairy and ice cream was graded as moderate, and for cream as very low. The low quality of the evidence from meta-analyses was primarily attributed to the small number of cohorts included (≤10), the absence of a meaningful effect size, and the non-significance of dose-response analyses.

Total, low-fat, and high-fat dairy and prediabetes

Total dairy intake (9 cohorts, 6 articles) was nonlinearly inversely associated with prediabetes in the most adjusted models ($P_{nonlinearity}$ <0.0001, I^2 = 18%) (**Table 2; Figure 2**). Consuming 1 serving/day of total dairy was associated with a 13% lower risk of prediabetes (RR: 0.87 95%CI 0.78-0.96), and the lowest risk (25%) at 3.4 servings/day (RR: 0.75, 95%CI 0.60-0.93). The association weakened at higher intakes.

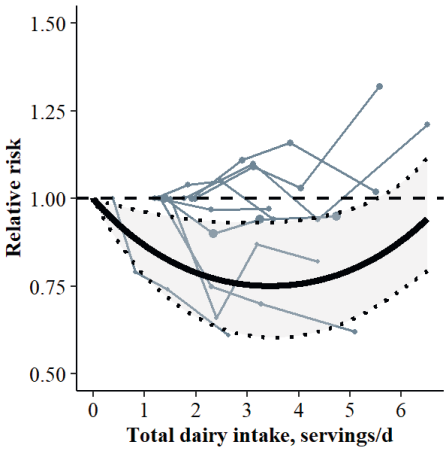


Figure 2. Spaghetti plot based on dose-response meta-analysis including 6 studies and 9 cohorts (6,653 cases among 95,844 participants) for the association between **total dairy** and prediabetes risk (lowest RR at 3.4 servings/day: 0.75, 95%CI 0.60-0.93 , I^2 = 18%). The solid black line represents the pooled RR at each quantity of intake. The light grey coloured area between the dotted black lines indicates the 95% confidence interval. The dashed grey line at RR = 1.00 represents the reference line. Each solid grey line represents a cohort with circles placed at the cohort-specific RRs at the corresponding intake level. The area of the circle is proportional to the study-specific weight. The associations were adjusted for age, sex, energy intake, educational level, smoking behaviour, physical activity, alcohol intake, family history of diabetes, intake of food groups, waist circumference or BMI, hypertension, and dyslipidaemia. Serving sizes were 200 g for liquid dairy foods and 20 g for solid dairy foods.

Table 1. Prospective cohort studies reporting associations between dairy product intake and prediabetes risk.

Author, year	Cohort (follow-up duration), baseline period, country	Female, %	Mean age, y	Mean BMI, kg/m2	Dietary assessment	Dairy types[1]	Prediabetes assessment	Cases, N	Adjustments
Hruby, 2017 [9]	Framingham Heart Study Offspring Cohort (10.5 y), 1991-1995, United States	57.0	52.6	26.4	126-item FFQ averaged across 4 exams. Validated against dietary records, highest for yogurt, r = .94–.97; lowest for cheese, r = .38–.57.	Dairy, milk (all total, high-fat, and low-fat), yogurt, cheese, and cream.	FPG 100-126 mg/dL, first incident measurement.	902/1867	Age, sex, energy intake, parental history of diabetes, smoking, dyslipidaemia, hypertension, coffee, nuts, fruits, vegetables, meats, alcohol, fish, glycaemic index, other dairy, baseline BMI and weight change over follow-up.
Slurink, 2022 [4]	Hoorn Study I (6 y), 1989-1992, The Netherlands	44.0	59.6	25.9	92-item FFQ. Validated against a dietary history, r = .72 for ranking of energy intake, .68 for animal protein, .73 for SFA, .69 for sodium, and .75 for calcium.	Dairy, fermented dairy, milk and milk products, plain milk, yogurt, cheese (all total, high-fat, and low-fat), cream and ice cream.	FPG 110-125 mg/dL or 2hPG 140-199 mg/dL or HbA1c 42-46 mmol/mol.	329/997	Age, sex, energy intake, follow-up duration, cohort, education, smoking, PA, alcohol intake, family history of diabetes, fruit, vegetables, tea, coffee, grains, meat, SSBs, BMI, LDL-c, and blood pressure.
Slurink, 2022 [5]	Hoorn Study II (6.7 y), 2006-2007, The Netherlands	55.0	53.0	25.6	104-item FFQ. Validated against actual energy intake in controlled feeding trials, r = .82 for energy intake, and validated against 3 24-h recalls, energy adjusted, deattenuated r = .60 for animal protein, .55 for SFA,.67 for calcium, .61 for cheese, and .75 for milk and milk products.	""	""	482/1265	""
Slurink, 2023 [6]	AusDiab (12 y), 1999-2000, Australia	56.7	49.0	26.1	74-item FFQ. Validation against 7-d weighted food records, energy adjusted r = .39 for protein, .64 for SFA, .59 for calcium, and .30 for sodium.	Dairy, milk, cheese (all total, high-fat, and low-fat), fermented dairy, yogurt, and ice cream.	FPG 110-125 mg/dL or 2hPG 140-199 mg/dL.	765/4891	Age, sex, energy intake, education, smoking, PA, alcohol intake, family history of diabetes, fruit, vegetables, grains, legumes, nuts, red and processed meat, fruit juice, WC, change in WC, LDL-c, and hypertension.
Slurink, 2022 [5]	Rotterdam Study, I (20.9 y), 1989-1993, The Netherlands	59.6	65.5	26.0	170-item FFQ. Validated against 15 24h-recalls; energy adjusted, deattenuated r = .66 for protein, .52 for SFA, .72 for calcium, .58 for sodium, and against 24h urine collection; .67 for protein.	Dairy, fermented dairy, milk and milk products, plain milk, yogurt, cheese (all total, high-fat, and low-fat), cream and ice cream.	FPG 110-125 mmol/L or PG 140-199 mg/dL, first incident measurement.	519/2617	Age, sex, energy intake, education, smoking, PA, alcohol intake, family history of diabetes (not in RS-III), fruit, vegetables, whole grains, legumes, nuts, tea, coffee, red meat and SSBs and longitudinal waist circumference.

Table 1. Prospective cohort studies reporting associations between dairy product intake and prediabetes risk (continued).

Author, year	Cohort (follow-up duration), baseline period, country	Female, %	Mean age, y	Mean BMI, kg/m2	Dietary assessment	Dairy types[1]	Prediabetes assessment	Cases, N	Adjustments
Slurink, 2022 [5]	Rotterdam Study, II (12.6 y), 2000, The Netherlands	55.1	63.6	27.0	""	""	""	290/ 1250	""
Slurink, 2022 [5]	Rotterdam Study, III (7 y), 2006, The Netherlands	59.9	56.8	27.2	389-item FFQ. Validated against 9-day dietary records and a 4-week dietary history, r = .74 for energy, .61 for animal protein intake, .73–.75 for SFA, .60 for calcium, .60 for milk and milk products and .61 for cheese.	""	""	330/ 2186	""
Slurink, 2023 [7]	Lifelines study (4.1 y), 2006-2013, The Netherlands	59.7	45.5	25.7	110-item FFQ (heart). The complete FFQ (heart + petals was compared with a conventional FFQ, r = .68 for energy, .61 for animal protein, .65 for SFA, .62 for calcium, .57 for cheese and .69 for dairy, and validated against urinary sodium (.40) and potassium (.37) excretions. 75% and 73% of participants, respectively, ranked in the same or adjacent quartile.	Dairy, fermented dairy, plain milk, yogurt, cheese (all total, high-fat, and low-fat), milk and milk products, cream, and ice cream.	FPG 110-125 mg/ dL or HbA1c 42-46 mmol/mol.	2746/ 74132	Age, sex, energy intake, follow-up duration, educational level, alcohol use, smoking, PA, family history of diabetes, fruit, vegetables, bread, legumes, nuts, red and processed meat, coffee, tea, SSBs, TAGs, LDL-c, waist circumference, and hypertension.
Slurink, 2024 [10]	Fenland study (6.7 y), 2005-2015, United Kingdom	51.9	48.7	26.4	130-item FFQ. Validated against 16-day weighted records, r = .52 for energy, .43. for protein, .56 for and SFA, .50 for calcium, and against 7-day food diaries, .56 for milk, .57 for yogurt and .33 for cheese.	Dairy, fermented dairy, plain milk, yogurt, cheese (all total, high-fat, and low-fat), cream and ice cream.	FPG 110-125 mg/ dL, 2hPG 140-199 mg/ dL or HbA1c 42-46 mmol/mol.	290/ 6639	Age, sex, study site, energy intake, educational level, age at completion of education, ethnic origin, alcohol use, smoking, PA, family history, fruit, vegetables, whole grains, refined grains, potatoes, legumes, nuts, red and processed meat, fatty fish, coffee, tea, SSBs, hypertension, dyslipidaemia, and waist circumference.

[1] The dairy types were defined into uniform categories by the authors. The original definitions are found in Supplemental Table 2.
Abbreviations: 2hPG, 2-hour Postprandial Glucose; FFQ, Food Frequency Questionnaire; FPG, Fasting Plasma Glucose; HbA1c, Haemoglobin A1c; LDL-c, low-density lipoprotein cholesterol; RS, Rotterdam Study; SFA, saturated fat acids; SSBs, Sugar-Sweetened Beverages; TAG, Triacylglycerol.

Table 2. Characteristics and outcomes of separate two-stage random effects dose-response meta-analyses, per dairy exposure.

Exposure	No. cohorts (articles)	Total N	N cases	Mean follow-up (years)	Range median intake (servings/d)[1]	Model fit ($P_{nonlinearity}$)	RR (95%CI) at 1 serving/day[2]	Lowest or highest RR of quadratic fit (95%CI)[2]	Heterogeneity		
									I^2	Q test	p-value
Total dairy	9 (6)	95,844	6,653	9.6	1.1-3.7	Quadratic (p<0.0001)	0.87 (0.78-0.96)	0.75 (0.60-0.93) at 3.4 servings /d	18%	19.5	0.24
High-fat dairy	9 (6)	95,844	6,653	9.6	0.3-2.0	Linear	0.99 (0.96-1.02)		37%	12.7	0.12
Low-fat dairy	9 (6)	95,844	6,653	9.6	0.8-2.6	Quadratic (p<0.0001)	1.00 (0.95-1.05)	0.93 (0.66-1.32) at 5.2 servings/d	28%	22.1	0.14
Fermented dairy	8 (5)	93,975	5,751	9.5	0.7-2.4	Quadratic (p<0.0001)	0.93 (0.86-1.01)	0.91 (0.81-1.02) at 2 servings/d	0%	9.4	0.81
High-fat fermented dairy	7 (4)	89,089	4,986	9.1	0.8-1.7	Quadratic (p=0.001)	0.95 (0.88-1.02)	0.94 (0.86-1.03) at 1.7 servings/d	0%	7.9	0.79
Low-fat fermented dairy	7 (4)	89,089	4,986	9.1	0.2-0.7	Linear	0.98 (0.94-1.03)		0%	1.8	0.94
Total milk	9 (6)	95,844	6,653	9.6	0.9-2.4	Quadratic (p<0.0001)	0.98 (0.88-1.08)	0.96 (0.84-1.09) at 3.5 servings/d	37%	25.3	0.07
High-fat milk	9 (6)	95,844	6,653	9.6	0.03-1.3	Linear	0.97 (0.88-1.08)		59%	19.3	0.01
Low-fat milk	9 (6)	95,844	6,653	9.6	0.6-1.5	Quadratic (p<0.0001)	1.03 (0.95-1.12)	0.87 (0.73-1.04) at 4.1 servings/d	30%	22.8	0.12
Total yogurt	9 (6)	95,844	6,653	9.6	0.02-0.5	Linear	0.98 (0.90-1.07)		8%	8.7	0.37
High-fat yogurt	7 (4)	89,086	4,986	9.6	0.2-0.4	Linear	1.04 (0.87-1.25)		21%	7.6	0.27
Low-fat yogurt	7 (4)	89,086	4,986	9.6	0.09-0.4	Linear	0.97 (0.82-1.15)		0%	2.9	0.83
Total cheese	9 (6)	95,844	6,653	9.6	0.6-2.9	Quadratic (p<0.0001)	0.89 (0.84-0.95)	0.86 (0.78-0.94) at 2.1 servings/d	0%	12.6	0.70
High-fat cheese	8 (5)	93,977	5,751	9.5	0.4-2.4	Quadratic (p<0.0001)	0.92 (0.87-0.98)	0.90 (0.81-0.99) at 2.1 servings/d	12%	15.8	0.32
Low-fat cheese	8 (5)	93,977	5,751	9.5	0.2-1.0	Linear	1.05 (0.98-1.14)		48%	13.5	0.06
Cream	8 (5)	90,953	5,888	9,3	0.02-0.8	Linear	0.85 (0.69-1.05)		0%	6.2	0.51
Ice cream	8 (5)	96,239	6,562	9,5	0.02-0.06	Linear	0.50 (0.26-0.94)		0%	2.4	0.93

[1] For composite dairy types, serving sizes were 200 g for liquid dairy foods and 20 g for solid dairy foods. For individual dairy types, serving sizes were 150 g for milk, yogurt, and ice cream; 20 g for cheese, and 15 g for cream. [2] The relative risks (95%CIs) were adjusted for age, sex, energy intake, educational level, smoking behaviour, physical activity, alcohol intake, family history of diabetes, intake of food groups, waist circumference or BMI, hypertension, and dyslipidaemia.
Abbreviations: CI, Confidence interval; RR, relative risk.

Neither high-fat nor low-fat dairy intake showed a statistically significant association with prediabetes risk ($RR_{serving/day}$ 0.99, 95%CI 0.96-1.02, I^2 = 37% and ($P_{nonlinearity}$ <0.0001, RR at 5.2 servings/day: 0.93, 95%CI 0.66-1.32, I^2 = 28%, respectively) (**Supplemental figure 18**). There was no evidence for publication bias in the meta-analyses for total, high-fat and low-fat dairy, as indicated by the funnel and DOI plots (**Supplemental Figures 19-21**).

Fermented dairy and prediabetes

No significant associations with prediabetes were found for fermented dairy intake, irrespective of fat content (**Table 2; Supplemental figure 22**), and there was no significant heterogeneity or evidence for publication bias (**Supplemental figures 23-25**).

Milk and prediabetes

No associations were observed for the consumption of total milk, high-fat milk, and low-fat milk with prediabetes risk (**Table 2; Supplemental figure 26**). For total and low-fat milk, there was moderate but non-significant heterogeneity, with I^2 of 37% and 30%, respectively, and no evidence was found for publication bias (**Supplemental figures 27-28**). Sensitivity analyses showed that with the exclusion of the FHS-OC from the meta-analysis, low-fat milk was associated with a 7% higher risk at 1.5 servings/day (RR 1.07, 95%CI 1.01-1.14) (**Supplemental table 7**). For high-fat milk, significant heterogeneity was observed (I^2 = 59%, p = 0.01). Accounting for moderation by the prediabetes assessment method resulted in a slightly better model fit for total milk (R^2 change 0.001 to 0.18) and high-fat milk (R^2 change 0.03 to 0.33), indicating that part of these association was explained by the prediabetes assessment method used (**Supplemental table 8**). This was especially noticeable in associations for high-fat milk. Cohorts that defined prediabetes based on FPG (FHS-OC) or FPG and non-fasting plasma glucose (RS) showed an inverse or no association between high-fat milk and prediabetes risk (**Supplemental figure 8**). Conversely, cohorts defining prediabetes using FPG, 2hPG, and HbA1c (i.e., the Fenland study and HS-II) showed a positive association between high-fat milk and prediabetes risk. The assessment of publication bias in the association with high-fat milk yielded inconclusive results. The Rotterdam Study II (RS-II) was an outlier in the funnel plot, although the Egger's test was not significant (p=0.45), and the LFK index indicated minor asymmetry (**Supplemental figure 29)**.

Yogurt and prediabetes

Total, high-fat and low-fat yogurt were not associated with prediabetes risk in the most adjusted models (**Table 2; Supplemental figure 30**). In the minimally adjusted model, total yogurt was linearly associated with lower prediabetes risk ($RR_{serving/day}$ 0.90, 95%CI 0.81-0.99, I^2 = 8%), but this association attenuated in models including adjustments for sociodemographic and health factors (**Supplemental table 9**). Including the prediabetes assessment method as a moderator in the model improved the model fit for high-fat yogurt (R^2 change 0.02 to 0.29) (**Supplemental table 8**). Although none of the associations in the individual cohorts was statistically significant, a discernible pattern

emerged (**Supplemental figure 11**). In cohorts in which prediabetes was defined using FPG, 2hPG and HbA1c (i.e., the Fenland study and HS-II), high-fat yogurt intake was associated with higher prediabetes risk. In contrast, in cohorts in which prediabetes was defined, based on the FPG and non-fasting plasma glucose (RS), an inverse association was found between high-fat yogurt intake and prediabetes risk, albeit not significant. There was no evidence of publication bias in the meta-analyses of total and high-fat yogurt (**Supplemental figures 31-32**). Evidence for publication bias for low-fat yogurt was inconclusive based on the Funnel plot and DOI plot (**Supplemental figure 33**).

Cheese and prediabetes

A nonlinear inverse association was found between total and high-fat cheese intake and prediabetes risk (both $P_{nonlinearity}$<0.0001) in the most adjusted models (**Table 2; Figure 4**). The risk of prediabetes was the lowest at intakes of 2.1 servings/day of total (RR 0.86, 95%CI 0.78-0.94, I^2 = 0%) and high-fat cheese intake (RR 0.90, 95%CI 0.81-0.99, I^2 = 12%), and increased to a RR of higher than 1 at intakes of more than 4 servings/day. Some asymmetry was observed in the funnel plot and the LFK index indicated major asymmetry; studies reporting positive associations tended to have smaller standard errors, whilst studies reporting inverse associations had higher standard errors (**Supplemental figures 34-35**). In sensitivity analyses, excluding the HS-1 cohort attenuated the association between high-fat cheese and prediabetes risk (RR at 1.8 servings/day: 0.93, 95%CI 0.87-1.01) (**Supplemental table 7**), but not with total cheese. The prediabetes assessment method explained some additional variance for total cheese (R^2 change from 0.46 to 0.68) and high fat cheese (R^2 change from 0.21 to 0.37) (**Supplemental table 8**). However, differences in model fit were minor and no evident pattern was found to explain this finding.

For low-fat cheese intake, a linear positive association with prediabetes risk was observed in models adjusted for age, sex, energy intake, sociodemographic and health factors, and food group intake ($RR_{serving/day}$ 1.10, 95%CI 1.02-1.19, I^2 = 51%) (**Supplemental table 9**). In models additionally adjusted for BMI or waist circumference, hypertension and dyslipidaemia, this association attenuated ($RR_{serving/day}$ 1.05, 0.98-1.14, I^2 = 48%) (**Table 2; Supplemental figure 36**). There was no evidence for publication bias (**Supplemental figure 37)**. Meta-regression analyses showed that the year of dairy intake assessment in individual studies explained some additional variance (R^2 change from 0.17 to 0.50) (**Supplemental table 8**). The two oldest cohorts, the HS-I and RS-I, both started in 1989, reported the highest risk for prediabetes with higher low-fat cheese intake, while the more recent cohorts did not report a statistically significant association (RS-III: 2006-07, Fenland: 2005-15 and Lifelines: 2006-2013) (**Supplemental figure 15**).

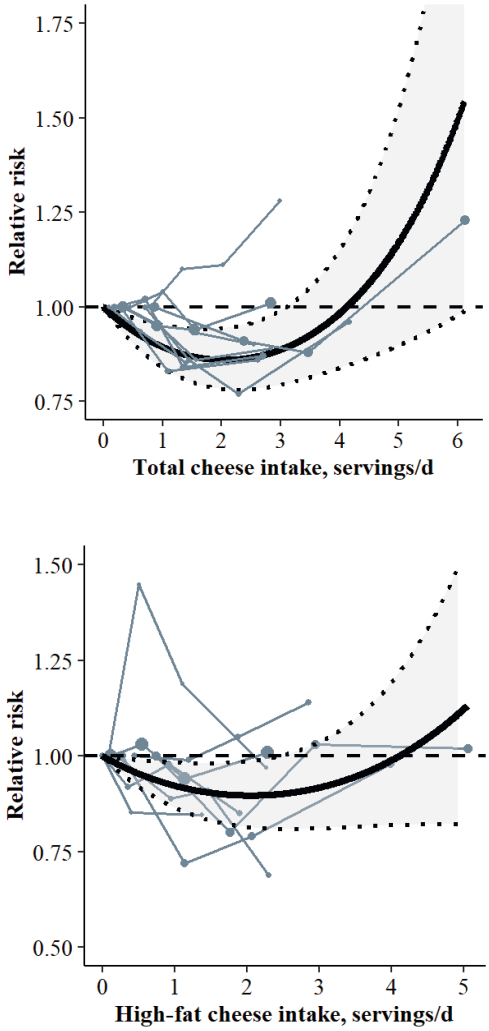


Figure 3. Spaghetti plot based on dose-response meta-analysis for the associations between **total cheese** (lowest RR at 2.1 servings/day: 0.86, 95%CI 0.78-0.94, I2 = 0%, including 6 studies and 9 cohorts, 6,653 cases among 95,844 participants)(left) and **high-fat cheese** (lowest RR at 2.1 servings/day: 0.90, 95%CI 0.81-0.99, I^2 = 12%, including 5 studies and 8 cohorts, 5,751 cases among 93,977 participants)(right) and prediabetes risk. The solid black line represents the pooled RR at each quantity of intake. The light grey coloured area between the dotted black lines indicates the 95% confidence interval. The dashed grey line at RR = 1.00 represents the reference line. Each solid grey line represents a cohort with circles placed at the cohort-specific RRs at the corresponding intake level. The area of the circle is proportional to the study-specific weight. The associations were adjusted for age, sex, energy intake, educational level, smoking behaviour, physical activity, alcohol intake, family history of diabetes, intake of food groups, waist circumference or BMI, hypertension, and dyslipidaemia. A serving size of cheese was 20 g.

Other dairy foods and prediabetes

Higher cream intake was not associated with prediabetes risk ($RR_{serving/day}$ 0.85, 95%CI 0.69-1.05, I^2 = 0%) (**Table 2; Supplemental figure 38**). Only the FHS-OC had high dairy intake levels and thereby contributed to 94% of the weight in the meta-analysis (**Supplemental figure 16**). Excluding this cohort from the model led to an additional risk reduction, although the association remained non-significant ($RR_{serving/day}$ 0.73, 95%CI 0.52-1.02) (**Supplemental table 7)**. The RS-II and RS-III had extremely low median intake levels with high standard errors resulting in major asymmetry of the funnel and Doi plot indicating potential publication bias (**Supplemental figure 39)**.

A linear association was found between higher ice cream intake and lower prediabetes risk ($RR_{serving/day}$ 0.50, 95%CI 0.26-0.94, I^2 = 0%) (**Table 2; Figure 4**). At the highest reported median intake of 0.23 servings/day, the RR was 0.85 (95%CI 0.73-0.99). However, this association was not statistically significant in models that were not adjusted for cardiometabolic risk markers ($RR_{serving/day}$ 0.57, 95% CI 0.30-1.08, I^2 = 0.0%) (**Supplemental table 9**). Minor asymmetry was identified in the funnel and Doi plot, with some smaller studies (HS-I and RS-I) reporting inverse associations, indicating some publication bias (**Supplemental figure 40)**.

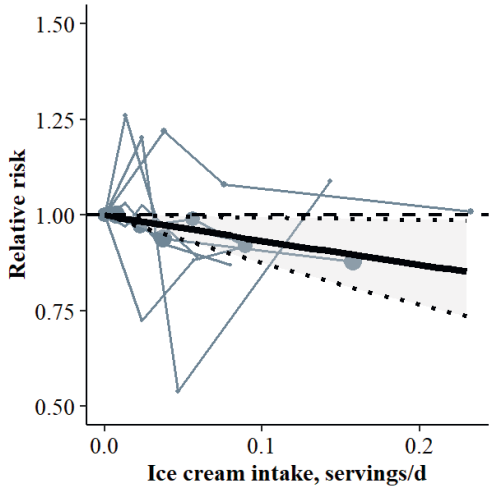


Figure 4. Spaghetti plot based on dose-response meta-analysis including 5 studies and 8 cohorts (6,562 cases among 96,239 participants) for the associations between **ice cream** intake and prediabetes risk ($RR_{serving/day}$ 0.50, 95%CI 0.26-0.94, I2 = 0%). The solid black line represents the pooled RR at each quantity of intake. The light grey coloured area between the dotted black lines indicates the 95% confidence interval. The dashed grey line at RR = 1.00 represents the reference line. Each solid grey line represents a cohort with circles placed at the cohort-specific RRs at the corresponding intake level. The area of the circle is proportional to the study-specific weight. The associations were adjusted for age, sex, energy intake, educational level, smoking behaviour, physical activity, alcohol intake, family history of diabetes, intake of food groups, waist circumference or BMI, hypertension, and dyslipidaemia. A serving size of ice cream was 150 g.

Dairy and glycaemic markers

The characteristics, results, and conclusions of 14 prospective cohort studies that reported associations between dairy products and glycaemic markers in a total of 38,441 participants are shown in **Table 3**. Furthermore, one meta-analysis including 182,041 participants from 18 studies conducted in the US, Denmark, Spain, Australia, and Finland was identified [27]. These studies were highly heterogeneous in analytical approaches, particularly in the use of baseline versus changes in dairy intake and in defining the outcome at follow-up versus change during follow-up. This variation prevented us from conducting meta-analyses on the relationship between dairy products and glycaemic markers.

Baseline dairy intake in relation to glycaemic markers at follow-up was reported in eight studies; two for FPG [28, 29], two for 2hPG [30, 31], two for HbA1c [27, 30], three for insulin resistance indices [5, 27, 30], and three for fasting insulin levels [27, 29, 32] with heterogeneous results. Inverse associations were observed in two studies, specifically for yogurt with FPG and fasting insulin levels [29], cheese and FPG, and fermented dairy and 2hPG [30]. Two studies reported positive associations between total dairy and HbA1c [27], and for total dairy, low-fat dairy, milk and milk products and HOMA2-B [30]. Other studies did not report any statistically significant associations.

Baseline dairy intake in relation to changes in glycaemic markers during follow-up was reported in five studies: four for changes in FPG [33-36], one for changes in 2hPG [34], one for estimates of insulin resistance [36], and one for fasting insulin levels [36]. Inverse associations were found for total dairy and FPG [35], milk, yogurt, and cottage cheese intake and FPG changes in men only [33]. All other studies reported no statistically significant associations.

Changes in dairy in relation to changes in glycaemic markers were reported in two studies: one for FPG and 2hPG [8], and two for HbA1c [8, 37]. Both studies assessed a wide range of dairy types, most of which were not associated with the outcome. One study found that changes in low-fat dairy and low-fat milk intake were positively associated with changes in FPG and 2hPG during follow-up [8], while another study found that changes in high-fat milk intake were positively associated with changes in HbA1c [37]. Both studies also reported associations of linear mixed models with repeated outcome measures, in which several additional or inconsistent associations were found. For example, repeated measures of high-fat milk and cheese were inversely associated with repeated measures of HbA1c [37]. Two other studies also estimated associations of repeated measures of dairy and glycaemic markers [5, 38]. Total dairy and cheese intake were associated with lower insulin levels in a study that involved repeated assessments of insulin [38]. In contrast, another study measuring HOMA-IR found no associations for total dairy and cheese [5]. However, this study showed that high-fat yogurt was associated with lower HOMA-IR, while low-fat dairy, total milk, and low-fat milk were associated with higher HOMA-IR.

Table 3. Prospective cohort studies reporting associations between dairy product intake and glycaemic outcomes.

Author, year	Cohort (follow-up), baseline and location	N, female %	Mean age, y	Mean BMI, kg/m2	Dietary assessment	Dairy types	Glycaemic markers	Adjustments	Main results	Conclusion
Feskens, 1995 [31]	Seven Countries Study including East Finland, West Finland, and Zutphen cohort (20 y), 1958-1964, Finland and The Netherlands.	338, 0%	75.5[2]	26.0[2]	Cross-check dietary history method by experienced dieticians and nutritionist.	Milk and milk products	2hPG at follow-up.	Cohort, age, past BMI, and past energy intake. Change estimates were additionally adjusted for past consumption.	Baseline milk and milk product intake or changes in intake were not associated with 2hPG at follow-up.	Fat intake (especially SFA) was associated with IGT and diabetes incidence, but not dairy intake. The regression coefficients were small.
Ma, 2006 [32]	Insulin Resistance Atherosclerosis Study (IRAS) (5 y), 1992-1994, United States.	1,087, 56.4%	54.8	28.5	114-item FFQ validated by comparison with 8 24-hour dietary recalls.	Total dairy	Insulin sensitivity at baseline and follow-up.	Age, sex, ethnicity, clinical centre, energy intake, family history of diabetes, smoking, alcohol intake, PA, protein, fat, fibre, calcium, magnesium, grains, fruit, vegetables, fish, meat, and BMI.	Dairy intake was not statistically significantly associated with insulin sensitivity (p = 0.41).	Magnesium and calcium intake, but not dairy intake, were independently associated with insulin sensitivity.
Snijder, 2008 [34]	Hoorn Study (6.4 y), 1989, The Netherlands.	1,124, 46.1%	60.0[2]	26.3	92-item FFQ. Validated against dietary history, controlled feeding trials, and 24-h recalls.	Total-, high-fat-, low-fat dairy, plain milk, dairy desserts, yogurt, and cheese.	Changes in FPG and 2hPG.	Age, sex, total energy intake, baseline outcome value, alcohol intake, smoking, and PA.	Dairy intake was not associated with changes in FPG and 2hPG. Separate dairy types were also not associated with glycaemic outcomes.	Dairy consumption was not associated with changes in weight, fat distribution, or other metabolic variables, nor with the incidence of the MetS.
Fumeron, 2011 [38]	DESIR (9 y), 2007-2010, France.	3,417, 50.1%	30-65	24.4[2]	23-item FFQ. Validated by comparison with dietary history method.	Total dairy Milk and other dairy products (except cheese) Cheese	Insulin at baseline and follow-up.	Age, sex, alcohol, smoking, PA, fat intake and BMI.	Total dairy (p=0.02) and cheese (p=0.05) intake were associated with lower insulin levels over the 9-year follow-up. Dairy except cheese intake was not associated with insulin levels (p=0.14).	Beneficial associations of dairy intake on metabolic syndrome and glycaemic disorders were found. Dairy product consumption could be protective against cardiovascular risk.

Table 3. Prospective cohort studies reporting associations between dairy product intake and glycaemic outcomes (continued).

Author, year	Cohort (follow-up), baseline and location	N, female %	Mean age, y	Mean BMI, kg/m2	Dietary assessment	Dairy types	Glycaemic markers	Adjustments	Main results	Conclusion
Struijk, 2012 [30]	Inter99 (5 y), 1999-2001, Denmark	5,953, 47.5%. Change analyses: 3,440-3,617	45.8	26.1	198-item FFQ. Validated against a 28-day dietary history with *r* = .55 for men and .60 for women, and for protein of .47 and .55, respectively.	Total-, high-fat-, low-fat dairy, milk and milk products, cheese, and fermented dairy.	FPG, 2hPG, HbA1c, HOMA2-IR and HOMA2-B at follow-up.	Age, sex, intervention group, educational level, diabetes family history, PA, smoking, alcohol intake, wholegrain cereal, meat, fish, coffee, tea, fruit, vegetables, energy intake, change in diet quality score and waist circumference.	Cheese intake was associated with 2hPG (β per serving/ day -0.048 mmol/L, 95%CI -0.095, 0.001). Fermented dairy intake was associated with FPG (-0.028 mmol/L, -0.048, -0.008) and HbA1c (-0.016 %, -0.03, -0.001). Intake of total dairy, low-fat dairy and milk and milk products were associated with a 0.8%-0.9% increase in HOMA2-B. Associations attenuated after adjustment for waist circumference.	The results indicate a potential modest beneficial role of cheese and fermented dairy on glucose regulation measures. However, all observed associations were small in magnitude, and this did not translate into a significant association with incident type 2 diabetes. The clinical relevance of the findings may therefore be limited.
Samara, 2013 [33]	STANISLAS study (5 y), 1994-1995, France	288 men, 300 women	42.8; 40.9	25.4; 23.5	3-day dietary records.	Milk, yogurt and cottage cheese, cheese	Changes in FPG.	Age, PA, alcohol, smoking, energy intake, education level, mean adequacy ratio index, and baseline outcome value.	For men, an inverse association was found between milk, yogurt, and cottage cheese intake and FPG changes (β -0.199 mmol/L per one tertile; SE 0.07, p=0.005). No associations were found for cheese. For women, no associations were found.	The relation of dairy products and 5-y changes in MetS components differed according to sex. Sex-related behaviours, knowledges, and attitudes toward diet and lifestyle could in part explain the inconsistent results.
Panahi, 2018 [29]	Quebec Family Study (6 y), 1978, Canada.	248, 53.7	40.5[2]	27.9[2]	3-day dietary records.	Total yogurt	FPG, insulin, C-peptide and AUC at follow-up	Age, diet quality, physical activity, and % body fat.	In men, consumers (≥1 servings/day) had lower fasting insulin, fasting C-peptide, glucose AUC, insulin AUC and C-peptide AUC (all p<0.05). In women, consumers had lower fasting insulin and glucose AUC. No interactions for yogurt intake and follow-up time were seen.	Yogurt consumption was associated with body composition and metabolic health benefits that are not entirely explained by a global effect on diet quality. There was no benefit of yogurt consumption over time on glycaemic markers.

Table 3. Prospective cohort studies reporting associations between dairy product intake and glycaemic outcomes (continued).

Author, year	Cohort (follow-up), baseline and location	N, female %	Mean age, y	Mean BMI, kg/m2	Dietary assessment	Dairy types	Glycaemic markers	Adjustments	Main results	Conclusion
Huang, 2019 [27]	Meta-analysis of 18 studies (1 to 26 years, 6 studies without follow-up), United States, Denmark, Spain, Australia, and Finland.	182,041, 48.8%-64%	19.9-70.9	24.3-30.1	FFQ in 14 studies. General questionnaire in 2 studies. 66-item interviewer administered in 1 study.	Total dairy	FPG HbA1c Log fasting insulin HOMA-IR HOMA-β at follow-up.	Sex, ethnicity, region, years of follow-up, and other baseline covariates if available (age, smoking status, physical activity, total energy intake, and alcohol intake).	Higher dairy intake was significantly associated with higher HbA1c (β 0.009 % per serving/day; SE 0.002; P < 0.001, I^2 = 9%). Associations with other glycaemic markers were not significant.	No causal association between dairy and glycaemic traits in MR analyses imply that the observational associations of dairy intake with glycaemic traits could be the result of confounding.
Trichia, 2020 [37]	EPIC-Norfolk study (3.7 y), 1993-1997, United Kingdom.	6,224, 56.2%	58.6	26.1	130-item FFQ. Validated against 7-d food diaries, *r* = 0.56 for milk, 0.57 for yogurt, 0.33 for cheese, and 0.54 for butter.	Total dairy, milk, yogurt, cheese (all total, high-fat, and low-fat), fermented dairy and ice cream.	Change in HbA1c and repeated measures.	Age, sex, educational level, age at completion of full-time education, marital status, occupation, follow-up time, PA, smoking, medication, energy intake, fruit, vegetables, potatoes, legumes, nuts, processed cereals, wholegrain cereals, poultry and eggs, red meat, processed meat, fish, sauces, margarine, sweet snacks, SSBs, ASBs, fruit juice, coffee, tea and alcoholic beverages, supplement use and BMI.	Increased full-fat milk intake was associated with increased HbA1c (β per serving/day 0.52 mmol/mol, 95% CI: 0.06, 0.97). Repeated measures of full-fat milk (β per serving/day -0.21 mmol/mol, 95%CI -0.41, -0.01), cheese (β per serving/day -0.31 mmol/mol, -0.59, -0.03), high-fat cheese (β per serving/day -0.63 mmol/mol 95%CI -1.01, -0.25) were associated with repeated measures of HbA1c, but not after multiple test correction.	Increased habitual full-fat milk consumption was associated with increased HbA1c concentrations, but reasons remain unclear. Various dairy types may differently influence cardiometabolic risk through adiposity and lipid pathways. The finding of an inverse association of fermented dairy products, yogurt (total and low-fat), or low-fat cheese with body weight provides a potential pathway for the previously described inverse association of fermented dairy consumption with type 2 diabetes.
Riseberg, 2022 [28]	Boston Puerto Rican Health Study (2 y), 2004-2007, United States.	1,126, 72.9%	56.0	31.1	FFQ validated against vitamins E and B12 and plasma carotenoids.	Milk and yogurt Cheese	FPG at follow-up.	Energy intake, sex, age, education, baseline outcome, smoking, alcohol intake, PA, psychological acculturation, fruit and vegetable intake score, omega-3 fatty acid intake, wholegrain intake, and medication.	No significant associations were observed for milk and yogurt intake (β per serving/day 0.91 mg/dL, 95%CI -1.27, 3.09) and cheese (-0.45 mg/dL, 95%CI -5.51; 4.62) and FPG.	Dairy was not related to cardiometabolic health.

Table 3. Prospective cohort studies reporting associations between dairy product intake and glycaemic outcomes (continued).

Author, year	Cohort (follow-up), baseline and location	N, female %	Mean age, y	Mean BMI, kg/m2	Dietary assessment	Dairy types	Glycaemic markers	Adjustments	Main results	Conclusion
Slurink, 2022 [5]	Rotterdam Study, I (20.9 y), 2000 Rotterdam Study, II (12.6 y), 2006 Rotterdam Study, III (7 y), 2006-2013, the Netherlands.	2,971, 59.6%; 1,413, 55.1%; 2,386, 59.9%	65.5; 63.6; 56.8	26.0; 27.0; 27.2	*See table 1.*	Total dairy, fermented dairy, milk and milk products, plain milk, yogurt, cheese (all total, high-fat, and low-fat), cream and ice cream.	HOMA-IR repeated measures.	Age, sex, energy intake, education, smoking, physical activity, alcohol intake, family history of diabetes (not in RS-III), fruit, vegetables, whole grains, legumes, nuts, tea, coffee, red meat and SSBs. Longitudinal waist circumference in sensitivity analyses.	High-fat yogurt was associated with lower HOMA-IR (β top vs. bottom quartile= −0.10, 95% CI −0.16, −0.05, p for trend = 0.0003, β = −0.08 per serving/day, 95% CI −0.13, −0.03). Low-fat dairy and low-fat milk were associated with higher HOMA-IR (β top vs. bottom quartile = 0.06, 0.03–0.10, p for trend = 0.0003 and 0.07, 0.03–0.11, p for trend = 0.001, and β per serving/day = 0.02, 0.01–0.03 and 0.02, 0.01–0.04, respectively). Total milk was associated with higher HOMA-IR (β top vs. bottom quartile = 0.05, 0.01–0.09, p for trend = 0.02). The positive associations of low-fat dairy and low-fat milk with HOMA-IR were observed in RS-II and RS-III but not in RS-I.	High-fat yogurt showed robust inverse associations with a prediabetes risk and longitudinal insulin resistance. Higher intake of high-fat milk was also associated with a lower prediabetes risk. Low-fat dairy, total milk, low-fat milk, and total and low-fat cheese were positively associated with the outcomes but inconsistently.
Yun, 2022 [35]	NHAPC (Nutrition and Health of Aging Population in China) (6 y), 2005, China	2,140, 59.4%	58.4	24.4	74-item FFQ. Validated against dietary records, *r* ranging from .68 to .72.	Total dairy, milk, non-milk dairy products (yogurt, ice cream and other).	Change in FPG.	Age, sex, region, residence, educational attainment, smoking, alcohol intake, PA, family history of chronic disease, lipid-lowering medication, energy intake, red meat, egg, fish, soy milk, vegetables, fruit, fibre, baseline outcome value.	Dairy intake was associated with changes in FPG (β per serving/day -0.14 mmol/L, 95%CI -0.21; -0.07). Milk and non-milk dairy products were not significantly associated with FPG.	The beneficial effects of dairy products on cardiovascular health might be mediated through specific sphingomyelins. Four unique sphingomyelins were identified reflecting total dairy consumption and showing favourable associations with changes in FPG.
Chatzidi-akou, 2023 [36]	Caerphilly prospective cohort study (5 y), United Kingdom	1,350, 0%	45-59		50-item FFQ. Validated against 7-day weighted dietary intake, *r* = .3 to .5 for all food items.	Total dairy Milk Cheese Cream	Changes in FPG, insulin and IR estimates.	Age, BMI, total energy intake and baseline outcome value.	No associations were observed between total dairy, milk, cheese, or cream consumption with changes in plasma glucose, insulin, and indices of insulin resistance.	A neutral effect of total dairy and types of dairy product intake on markers of glycaemic control was observed.

Table 3. Prospective cohort studies reporting associations between dairy product intake and glycaemic outcomes (continued).

Author, year	Cohort (follow-up), baseline and location	N, female %	Mean age, y	Mean BMI, kg/m2	Dietary assessment	Dairy types	Glycaemic markers	Adjustments	Main results	Conclusion
Slurink, 2024 [10]	Fenland study (6.7 y), 2005-2015, United Kingdom	7,410, 51.9%	48.7	26.4	*See table 1.*	Total dairy, fermented dairy, plain milk, yogurt, cheese (all total, high-fat, and low-fat), cream and ice cream	Changes in FPG, 2hPG and HbA1c and repeated measures.	Age, sex, study site, (change in) energy intake, educational level, age at completion of education, ethnic origin, (change in) alcohol intake, smoking, (change in) PA, family history of diabetes, (change in) fruit, vegetables, whole grains, refined grains, potatoes, legumes, nuts, red and processed meat, fatty fish, coffee, tea, SSBs, hypertension, dyslipidaemia and (change in) waist circumference and baseline outcome value.	Changes in low-fat dairy and low-fat milk were positively associated with changes in FPG ($\beta_{serving/day}$=0.02, 0.00-0.04, and 0.03, 0.01-0.05, respectively) and 2hPG (0.04, 0.00-0.08 and 0.06, 0.01-0.11, respectively). Changes in high-fat milk were inversely associated with changes in FPG ($\beta_{serving/day}$=-0.03, -0.05 to 0.00). Other associations of changes in dairy intake during follow-up and changes in FPG, OGT and HbA1c were not significant. Repeated measures of total dairy, milk, low-fat milk were positively associated with repeated measures of FPG ($\beta_{serving/day}$*time 0.002-0.003 mmol/L). Fermented dairy and high-fat fermented dairy were positively associated with 2hPG (both $\beta_{serving/day}$*time 0.01 mmol/L). Low-fat dairy, low-fat milk, and ice cream were positively associated with HbA1c, and high-fat dairy, high-fat milk were inversely associated with HbA1c.	The associations of dairy types and glycaemic markers differed between the parallel change analysis and the analysis of repeated measurements. Only the positive association between low-fat milk intake and FPG was found based on both analytical strategies. However, associations found for high-fat dairy, low-fat dairy, milk, high-fat milk, and low-fat milk with different glycaemic outcomes might prone to misclassification bias due to participants who consumed milk with unknown fat content at baseline or follow-up.

[1] The dairy types are shown according to definitions used by the original articles.
[2] Pooled across categories.
Abbreviations: 2hPG, 2-hour Postprandial Glucose; ASBs: Artificially Sweetened Beverages; AUC: Area Under the Curve; BMI: Body Mass Index; EPIC, European Prospective Investigation into Cancer and Nutrition; FFQ, Food Frequency Questionnaire; FPG, Fasting Plasma Glucose; HbA1c, Haemoglobin A1c; HOMA-2B, Homeostatic Model Assessment of Beta-cell Function; HOMA-β, Homeostatic Model Assessment of Beta-cell Function; HOMA-IR, Homeostatic Model Assessment of Insulin Resistance; HOMA2-IR, Homeostatic Model Assessment of Insulin Resistance; IR, Insulin Resistance; METS, Metabolic Equivalents; NR, Not Reported; PA, Physical Activity; SFA, saturated fatty acids; SSBs, Sugar-Sweetened Beverages.

The quality score ranged from 6 to 9 (**Supplemental table 5**), indicating moderate to high quality of individual studies. Most studies used FFQs that were validated against other dietary assessment methods for dairy components (**Table 3**). All studies had sufficient follow-up duration (≥3 years) except for one [28]. Nevertheless, the follow-up rate was less than 80% in six studies [8, 29, 34, 35, 37, 38], or no information was given on follow-up rates in two studies [30, 33]. Three studies excluded participants with a history of or prevalent diabetes [28, 33, 35]. Lastly, five studies lacked sufficient adjustment for confounding factors [29, 31, 34, 36, 38].

Discussion

The main findings of the dose-response meta-analysis of 6 articles based on 9 prospective cohorts indicated a nonlinear inverse association between total dairy intake and prediabetes risk. A 25% lower risk was identified at 3.4 servings/day, and the risk attenuated at higher intake levels, considering the limited data points available due to narrower intake ranges in underlying studies. Nevertheless, no associations were found for either high-fat or low-fat dairy intake. Furthermore, total, and high-fat cheese were nonlinearly inversely associated with prediabetes risk, with a 14% and 10% lower risk at 2.1 servings/day, respectively, but a positive risk at intakes of more than 4 servings/day. A linear inverse association was found for ice cream intake. No associations were found for fermented dairy, milk, yogurt, and cream intake, irrespective of fat content.

Furthermore, a systematic review was conducted of 14 prospective cohort studies exploring the relationship between dairy consumption and glycaemic outcomes. These studies presented a mix of inverse, positive, and null associations, reflected by the diversity in analytical strategies used. Thus, the existing body of evidence remains inconclusive regarding the relationship between subtypes of dairy and glycaemic outcomes.

The reduced risk of prediabetes observed at 3.4 servings of total dairy per day appeared not to be attributable to a specific fat content. This finding is plausibly driven by inverse associations between total and high-fat cheese intake and prediabetes. This association was in the same direction as in a preliminary meta-analysis of total cheese and prediabetes of 3 studies also included in our meta-analysis (RR highest versus lowest intake 0.93, 95%CI 0.78-1.12, I^2 = 66%) [39]. As we utilized serving size definitions weighed for the liquid content, the relative contribution of cheese is higher, as opposed to operationalization in grams. This inverse association of total and high-fat cheese might relate to cheese matrix effects. Specific components, including proteins, milk fat globule membrane (MFGM), medium- and branched chain fatty acids, calcium and vitamin K2 are abundantly present in high-fat cheese, which may exert possible beneficial effects on glucose homeostasis [40]. In line with our findings, the most comprehensive meta-analysis, based on 25 prospective studies with T2D as the outcome, showed that high compared to low total cheese intake was associated with lower risk (44,584 cases among

674,107 participants, RR: 0.93, 95%CI 0.88-0.98, I^2 = 45%), but there was no evidence for a dose-response association (*n* = 18, 35,449 cases among 394,508 participants, RR per 30 g/day: 1.00 95%CI 0.95-1.06, I^2 = 57%) [39]. Limited trial data to date, comparing cheese with different fat contents, suggested no effects on FPG, fasting insulin, and HOMA-IR [41, 42].

The inverse association we observed between total dairy and prediabetes risk is consistent in direction with the associations found for T2D, with the most complete meta-analysis showing a 3% lower risk per 200 g/day of total dairy (*n* = 22, RR: 0.97, 95%CI 0.95-1.00, I^2 = 63%) [2]. However, these associations with T2D are mainly driven by yogurt intake [39, 43, 44], whereas we did not observe associations between yogurt intake and prediabetes. Prior inverse associations have been attributed to beneficial effects of probiotics on weight and blood glucose regulation [40, 45], although the precise impact depends on the specific probiotic strains and their dosage [46]. The dietary assessment methods used in the included studies often lack detailed information of specific yogurt types and sugar content. As a result, the inconsistent findings observed may be attributed to variations in the yogurt types consumed. Furthermore, the meta-analysis with T2D as outcome found that the association was strongest in US populations (*n* = 5, RR per 50 g/day 0.91, 95%CI 0.86-0.96) compared to Asian (0.95, 0.79-1.14) and European (0.96, 0.92-1.01) populations [47]. Most of these US cohorts had limited yogurt intake with consumption being linked to healthy dietary patterns and behaviours [48, 49]. In contrast, yogurt consumption is more common in European populations, as well as for the Dutch cohorts included in the current meta-analysis. A wider range of yogurt intake levels, along with greater variability in associated participant characteristics (i.e., age, diet, health factors) could weaken the observed associations.

A linear association between ice cream intake and prediabetes emerged only after adjusting for confounding effects of cardiometabolic risk factors. However, caution is needed in interpretation, given that the highest median intake level was less than one serving per week (9 g/day), and the assessment is limited by seasonal variation. Similarly, a non-linear inverse association between ice cream intake and T2D risk was reported in a meta-analysis of 5 studies (19,730 cases among 258,571 participants) with the lowest risk observed at 10 g/day (RR 0.81, 0.78-0.85, I^2 = 86%) and no further decrease in risk found at higher intakes [47]. In three US prospective cohorts, the inverse association of ice cream intake and T2D attenuated when dietary data were no longer updated after participants reported a hypertension or hypercholesterolemia diagnosis [48]. Similarly, the potential for reverse causation, wherein dietary changes occur in response to cardiometabolic risk or a diagnosis, may explain our findings.

We observed no association between total and low-fat milk intake and prediabetes risk. For T2D, the most comprehensive meta-analysis of observational studies similarly shows no associations for total, high-fat or low-fat milk intake [2]. Also, no evidence for the causality

of these associations was found in Mendelian Randomization studies [50, 51]. A recent meta-analysis (*n* = 24, 13,211 cases among 1,297,951 participants) showed that milk intake was associated with lower T2D risk in non-white populations (RR per 245 g/day: 0.80, 95%CI 0.66-0.96), while in white populations, a modest positive association was found (RR per 245 g/day: 1.03, 95%CI 1.01-1.04) [52]. They suggested that variations in lactase persistence prevalence across populations could contribute to this heterogeneity and showed that this effect modification could stem from favourable alterations in gut microbiota and circulating metabolite profiles among individuals with lactase non-persistence. The populations included in our meta-analysis were predominantly of White origin.

Strengths and limitations

This is the first meta-analysis to investigate linear and non-linear dose-response relationships of various dairy types, categorized by high versus low-fat content, in association with prediabetes risk. We evaluated several factors that could explain heterogeneity between individual cohorts, including differences in confounder adjustments, baseline year, follow-up duration, and the prediabetes definition used. The results should be interpreted considering several limitations. First, the limited number of studies with low variation of intake for some dairy types may have led to overfitting, increasing the risk of spurious associations. Second, for meta-analyses of total and high-fat cheese, there were indications of publication bias. Third, the Lifelines study has a much larger sample size compared to the other cohorts and therefore received more weight in certain meta-analyses (e.g. 53.9% for fermented dairy). Fourth, given the observational nature of the included studies, the potential for residual confounding, reverse causation, and information bias cannot be dismissed. Fifth, the glycaemic markers and cut-off values to define prediabetes differed across the cohort, which was a source of heterogeneity for associations of high-fat milk and high-fat yogurt with prediabetes. Finally, all studies were conducted in Western populations limiting the generalizability of findings to other ethnic backgrounds, low- or middle-income countries, and varying dietary patterns.

Conclusion

The current evidence of population-based prospective cohort studies suggests that overall, dairy intake as measured at baseline, regardless of fat content, does not elevate the risk of prediabetes. Moderate beneficial associations were observed for total dairy intake, total cheese, and high-fat cheese. Milk, yogurt, and cream were not associated with prediabetes, irrespective of their fat content. The potential for reverse causation and residual confounding, especially considering our finding of an inverse linear association for ice cream intake, warrants the need for studies with comprehensive repeated measurements. Additionally, to inform more targeted preventive strategies and interventions, there is a need for randomized controlled trials validating potential underlying mechanisms and for exploring possible intra-individual variability in responses to dairy intake.

List of supplementary materials chapter 7

Supplemental table 1. Search strategy.

Supplemental table 2. Definitions of dairy types included in each individual study included in the meta-analysis.

Supplemental table 3. Characteristics and outcomes of separate two-stage random-effect meta-analyses comparing the highest versus lowest category of intake, per dairy exposure, for the scoring of the effect size (7) of NutriGrade.

Supplemental table 4. Characteristics and outcomes of separate two-stage fixed-effects dose-response meta-analyses, per dairy exposure.

Supplemental table 5. Quality assessment of cohort studies on dairy intake, prediabetes, and glycaemic outcomes.

Newcastle – Ottawa Quality Assessment Scale Cohort Studies

Supplemental table 6. NutriGrade grading of evidence from separate meta-analyses, per dairy exposure.

NutriGrade Scoring system for meta-evidence based on prospective cohort studies

Supplemental table 7. Sensitivity analyses of associations between dairy foods and prediabetes risk based on one-stage random-effects dose-response meta-analysis.

Supplemental table 8. Goodness-of fit tests for models with potential moderators of the association between dairy foods and prediabetes risk based on one-stage linear or quadratic dose-response meta-regression.

Supplemental table 9. Associations between dairy foods and prediabetes risk based on linear or quadratic two-stage random-effects dose-response meta-analysis for different confounder models.

Supplemental Figure 1-17. Forest plot for the associations between dairy intake and prediabetes risk and study variation based on two-stage linear meta-analysis.

Supplemental figure 18. Spaghetti plot based on dose-response meta-analysis for the associations between high-fat dairy, and low-fat dairy intake and prediabetes risk.

Supplemental figure 19. Contour-enhanced funnel plot and Doi plot for studies of the association between total dairy intake and prediabetes risk.

Supplemental figure 20. Contour-enhanced funnel plot and Doi plot for studies of the association between high-fat dairy intake and prediabetes risk.

Supplemental figure 21. Contour-enhanced funnel plot and Doi plot for studies of the association between low-fat dairy intake and prediabetes risk.

Supplemental figure 22. Spaghetti plot based on dose-response meta-analysis for the associations between **f**ermented dairy, high-fat fermented dairy, and low-fat fermented dairy intake and prediabetes risk.

Supplemental figure 23. Contour-enhanced funnel plot and Doi plot for studies of the association between fermented dairy intake and prediabetes risk.

Supplemental figure 24. Contour-enhanced funnel plot and Doi plot for studies of the association between high-fat fermented dairy intake and prediabetes risk.

Supplemental figure 25. Contour-enhanced funnel plot and Doi plot for studies of the association between low-fat fermented dairy intake and prediabetes risk.

Supplemental figure 26. Spaghetti plot based on dose-response meta-analysis for the associations between total milk, high-fat milk, and low-fat milk intake and prediabetes risk.

Supplemental figure 27. Contour-enhanced funnel plot and Doi plot for studies of the association between total milk intake and prediabetes risk.

Supplemental figure 28. Contour-enhanced funnel plot and Doi plot for studies of the association between low-fat milk intake and prediabetes risk.

Supplemental figure 29. Contour-enhanced funnel plot and Doi plot for studies of the association between high-fat milk intake and prediabetes risk.

Supplemental figure 30. Spaghetti plot based on dose-response meta-analysis for the associations between total yogurt, high-fat yogurt, and low-fat yogurt intake and prediabetes risk.

Supplemental figure 31. Contour-enhanced funnel plot and Doi plot for studies of the association between total yogurt intake and prediabetes risk.

Supplemental figure 32. Contour-enhanced funnel plot and Doi plot for studies of the association between high-fat yogurt intake and prediabetes risk.

Supplemental figure 33. Contour-enhanced funnel plot and Doi plot for studies of the association between low-fat yogurt intake and prediabetes risk.

Supplemental figure 34. Contour-enhanced funnel plot and Doi plot for studies of the association between total cheese intake and prediabetes risk.

Supplemental figure 35. Contour-enhanced funnel plot and Doi plot for studies of the association between high-fat cheese intake and prediabetes risk.

Supplemental figure 36. Spaghetti plot based on dose-response meta-analysis for the associations between low-fat cheese and prediabetes risk.

Supplemental figure 37. Contour-enhanced funnel plot and Doi plot for studies of the association between low-fat cheese intake and prediabetes risk.

Supplemental figure 38. Spaghetti plot based on dose-response meta-analysis for the associations between cream intake and prediabetes risk.

Supplemental figure 39. Contour-enhanced funnel plot and Doi plot for studies of the association between cream intake and prediabetes risk.

Supplemental figure 40. Contour-enhanced funnel plot and Doi plot for studies of the association between ice cream intake and prediabetes risk.

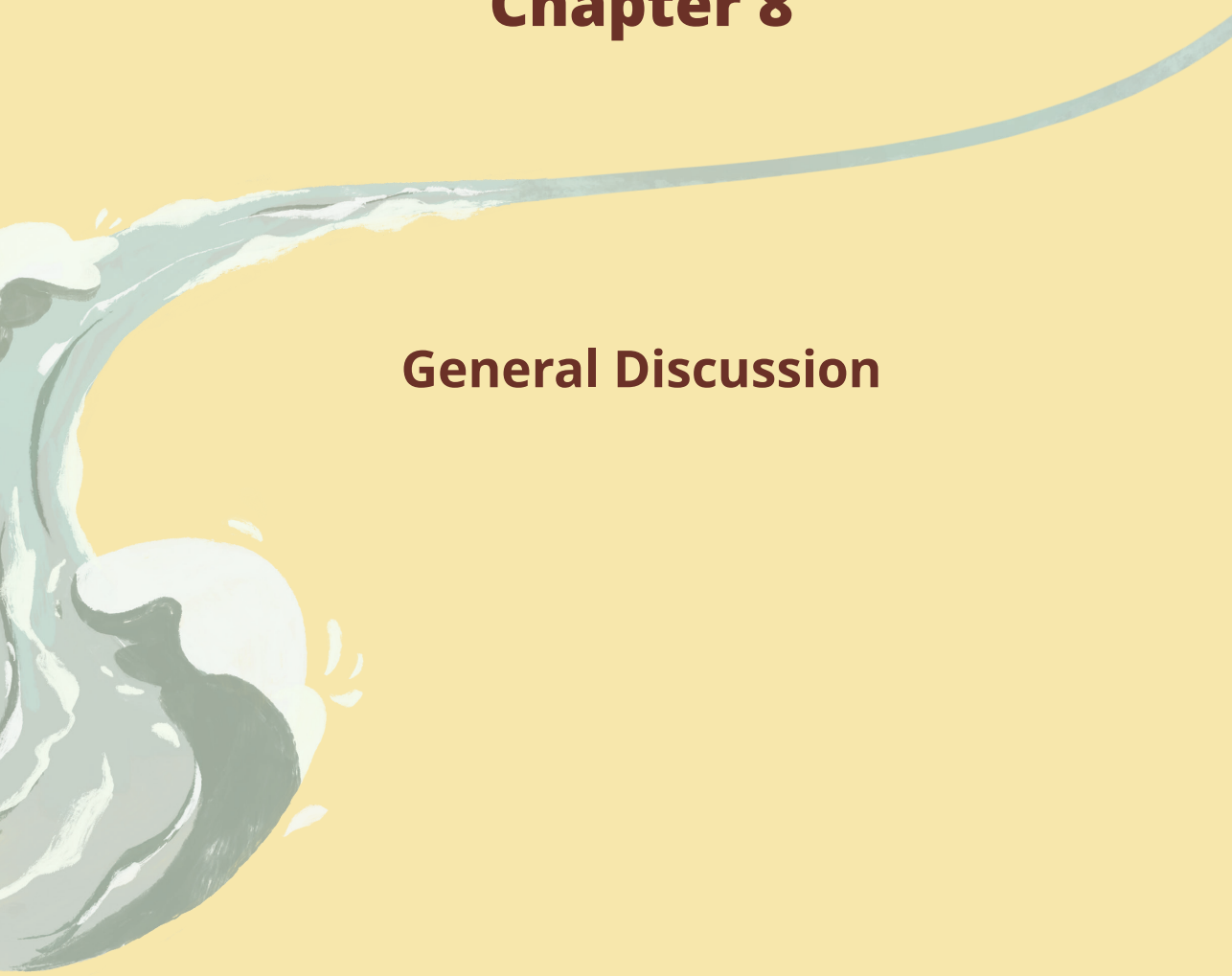
Scan this QR code to download the supplementary materials.

References

1. Rooney, M.R., et al., *Global Prevalence of Prediabetes.* Diabetes Care, 2023: p. dc222376.
2. Soedamah-Muthu, S.S. and J. de Goede, *Dairy Consumption and Cardiometabolic Diseases: Systematic Review and Updated Meta-Analyses of Prospective Cohort Studies*. Curr Nutr Rep, 2018. **7**(4): p. 171-182.
3. Comerford, K.B., et al., *Global Review of Dairy Recommendations in Food-Based Dietary Guidelines.* Front Nutr, 2021. **8**: p. 671999.
4. Slurink, I.A.L., et al., *Dairy product consumption and incident prediabetes in Dutch middle-aged adults: the Hoorn Studies prospective cohort.* Eur J Nutr, 2022. **61**(1): p. 183-196.
5. Slurink, I.A.L., et al., *Dairy Product Consumption in Relation to Incident Prediabetes and Longitudinal Insulin Resistance in the Rotterdam Study.* Nutrients, 2022. **14**(3).
6. Slurink, I.A., et al., *Dairy Product Consumption and Incident Prediabetes in the Australian Diabetes, Obesity, and Lifestyle Study With 12 Years of Follow-Up.* J Nutr, 2023. **153**(6): p. 1742-1752.
7. Slurink, I.A., et al., *Dairy consumption and incident prediabetes: prospective associations and network models in the large population-based Lifelines study.* Am J Clin Nutr, 2023. **118**(6): p. 1077-1090.
8. Slurink, I. and F. Imamura, *Dairy prediabetes Fenland.*
9. Hruby, A., et al., *Associations of Dairy Intake with Incident Prediabetes or Diabetes in Middle-Aged Adults Vary by Both Dairy Type and Glycemic Status*. J Nutr, 2017. **147**(9): p. 1764-1775.
10. Slurink, I., et al., *Dairy consumption and risk of prediabetes and type 2 diabetes in the Fenland study [Manuscript submitted for publication].*
11. Higgins, J.P. and S. Green, *Cochrane handbook for systematic reviews of interventions. The Cochrane Collaboration. London, UK, 2011.*
12. Dutch Health Council (Gezondheidsraad), *Dutch dietary guidelines 2015 (Richtlijnen goede voeding 2015). Publication nr. 2015/24. ISBN 978-94-6281-089-1. The Hague.* 2015.
13. Stroup, D.F., et al., *Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group.* JAMA, 2000. **283**(15): p. 2008-12.
14. Ouzzani, M., et al., *Rayyan-a web and mobile app for systematic reviews.* Syst Rev, 2016. **5**(1): p. 210.
15. World Health Organization (WHO), *Definition and diagnosis of diabetes mellitus and intermediate hyperglycaemia: report of a WHO/IDF consultation.* 2006.
16. American Diabetes Association Professional Practice, C., *2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes-2022.* Diabetes Care, 2022. **45**(Suppl 1): p. S17-S38.
17. International Expert, C., *International Expert Committee report on the role of the A1C assay in the diagnosis of diabetes.* Diabetes Care, 2009. **32**(7): p. 1327-34.
18. Wells, G.A., et al., *The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses*. 2000, Oxford.
19. Schwingshackl, L., et al., *Perspective: NutriGrade: A Scoring System to Assess and Judge the Meta-Evidence of Randomized Controlled Trials and Cohort Studies in Nutrition Research.* Adv Nutr, 2016. **7**(6): p. 994-1004.
20. Crippa, A. and N. Orsini, *Multivariate dose-response meta-analysis: the dosresmeta R package.* Journal of Statistical Software, 2016. **72**: p. 1-15.
21. Viechtbauer, W., *Conducting meta-analyses in R with the metafor package.* Journal of statistical software, 2010. **36**: p. 1-48.
22. Harrell Jr, F.E., *rms: Regression modeling strategies.* R package version, 2016. **5**(2): p. 1-263.
23. Peters, J.L., et al., *Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry.* J Clin Epidemiol, 2008. **61**(10): p. 991-6.
24. Furuya-Kanamori, L., J.J. Barendregt, and S.A.R. Doi, *A new improved graphical and quantitative method for detecting bias in meta-analysis.* Int J Evid Based Healthc, 2018. **16**(4): p. 195-203.
25. Greenland, S. and M.P. Longnecker, *Methods for trend estimation from summarized dose-response data, with applications to meta-analysis.* Am J Epidemiol, 1992. **135**(11): p. 1301-9.
26. Higgins, J.P. and S.G. Thompson, *Quantifying heterogeneity in a meta-analysis.* Stat Med, 2002. **21**(11): p. 1539-58.
27. Huang, L., et al., *Circulating Saturated Fatty Acids and Incident Type 2 Diabetes: A Systematic Review and Meta-Analysis.* Nutrients, 2019. **11**(5): p. 998.

28. Riseberg, E., et al., *Specific Dietary Protein Sources Are Associated with Cardiometabolic Risk Factors in the Boston Puerto Rican Health Study.* J Acad Nutr Diet, 2022. **122**(2): p. 298-308 e3.
29. Panahi, S., et al., *Yogurt consumption, body composition, and metabolic health in the Quebec Family Study.* Eur J Nutr, 2018. **57**(4): p. 1591-1603.
30. Struijk, E.A., et al., *Dairy product intake in relation to glucose regulation indices and risk of type 2 diabetes.* Nutr Metab Cardiovasc Dis, 2013. **23**(9): p. 822-8.
31. Feskens, E.J., et al., *Dietary factors determining diabetes and impaired glucose tolerance. A 20-year follow-up of the Finnish and Dutch cohorts of the Seven Countries Study.* Diabetes Care, 1995. **18**(8): p. 1104-12.
32. Ma, B., et al., *Dairy, magnesium, and calcium intake in relation to insulin sensitivity: approaches to modeling a dose-dependent association.* Am J Epidemiol, 2006. **164**(5): p. 449-58.
33. Samara, A., et al., *Dairy product consumption, calcium intakes, and metabolic syndrome-related factors over 5 years in the STANISLAS study.* Nutrition, 2013. **29**(3): p. 519-24.
34. Snijder, M.B., et al., *A prospective study of dairy consumption in relation to changes in metabolic risk factors: the Hoorn Study.* Obesity (Silver Spring), 2008. **16**(3): p. 706-9.
35. Yun, H., et al., *Lipidomic signatures of dairy consumption and associated changes in blood pressure and other cardiovascular risk factors among chinese adults.* Hypertension, 2022. **79**(8): p. 1617-1628.
36. Chatzidiakou, Y., et al., *Relationship between the consumption of dairy foods and markers of glycaemic control: evidence from the Caerphilly prospective cohort study.* Proceedings of the Nutrition Society, 2023. **82**(OCE1): p. E6.
37. Trichia, E., et al., *The associations of longitudinal changes in consumption of total and types of dairy products and markers of metabolic risk and adiposity: findings from the European Investigation into Cancer and Nutrition (EPIC)-Norfolk study, United Kingdom.* Am J Clin Nutr, 2020. **111**(5): p. 1018-1026.
38. Fumeron, F., 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): p. 813-7.
39. Zhang, M., et al., *Cheese consumption and multiple health outcomes: an umbrella review and updated meta-analysis of prospective studies.* Adv Nutr, 2023. **14**(5): p. 1170-1186.
40. Mozaffarian, D. and J.H.Y. Wu, *Flavonoids, Dairy Foods, and Cardiovascular and Metabolic Health: A Review of Emerging Biologic Pathways.* Circ Res, 2018. **122**(2): p. 369-384.
41. Nilsen, R., et al., *Effect of a high intake of cheese on cholesterol and metabolic syndrome: results of a randomized trial.* Food Nutr Res, 2015. **59**: p. 27651.
42. Raziani, F., et al., *High intake of regular-fat cheese compared with reduced-fat cheese does not affect LDL cholesterol or risk markers of the metabolic syndrome: a randomized controlled trial.* Am J Clin Nutr, 2016. **104**(4): p. 973-981.
43. Fan, M., et al., *Dietary Protein Consumption and the Risk of Type 2 Diabetes: ADose-Response Meta-Analysis of Prospective Studies.* Nutrients, 2019. **11**(11): p. 2783.
44. Zhang, K., P. Bai, and Z. Deng, *Dose-Dependent Effect of Intake of Fermented Dairy Foods on the Risk of Diabetes: Results From a Meta-analysis.* Can J Diabetes, 2022. **46**(3): p. 307-312.
45. Kim, D.H., et al., *Kefir alleviates obesity and hepatic steatosis in high-fat diet-fed mice by modulation of gut microbiota and mycobiota: targeted and untargeted community analysis with correlation of biomarkers.* J Nutr Biochem, 2017. **44**: p. 35-43.
46. Nikbakht, E., et al., *Effect of probiotics and synbiotics on blood glucose: a systematic review and meta-analysis of controlled trials.* Eur J Nutr, 2018. **57**(1): p. 95-106.
47. Gijsbers, L., et al., *Consumption of dairy foods and diabetes incidence: a dose-response meta-analysis of observational studies.* Am J Clin Nutr, 2016. **103**(4): p. 1111-24.
48. Margolis, K.L., et al., *A Diet High in Low-Fat Dairy Products Lowers Diabetes Risk in Postmenopausal Women.* The Journal of Nutrition, 2011. **141**(11): p. 1969-1974.
49. Chen, M., et al., *Dairy consumption and risk of type 2 diabetes: 3 cohorts of US adults and an updated meta-analysis.* BMC Med, 2014. **12**: p. 215.
50. Jensen, C.F., M. Timofeeva, and G. Berg-Beckhoff, *Milk consumption and the risk of type 2 diabetes: A systematic review of Mendelian randomization studies.* Nutr Metab Cardiovasc Dis, 2023. **33**(7): p. 1316-1322.
51. Mendelian Randomization of Dairy Consumption Working, G. and C. consortium, *Dairy Intake and Body Composition and Cardiometabolic Traits among Adults: Mendelian Randomization Analysis of 182041 Individuals from 18 Studies.* Clin Chem, 2019. **65**(6): p. 751-760.
52. Luo, K., et al., *Variant of the lactase LCT gene explains association between milk intake and incident type 2 diabetes.* Nature Metabolism, 2024: p. 1-18.

Chapter 8



General Discussion

The aim of the research described in this thesis was to investigate intake of total dairy and a variety of high-fat and low-fat dairy types, including milk, yogurt, and cheese, cream, and ice cream, in relation to glycaemic outcomes and the incidence of prediabetes. To achieve this aim, an in-depth nutritional epidemiological study was undertaken applying a similar data analytical approach in five large prospective cohort studies in the Netherlands, Australia, and the UK. The current chapter provides an overview and interpretation of the findings described in the previous chapters, strengths and limitations, implications for public health, and suggestions for future research.

Main findings of this thesis

- Moderate dairy intake, particularly driven by high-fat cheese intake, was associated with a lower risk of prediabetes. The presence of fatty acids, the milk fat globule membrane, calcium, and vitamin K in the cheese matrix might exert beneficial effects on hyperglycaemia. Dietary guidelines focusing solely on low-fat cheese may not be adequately supported by the current literature.
- Milk intake was not associated with prediabetes risk, which is consistent with prior evidence in the literature.
- Yogurt intake showed no significant association with prediabetes risk in our study. The observed inverse associations with type 2 diabetes in previous research could be attributed to low intake levels and health behaviours associated with yogurt consumption. Additionally, inconsistencies might stem from the diversity of yogurt types consumed, including different probiotic strains.
- Ice cream intake was inversely associated with prediabetes risk in our analysis. However, it is important to note that ice cream intake levels were relatively low, its measurement may be subject to seasonal variation.
- There is a high likelihood of residual confounding by background diet and health behaviours, and reverse causation due to health or risk awareness.

Main findings in context

Associations between dairy intake and prediabetes in individual cohort studies

In the individual cohort studies, most dairy types were not significantly associated with prediabetes. There were inconsistencies in the associations between dairy types and prediabetes risk in each of the individual cohorts. In the Dutch Hoorn Studies (HS) (**Chapter 2**), we found that the top quartiles of high-fat fermented dairy, total cheese and high-fat cheese were associated with a lower risk of prediabetes. Associations of high-fat fermented dairy and total cheese were driven by high-fat cheese intake, as 52% and 83% consisted of high-fat cheese, respectively. Only higher consumption of high-fat cheese was continuously associated with lower risk of prediabetes. No significant associations of substituting high-fat cheese with other dairy types were found, although the effect estimates pointed to a higher risk of prediabetes when substituting high-fat cheese by other dairy types, except for ice cream.

In the Dutch Rotterdam Studies (**Chapter 3**), higher intake of high-fat milk and high-fat yogurt were associated with lower prediabetes risk, in the highest versus lowest intake categories as well as continuously in servings/day. Weaker associations for low-fat dairy, total milk, low-fat milk, and total and low-fat cheese were observed, across sub-cohorts and by exposure operationalization. Furthermore, high-fat yogurt intake was associated with lower longitudinal insulin resistance, while higher intake of low-fat dairy and total and low-fat milk were associated with a higher longitudinal insulin resistance. The results of this study suggested that mainly longitudinal waist circumference mediated associations of dairy intake and prediabetes, while blood lipids and hypertension did not play a major confounding or mediating role.

In the Australian AusDiab study (**Chapter 4**), higher intakes of high-fat dairy, high-fat milk, and total cheese were associated with lower prediabetes risk. Low-fat milk intake was nonlinearly associated with prediabetes risk; the risk was highest at 1.5 servings per day, with a decreasing risk at lower and higher intakes.

In the Dutch Lifelines study (**Chapter 5**), higher plain and low-fat milk intake were associated with higher risk of prediabetes in the top compared with bottom quartile, but no dose-response associations were observed. We showed that reverse causation may play a role in our findings, despite the utilization of a prospective design. In nutritional epidemiological studies, reverse causation occurs when individuals change their eating habits because of risk or disease awareness. Risk awareness, evident from a diabetes risk score or a desire to lose weight, was associated with a lower intake of high-fat dairy types and a higher intake of low-fat dairy types. Possibly, individuals without disease risk (e.g., no obesity) might choose their diet more hedonically instead of health conscious which may lead to a lower prediabetes risk associated with dairy types not included in dietary recommendations (i.e., high fat dairy types and ice cream).

In the UK Fenland study (**Chapter 6**), a higher intake of high-fat dairy and high-fat milk were associated with higher prediabetes risk. Associations of specific high-fat dairy types and cheese were similar but not significant. In contrast, a higher intake of low-fat milk was associated with a lower prediabetes risk. Increases in intake of low-fat dairy and low-fat milk during follow-up were associated with parallel changes in increasing fasting plasma glucose (FPG) and 2-hour plasma glucose (2hPG), but not with glycated haemoglobin (HbA1c). Increases in high-fat milk during follow-up were associated with decreasing FPG and with lower risk of progressing to prediabetes or type 2 diabetes (T2D). These findings suggest that baseline dairy intake and changes of dairy intake represent distinct exposure statuses, with the results for the parallel change analysis potentially affected by regression to the mean.

Interconnections of dairy intake, sociodemographic, health and cardiometabolic risk factors, dietary characteristics, and prediabetes in individual cohort studies

In the Lifelines and Fenland study (**Chapters 6 and 7)**, we employed a network approach to understand the key variables and clusters among dairy intake, dietary characteristics, and sociodemographic, health and cardiometabolic risk factors and prediabetes. Network models align with the multifactorial nature of prediabetes because it describes clusters of risk factors and dietary patterns as potential causes, moving beyond single causes as would be assessed in reductionist regression analyses. With this approach, we aimed to assess if the heterogeneity in associations of dairy intake and prediabetes risk is explained by covariation of health behaviour and food intake with the several types of dairy foods. We expected that individual dairy types would show distinct interrelations or clustering with other variables in the network, which would explain discrepancies in associations with prediabetes risk found in regression models.

The networks showed a preference of participants for either high-fat or low-fat dairy foods. Furthermore, both networks showed a cluster of food groups considered beneficial for health including vegetables, fruit, fish, legumes, tea, nuts, and wholegrains. This clustering may reflect health-conscious behaviour but could also to some extent reflect reporting bias of socially desirable answers. This cluster also included physical activity in the Lifelines cohort. In the Fenland study, high-fat yogurt, low-fat yogurt, and low-fat cheese intake connected with this cluster of healthy food groups, while only low-fat yogurt connected with this cluster in the Lifelines study. High-fat yogurt and low-fat cheese intake contributed more to total energy intake in the Lifelines study due to their higher consumption levels, positioning them closer to overall energy intake. Both networks also showed a cluster of energy dense food groups around energy intake, and a bridging role of waist circumference between prediabetes and sociodemographic characteristics.

The findings of the regression and network analyses showed limited alignment. In the Fenland study (**Chapter 7**), the network analyses showed that low-fat and high-fat yogurt intake were linked with the cluster of healthy food groups. Despite this similar placement, the associations in regression analyses (albeit non-significant) for high-fat and low-fat yogurt with prediabetes pointed in opposite directions. Furthermore, high-fat milk and low-fat milk were similarly placed in the network due to their strong connection with energy intake, despite opposite associations with prediabetes risk in regression analyses. The negative connections between high-fat and low-fat dairy types indicate an inverse relationship between their intake levels, explaining conflicting associations in regression analysis. Moreover, while low-fat cheese intake was connected to the cluster of healthy food groups, high-fat cheese connected to energy intake. Nonetheless, the associations for high-fat and low-fat cheese were in similar positive direction in regression analysis. In the Lifelines cohort (**Chapter 6**), low-fat milk did not have a much more pronounced

role in the network based on predictability and centrality compared to other dairy types, while being the only dairy type associated with prediabetes risk in regression models. Its connection to energy intake, energy-dense foods and waist circumference might explain the positive association found. While these connections offer insights into the observed relationship, it is noteworthy that other dairy types were also connected with these factors, suggesting that these connections alone may not fully explain the positive association found for low-fat milk intake and prediabetes. Overall, the networks analyses showed a nuanced understanding of dietary behaviour beyond multiple regression models but did not fully explain the heterogeneity in observed associations.

Integration of evidence on dairy intake in relation to continuous glycaemic markers and prediabetes

The systematic review including 14 prospective cohort studies on the relation between dairy intake and continuous glycaemic measurements with diverse analytical approaches presented a mix of inverse, positive, and non-significant associations (**Chapter 7**). In the meta-analysis of six studies across nine prospective cohorts examining dairy intake in relation to prediabetes risk, several key findings emerged. A quadratic nonlinear inverse association was observed for total dairy intake, indicating the lowest risk at 3.4 servings per day, with no clear trend for high-fat or low-fat dairy. Both total and high-fat cheese displayed nonlinear associations with prediabetes risk, showing the lowest risk at 2.1 servings per day, but a positive risk at intakes surpassing 4 servings per day. High-fat cheese intake likely drives the association for total dairy due to its high relative contribution to a dairy serving. These results are somewhat in line with evidence for T2D, with a comprehensive meta-analysis of prospective cohort studies showed that high compared to low total cheese intake was associated with lower risk, though no dose-response association was evident [1]. Additionally, ice cream intake showed an inverse linear association with prediabetes risk, but caution is warranted given that the highest median intake level recorded was less than a serving per week. At these low intake levels, ice cream intake is unlikely to have such a strong significant impact on the development of prediabetes. Furthermore, potential reverse causation may explain this inverse association as well as in studies with T2D as outcome [2, 3]. We found no associations for total, high-fat and low-fat milk, and cream intake, consistent with evidence for T2D [4-6].

Also, we found no association between fermented dairy and yogurt intake and prediabetes, despite yogurt being consistently linked to T2D in previous prospective cohort studies [1, 2, 7-9]. The meta-analysis by Gijsbers *et al.* (2016) with subgroup analyses by continent showed that the associations between yogurt and T2D were strongest in the US (*n* = 5, RR per 50 g/day 0.91, 95%CI 0.86-0.96, I^2 = 73%) populations, followed by Asian (*n* = 2, RR 0.95 per 50 g/day, 0.79-1.14, I^2 = 65%) and European (*n* = 4, RR 0.96, 95%CI 0.92-1.01, I^2 = 68%) populations, but not in one Australian population (*n* = 1, RR 1.08, 95%CI 0.92-1.27) [2]. Most of these US cohorts had baseline measurements

before 2000, during which the yearly per capita yogurt consumption in the US was low (1980: 1.1 kg per person, compared to 16.6 kg in the Netherlands, 2000: 2.9 and 20.1 kg, respectively) [10, 11]. Consequently, these older US cohorts typically have limited yogurt intake. This restricted consumption of yogurt has been associated with healthy dietary patterns and behaviours [3, 12], and this may impact the observed associations within these US cohorts. In contrast, in the European populations, and the cohorts included in our meta-analysis, especially those in the Netherlands, yogurt consumption is more common. While correlations with healthy behaviours may still exist, they may be less pronounced due to a wider range of yogurt intake levels. This variability could dilute the strength of the observed associations between yogurt intake and T2D in prior meta-analyses, as well as with prediabetes in our meta-analysis.

The single prior observational study on this topic within the FHS-OC proposed that the associations between dairy intake and T2D could differ based on participant's baseline glycaemic status (i.e., normoglycaemia, prediabetes or T2D) [13]. This effect modification was mainly observed for total milk and cheese intake, but not for yogurt intake. By stratifying individuals with prediabetes from normoglycaemia and T2D, we aimed to address for this potential source of heterogeneity in prior meta-analyses. However, despite this adjustment aimed at creating a more homogenous baseline sample, inconsistencies in associations across individual cohorts described in this thesis persisted. Furthermore, our findings were consistent with associations shown in previously published meta-analyses for total dairy, total milk, and ice cream, and somewhat with cheese intake, with findings regarding T2D outcomes [2, 4]. This suggests that the baseline glycaemic status of participants may not be the primary factor driving heterogeneity across studies.

Prediabetes is a multifactorial condition influenced by numerous risk factors exerting small effects, and no single exposure can be solely causally linked to its development. In the systematic review (**Chapter 7**), we used NutriGrade, which is based on criteria-based inferential methods, to evaluate the evidence and establish causal or preventive conclusions aimed at informing dietary guidelines. The level of confidence in the estimate was moderate for total dairy and ice cream, very low for cream, and low for all other dairy types. This was primarily attributed to the small number of cohorts included, limited strength of association and the non-significance of dose-response associations. The risk estimates ranged from 0.50 to 1.05, with most associations within the range of 0.90 to 1.00. However, despite their weakness, these associations may have a significant public health impact, given the widespread consumption of dairy and the high incidence of prediabetes. A significant U-shape dose-response curve was observed for total dairy, total cheese, and high-fat cheese intake. These curves suggested an optimal level of consumption beneficial for lowering prediabetes risk, but deviating from this level in either direction could lead to reduced effectiveness or increased risk.

Evaluation of the consistency of associations across populations and study designs is crucial for determining the plausibility and generalizability of findings. Inconsistencies in findings do not necessarily negate the validity of observed associations. Some associations might be specific to certain population groups. The meta-analysis by Gijsbers *et al.* (2016) suggested inverse associations in Asian populations of total dairy (*n* = 3, RR 0.85 per 200 g/day, 95% CI: 0.65, 1.12, I^2 = 88%) and milk (*n* = 3, RR 0.87, 95%CI 0.72-1.05, I^2 = 71%) with incident T2D, but not in American, European, or Australian populations (*n* = 1-6, RR range 0.93-1.03). This pattern was not seen for total yogurt and cheese intake. Differential associations by fat content of dairy foods and milk could not be compared by continent due to lack of studies in Asian populations. Another meta-analysis including additional cohorts (*n* = 24, 13,211 cases among 1,297,951 participants) also showed that milk intake was associated with lower T2D risk in non-white populations (RR per 245 g/day: 0.80, 95%CI 0.66-0.96), while in white populations, a modest positive association was found (RR per 245 g/day: 1.03, 95%CI 1.01-1.04) [14]. The findings of these two meta-analyses align with our findings for milk and low-fat milk in relation to prediabetes risk based on predominantly White European populations and one Australian populations.

To further infer on consistency of these findings, they can be compared across several study designs including Mendelian randomization (MR) studies and randomized controlled trials (RCTs). MR studies provide insights on the potential causality of exposures on outcomes by utilizing genetic variants as instrumental variables (IVs). These studies leverage the random assignments of genetic variations during conception, reducing residual confounding and reverse causation biases. For milk intake, the single nucleotide polymorphism (SNP) rs4988235 on the lactase (LCT) gene was found to be associated with lactase persistence (LP) in European populations and therefore allows for examination of natural randomization in milk intake. Thereby, MR studies are confined to assessing milk intake solely as the LCT-rs4988235 is not associated with intake of cheese or yogurt as these contain less lactose. Thus, these studies may underestimate the full impact of dairy consumption on health outcomes [15]. The findings of MR studies align with the current observational evidence for milk intake [4-6]. A systematic review of 6 MR studies in 12 countries concluded that this genetic marker was not associated with T2D risk or HbA1c levels [5]. Furthermore, a MR study among 182,041 participants from 18 cohorts also showed no association between the genetic marker and glycaemic markers including FPG, fasting insulin, insulin sensitivity and insulin resistance [6]. Our inverse association of total cheese and prediabetes risk was consistent with a MR analysis using a genome wide association study (GWAS) to identify SNPs associated with cheese intake as measured with the FFQ [16]. This study showed that genetically predicted cheese intake was associated with lower fasting glucose (β -0.20, 95%CI -0.33; -0.07) and odds of T2D (OR 0.46, 95%CI 0.34-0.63). GWAS has the potential to uncover novel genetic markers and pathways associated with dairy intake, beyond known biological markers such as the LCT-rs4988235. However, a statistically significant association in GWAS does not necessarily imply causality.

Currently, no randomized controlled trials (RCT) examine the effects of dairy intake on prediabetes. Regarding glycaemic markers, RCTs suggest inconsistent effects of high dairy intake compared to the control arms [17-23]. A meta-analysis of 34 RCTs (*n* = 2,678) showed that high compared to low dairy intake (mean 3.1 versus 0.5 servings/day) was not associated with FPG changes in long-term studies (> 24 weeks, the maximum duration was 48 weeks), with energy restriction, and glucose metabolism as primary outcome [18]. Positive associations were observed for RCTs with liquid dairy types, low-fat dairy types and strictly controlled dietary interventions. Two meta-analyses including fewer studies found no effects on FPG [17, 20]. For HbA1c, a meta-analysis of 8 RCTs (*n* = 682) indicated no effect [20], while another one found an inverse association (4 RCTs, *n* = 512, mean difference -0.09%, 0.16; -0.03, I^2 = 0%) [18]. Regarding HOMA-IR, two meta-analyses showed no effects of high dairy diets [17, 18], while one meta-analysis including most RCTs reported an inverse association (14 RCTs, *n* = 314, -1.21, 95%CI -1.74; -0.67, I^2 = 92%) [19]. One meta-analysis examined effects of high dairy diets on fasting insulin, reporting no mean differences compared to low dairy control arms (29 RCTs, *n* = 1,902) [18].

Consistent with the findings from our meta-analysis (**Chapter 7**), limited trial evidence also does not indicate discernible effects of dairy fat content on glycaemic outcomes [23-25]. A landmark trial by Schmidt *et al.* (2020) compared the effects of high-fat versus low-fat dairy diets (3.3 servings/day) versus a control (≤3 servings/week of non-fat milk) in a 12-week parallel-design RCT with a 4-week wash-in period, including 72 participants with the metabolic syndrome [23]. After 12 weeks, no effects were observed on the primary endpoint glucose tolerance, as well as on the secondary endpoints FPG and HbA1c. Both dairy diets resulted in an increase in fasting insulin and HOMA-IR, and a decrease in insulin sensitivity. These effects were independent of pancreatic β-cell function, liver fat function, biomarkers of systemic inflammation, total adiponectin, and the slight weight gain observed in both dairy arms. Adjusting for dietary changes and physical activity did also not change these intervention effects. Similar results were found in a RCT by Eelderink *et al.* (2019) comparing a low-fat dairy diet (5-6 servings/day of milk, yogurt, and cheese) to a low dairy control arm (≤1 serving/day) for 6 weeks in a cross-over design with a 4-week wash-out period in 45 overweight individuals [22]. They found no differences in FPG, postprandial glucose and insulin response. After the high dairy diet, fasting insulin was higher (8.9 ± 3.3 mU/L versus 8.1 ± 2.8 mU/L), resulting in higher HOMA-IR (2.21 ± 0.91 mU/L versus 1.99 ± 0.72). Differences in HOMA-β and the Matsuda index exhibited a similar trend. This RCT also accounted for the increase in body weight in the dairy arm. Engel *et al.* (2018) compared the effect of skimmed milk versus whole milk intake (0.5 L/d) on FPG and IF a 3-week crossover trial with no wash-out period in 18 healthy adults. No significant differences were observed in FPG (5.24 ± 0.07 versus 5.32 ± 0.09 mmol/L), FI (41.99 ± 4.13 versus 45.66 ± 4.23 mmol/L) and HOMA-IR (1.37 ± 0.14 versus 1.50 ± 0.14).

We observed an inverse association for high-fat cheese at moderate intake levels, but not low-fat cheese with prediabetes. Limited trial evidence shows no distinct effects of cheese with varying fat content on glycaemic outcomes [24, 25]. Nilsen *et al.* (2015) compared the effects of a traditional fat-and salt-free Norwegian cheese (50 g/day), Gouda-type cheese with 27% fat (80 g/day), and a low cheese control group in an 8-week parallel-design RCT in 153 participants [24]. FPG showed a slight increase in the entire population, with no effects attributable to the different interventions. Raziani *et al.* (2016) compared the effects of 80g (mean 64-112 g) regular-fat cheese with an equal amount of reduced-fat cheese and a non-cheese carbohydrate control in a 12-week parallel-design RCT preceded by a 2-week run-in period in 139 participants [25]. No effects on FPG, fasting insulin and HOMA-IR were found.

Nevertheless, many RCTs on examining the effects of total dairy and dairy types on glycaemic outcomes are of low quality and difficult to compare due to heterogeneity in the study population (i.e., age, metabolic status, comorbidities), duration, different interventions, and control arms (e.g., SSBs, tea or water, low dairy, or usual diet) [17, 18, 26]. Often, there is no distinction made or clarification provided regarding the types of dairy and fat content. Most RCTs used skimmed or low-fat fluid milk in their intervention groups. Few studies used high-fat or fermented foods, either as a minor component of the diet or without specification. Additional well-designed trials are therefore needed.

Relevant molecular pathways, confirmed by short-term animal and human trials, of dairy constituents and the dairy matrix are summarised in **Chapter 1**. These potential mechanisms mainly support biological plausibility for neutral or inverse associations of moderate dairy intake with glycaemic outcomes. Whether these mechanisms explain the long-term observational associations of dairy and hyperglycaemia found in this thesis and prior evidence is yet to be determined. It remains not possible to pinpoint a specific mechanism considering the synergistic or interactive effects of various components in dairy foods. Potential counteracting effects of beneficial and harmful nutrients may result in null effects, further complicating the interpretation. Furthermore, their exposure is never null as most of these nutrients are derived from various sources. It is not entirely clear if dairy foods exert direct effects on glycaemia, or if the associations are due to indirect effects on satiety [27], body composition [6, 28], lipid profile [29], and gut microbiota [30]. Different associations for each dairy type could be due to dairy matrix effects. Inverse associations of milk intake and T2D in non-White populations could potentially be attributed to the high proportion of lactose intolerance, with undigested lactose affecting microbiota composition and activity, resulting in alterations in circulating metabolites with beneficial effects on hyperglycaemia [31]. The beneficial associations of yogurt intake and T2D have been attributed to probiotics [32, 33], with the precise impact depending on the specific probiotic strains and their dosage [34], and inconsistent findings observed may be attributed to variations in the yogurt types consumed. Inverse associations between high-fat cheese intake and prediabetes

risk could be attributed to the unique cheese matrix with effects of the intact MFGM, high calcium content, and semi-solid protein matrix on fat absorption [35] and energy balance [36, 37], as well as effects of vitamin K2 on insulin sensitivity [38-40]. The nutritional content is highly similar between high-fat and low-fat dairy varieties, except for the MFGM, fatty acids and fat-soluble vitamins. This may significantly impact gastric emptying and dairy fat kinetics [35]. Nevertheless, it is unlikely that these nutritional differences can fully account for the sign reversal for high-fat versus low-fat dairy types observed in some cohorts described in this thesis. Other factors such as individual metabolic responses, compensatory dietary patterns and health behaviour could play significant roles in shaping these associations.

Methodological considerations

Study design and populations

In this thesis data from five observational prospective cohort studies in Western adult general populations were used. The findings of these studies need to be interpreted considering internal and external validity. Internal validity assesses the degree of confidence that the observed associations in a study are accurate and not influenced by biases. These biases include selection bias, confounding bias, and information bias, particularly in nutritional studies related to dietary assessment. Selection bias arises when the associations being assessed differ between the individuals included in the analyses and those who were eligible but did not participate. This bias can result from selective participation at baseline or from selective lost-to-follow up. The Fenland study and AusDiab study had relatively low response rates at baseline (27% and 37%, respectively) and exhibited a slight underrepresentation of individuals with a low SEP compared to the general population figures [41, 42]. The higher baseline response rates in the HS-1 (71%) and Rotterdam Study (72%) suggest good representativeness. Both the HS-1 and HS-2 (response rate 45%) were comparable to the study population regarding sex and self-reported diabetes [43]. Determining the exact participation rates of the Lifelines study is challenging due to its inclusion strategy, but overall, the Lifelines study was found to be representative of the population of the north of the Netherlands [44]. In general, participants in population studies tend to be healthier and more affluent than nonparticipants. However, selective participation in cohort studies at baseline has been shown to not bias exposure-outcome associations [45-47]. The response rates at follow-up ranged from approximately 60-75% in each of the included cohorts which has been suggested as acceptable when follow-up data is missing at random [48]. An overrepresentation of a healthier population with higher educational levels may impact the external validity of the results, referring to the generalizability of findings to the broader population of interest. The prediabetes incidence rate might be underestimated compared to the general population. Additionally, non-Western ethnicities and individuals with lower socio-economic status (SEP) were underrepresented in our studies. Only in the Fenland study, we were able to differentiate between White and non-

White ethnic background (**Chapter 6**). Therefore, our findings are primarily generalizable to the general Western population with substantial consumption of dairy foods, high lactose persistence, and background diets with considerable SFA, protein, vitamin, and mineral intake. For example, the non-Western Prospective Urban Rural Epidemiology (PURE) study includes low- and middle-income countries with low background SFA intake (i.e., lowest intake was 2% SFA, which is considered inadequate) [49]. Higher intake of high-fat dairy in these countries would result in nutritionally adequate diets rather than potential overconsumption of SFA. Indeed, higher total dairy intake was associated with lower diabetes incidence (5,351 cases among *n* = 131,481 across 21 countries worldwide, HR ≥2 versus 0 servings/d 0.88, 95%CI 0.76-1.02, P_{trend} = 0.01), with similar directionality for high-fat dairy, milk, yogurt, and cheese, but not for low-fat dairy.

Confounding

In observational cohort studies, participants are not randomly assigned to an exposure, which can lead to potential bias due to confounding. Since participants self-select their exposure levels, these levels may be related to other characteristics. If these characteristics are also associated with the outcome, they could explain the observed associations. Contrasting the baseline characteristics between individuals with high and low intake of certain dairy types revealed some interesting patterns (**Chapter 2-6 and *Intermezzo***). Baseline tables provided insights into potential confounding factors associated with intake of certain dairy types. Individuals with high total dairy intake were generally more often male, with higher physical activity and diet quality, and higher intake of total energy, calcium, sodium, fruits, and SFAs. Other characteristics were inconsistently related to higher total dairy intake, such as smoking behaviour, alcohol intake, educational level, obesity, hypertension, and dyslipidaemia. Individuals with high intake of high-fat milk were more often male, current smoker, with higher energy intake but lower diet quality compared to individuals with low intake of high-fat milk. In contrast, those with high intake of low-fat milk were more often non-smokers with higher diet quality. Individuals with high yogurt intake were more often female, never smokers with higher diet quality and lower waist circumference. Individuals with high intake of high-fat cheese were more often male, with higher education, but smoking behaviour, energy intake and diet quality varied. Individuals with a high intake of low-fat cheese were more often female, reported higher physical activity, lower energy intake, and higher diet quality. Specific dairy types, especially high-fat yogurt, and low-fat cheese are consumed by a limited number of individuals with distinct characteristics related to health behaviour and cardiometabolic health. A low proportion of consumers with unique characteristics resulted in strong, but imprecise associations with prediabetes risk in the individual cohorts (**Chapter 2-6**). The inconsistent associations between dairy foods and the risk of prediabetes and T2D may be due to the varying proportions of consumers with specific characteristics within each cohort for each type of dairy.

Each of the included cohorts extensively measured potential confounders and therefore we could adjust for a wide range of confounding factors in different domains. Adjusting for total energy intake had the most impact on effect estimates, likely as this also addresses confounding factors related to body size, metabolic efficiency, and physical activity [50]. With adjustments made for sex, age, and energy intake, further adjustments did not result in major differences in effect estimates. Baseline body weight, hypertension, and dyslipidaemia may affect (reporting of) dietary patterns, and are risk factors for prediabetes, and therefore meet the criteria for confounding [51, 52]. Adjusting for these factors might also help to limit confounding of health factors associated with body weight, hypertension, and dyslipidaemia, including genetic predisposition, overall health, lifestyle, and SEP. Adjusting for obesity also aims to account for the potential mediating pathway of adipose tissue dysfunction in obesity-related insulin resistance [53], estimating the direct effect of dairy on prediabetes. However, distinguishing between confounding and mediating effects is not possible due to the use of only baseline measures and the absence of measurements to assess temporality. Comparing effect estimates based on models with and without adjustment for mediating effects remains imperative. The degree of attenuation observed may further depend on operationalization of the confounder; whether baseline values or changes during follow-up are considered, as well as whether continuous markers or dichotomized risk states are evaluated.

The potential for residual confounding remains, either by measurement errors in the assessment of confounding variables or because of unmeasured confounders. Unmeasured differences in population characteristics (i.e., SEP, lifestyle, dietary intake, genetic predispositions), health status, eating habits and replacement choices for dairy consumption offer a significant explanation for heterogeneous associations of dairy and prediabetes reported in each chapter and in prior literature. An important confounder in diet-disease relations which is not optimally assessed in prospective cohort studies is physical activity. Except for the Fenland study which used heart rate and movement sensors (**Chapter 6**), data collection of physical activity in all cohorts involved self-reported questionnaires, which are useful for obtaining population level insights but are prone to recall and response bias and inability to capture absolute physical activity [54]. In **Chapter 5**, one explanation for weaker interconnections in the network model for health risk factors including physical activity and food groups compared with clinical markers is that they reflect a greater extent of measurement uncertainty. In comparison, in **Chapter 6**, objectively measured physical activity was linked more closely to the clinical marker cluster possibly suggests a higher degree of precision in measurement and association.

Network models

Our application of different confounder models, each with consecutive additional adjustments for different variable domains, provides a structural and systematic manner

to identify the impact of different sets on variables. However, the exact impact of each variable is difficult to discern. With Gaussian graphical network (GGM) models, the complex relationships between dairy intake, prediabetes incidence, sociodemographic, health and cardiometabolic risk factors are assessed at once and visualized in a network plot [55-57]. GGMs may elucidate how dietary factors interact with each other and with health outcomes, providing insights in direct and indirect relationships [55]. Furthermore, this is a sound methodological approach to obtain holistic insights in structures among variables, providing insights in influential factors, clusters, or specific pathways [55, 58]. GGM addresses confounding present in nutritional epidemiology by obtaining conditional dependencies derived from regularized joint distributions. We applied these network models in an exploratory fashion in two populations. Consistent patterns across cohorts increase confidence in reliability and may strengthen the evidence for causal relationships or underlying mechanisms, while discrepancies highlight population-specific factors or differences in variable measurement methods. Nevertheless, visual interpretation of these large networks with variables from different domains proved to be challenging. The presence of many small connections makes it difficult to discern meaningful patterns, and the placement of variables is algorithm-based, leading to potential fluctuations in their relative positions with each re-estimation. As a result, the visual representation may not fully reflect the strength or nature of interrelationships [59]. Furthermore, networks also suffer from the biases in nutritional epidemiology including skewed exposures, limited variability in exposures, missing data, measurement errors and residual confounding. Hypothetically, if a variable that lies on the causal pathway of two other variables is not included in the network, a direct relationship between these two other variables can be created. Then, predictability is not defined as influence by neighbouring variables, as they are not influenced by each other but are caused by this variable on the causal pathway which was not included in the network [60].

Dietary assessment

A strength of this thesis lies in the investigation of various dairy types, including total dairy, fermented dairy, and specific dairy varieties, considering alignment with previous research while acknowledging limitations in the level of detail of the FFQs in each cohort. Self-reported dietary data provides information that is difficult to obtain any other way. FFQs have become the standard measurement in nutritional epidemiology, as they are self-reported in a single administration, easily processed and cheap. The fundamental concept of FFQs is that the exposure of interest for health outcomes is the average long-term dietary pattern, spanning periods of a month or a year, rather than isolated intakes on individual days. This approach is based on the notion that a dietary pattern maintained over an extended period may be more appropriate and practical than single daily diets for certain research questions. Indeed, numerous studies showed that FFQs can be used to rank individuals according to their intake, and that FFQ based intakes are predictive of health outcomes [61]. Each FFQ of the individual cohorts included in this

thesis demonstrated moderate to strong correlations with dairy intake or major dairy nutrients compared to other independent dietary assessment methods, suggesting good validity. However, it is important to note that validation specifically for distinguishing between high-fat and low-fat dairy intake is lacking.

The inaccuracy of FFQs to measure absolute intake levels has implications for determining consistency across studies using different cut-off points. Variability in number, formulation and grouping of dairy types in each FFQ limits comparability of dietary exposures and likely contributes to inconsistencies in findings between studies. The pre-determined food list on which a FFQ is based is often defined based on explaining the greatest amount of variance in population dietary habits, or on foods contributing most to nutrients of interest to certain populations [62]. In the Netherlands, where dairy consumption is widespread compared to other countries, FFQs are tailored to include detailed assessment of dairy products due to their significant role in the diet. The Dutch FFQs used in more recent cohorts (HS-II, Lifelines, RS-III) provided examples of several dairy types, frequency of consumption as well as usual serving sizes. In contrast, the AusDiab and Fenland FFQs only assessed the frequency, type, and amount for milk, other dairy types were assessed with a single frequency question. Also unmeasured differences in nutrient content contribute to inconsistent findings. Although some information about fat content is assessed (e.g. skimmed, semi-skimmed or full-fat), protein, sugar and micronutrient content is not assessed. Furthermore, FFQs are limited in assessment of seasonal variation in intake, for example ice cream consumption is likely to differ between winter and summer.

The main limitation of FFQs is that its self-reported retrospective nature is prone to recall bias resulting in substantial measurement error. This measurement is assumed to be mainly non-differential (i.e., independent of the outcome) and therefore results in attenuation of the effect estimates with larger uncertainty [63]. Although the misclassification errors might not relate to prediabetes, they might relate to risk factors of prediabetes including obesity and might thus be differential. People with obesity are more likely to underreport energy intake and certain foods or overreport others such as fruits and vegetables due to social desirability bias [64, 65]. Possibly, people at higher risk of prediabetes may overreport their intake of low-fat dairy types, to align with dietary recommendations and perceived norms emphasizing to consume low-fat or low-calorie foods. With differential measurement error, the direction of bias in effect estimates cannot be predicted.

We employed different strategies to account for measurement errors. We performed sensitivity analyses with dairy intake adjusted for total energy intake with the residual method. Energy-adjustment can partly mitigate the effects of correlated measurement errors in self-reported dietary assessment methods, isolating the effects of specific dietary factors on health outcomes. No differences in effect estimates were found.

Furthermore, in **Chapter 2**, we applied the Goldberg method, to identify and exclude energy under- and over reporters based on the ratio of energy intake (EI) and formula based basal metabolic rate (BMR) [66, 67]. Approximately a quarter of the sample was identified as energy misreporters. Their exclusion resulted in strengthened associations between dairy intake and prediabetes risk, highlighting the potential attenuation of effect sizes in main analyses due to energy misreporting. Applying this method as a sensitivity analysis in nutritional epidemiological studies could offer insights into the impact of potential measurement errors on the accuracy and reliability of effect estimates. However, it is important to recognize that assumptions underlying formulas and cut-off points may introduce misclassification bias and excluding a substantial portion of the sample results in loss of power [68]. Moreover, in well-designed studies and FFQs, reliance on the Goldberg method may be unnecessary. In each study described in this thesis, the FFQs were rigorously validated against energy intake, and participants with implausible energy intakes outside the sex-specific ranges were already excluded (<500 and ≥3500 kcal/day in women and <800 and ≥4000 kcal/day in men) [50]. The way forward is to improve validation of dietary assessment methods and identification of energy misreporters using objective measurements of total energy expenditure. This data could also be used for employing regression calibration to correct for measurement errors in dietary assessment methods.

Biomarkers of dairy intake

Objectively measured biomarkers of dairy fat intake may be used to reduce misclassification errors in self-reported dietary questionnaires and unmeasured dairy intake in mixed dishes (e.g., cheese on pizza or cream in cake). Hypothesis-driven derived biomarkers of dairy fat include circulating and adipose proportions of pentadecanoic acid (C15:0), heptadecanoic acid (C17:0), and trans-palmitoleic acid (*t*16:1n7), as these occur in ruminant milk and are no or little endogenous FA production in the body [69, 70]. While these OCSFA are not specific for dairy and are also present in meat and fish, they have been widely studied and recognized for their usability as dairy fat biomarker [71]. These fatty acids correlate well with self-reported intake of total dairy, high-fat dairy, and dairy fat (r = 0.4 to 0.7) based on 24h recalls or 7-day food records [72-74]. These fatty acids increase in response to dairy intake or decrease when replacing high-fat with low-fat dairy in trials [75-78], and are intercorrelated (r = 0.3 to 0.8), while representing two distinct fatty acids classed with distinct chemical structures and associated metabolic pathways [71]. Correlations of dairy fat biomarkers with dairy intake from FFQs are lower (r = 0.10 to 0.33) as compared to 24h recalls or 7-day food records, as these might miss more 'hidden' sources of dairy intake [79]. Generally, correlations with dairy intake are stronger for C15:0 and endogenous FA production contributes less to C15:0 than to C17:0. Furthermore, of prominent SFA in dairy fat, myristic acid (C14:0) can be used as hypothesis-driven biomarker of dairy fat [80, 81], considering that in contrast to C16:0 and C18:0 it is not produced by de novo lipogenesis. However, as C14:0 is also present

in meat, fish and grain, the correlations with total dairy intake are lower as compared to C15:0, C17:0 and *t*16:1n7 (r = 0.15, 95%CI 0.11-0.19) [81].

The dairy fat biomarkers C15:0, C17:0 and *t*16:1n7 have been linked to lower incidence of T2D [71, 82], while C14:0 is linked to higher T2D incidence [82]. A pooled analysis of 16 prospective studies (7 in the US, 7 in Europe, 1 in Australia and 1 in Taiwan, 15,180 cases among 63,682 participants, mean 9 y follow-up) showed that higher C15:0 levels were associated with 26% lower T2D risk (RR per 10^{th} to 90^{th} percentile range 0.74, 95% CI 0.68-0.8=80.2%0), adjusted for demographic, clinical, socioeconomic, and health factors [71]. For 17:0 and *t*16:1n7 were associated with a 45% (RR per 10^{th} to 90^{th} percentile range 0.55, 95% CI 0.49-0.62=88.4%) and 19% (RR per 10^{th} to 90^{th} percentile range 0.81, 95% CI 0.69-0.94, I^2=7.1%) lower risk, respectively. A meta-analysis of 7 studies (13,596 cases among 38,813 participants) showed that C14:0 was associated with a 13% higher T2D risk (RR per SD 1.13, 1.09-1.18, I^2 = 42%, mean 10.8 y follow-up) [82].

Metabolite scores or profiles are data-driven biomarkers of dairy fat calculated as composite measures of many different metabolites. In the UK EPIC-Norfolk diabetes case-cohort study (641 cases among 1,440 participants, 16,350 person-years), metabolite scores for total dairy and milk were inversely associated with T2D (HR 0.66, 95%CI 0.60-0.72 and HR 0.75, 95%CI 0.65-0.82, respectively) [80]. In PREDIMED, the multimetabolite profiles for total dairy, high-fat dairy, low-fat dairy, milk, yogurt, and energy-adjusted cheese were each associated with lower T2D risk [83]. In contrast, in the US confirmatory cohort, only the multimetabolite profile for total dairy and milk intake were associated with lower risk of lower T2D risk. Biases due to the observational nature of these prospective cohorts and different consumption patterns in the various countries contribute to the differences in strength of relations of these dairy fat biomarkers with health outcomes [84]. Therefore, the direct role of existing and novel biomarkers needs to be validated in well-designed intervention studies [85].

To further understand the associations of dairy intake with prediabetes risk, we aimed to assess how these dairy fat biomarkers influence these associations in the Fenland study (**Chapter 7)**. Attenuation of the association between high-fat dairy intake and prediabetes risk by C14:0 may as we found in the Fenland study indicate a mediating role of this metabolite in this relationship. After adjustment for other biomarkers such as C15:0, C17:0, *t*C16:1n7, and all metabolites together, the associations between dairy intake and prediabetes risk were slightly strengthened. This might indicate higher specificity of findings and more reliable estimates with the use of biomarkers due to less influence of measurement errors in self-reported dairy intake. The biomarkers did not significantly contribute to the inverse associations for low-fat milk intake and prediabetes risk, as these biomarkers are specific to dairy fat and thus do not correlate with low-fat dairy types.

Prediabetes definition

Considerations regarding the definition of prediabetes have implications for both prevalence rates and prognosis. Controversy about the optimal definition of prediabetes and inconsistencies in its association with T2D and CVD has raised doubts about its utility in the medical community [86, 87]. For example, the T2D incidence ranged from 9-84% depending on follow-up duration and prediabetes definition in a systematic review of 103 prospective studies up to 2018 from the Cochrane library [88].

We utilized various analytical strategies to explore the impacts of the prediabetes definition on the association between dairy intake and prediabetes risk. In the Hoorn Studies **(Chapter 2)**, we addressed possible misclassification in prediabetes defined at baseline by repeating analyses including participants with prediabetes at baseline (20.9%), resulting in attenuation of associations. Rather than removing misclassification errors, this approach likely introduced non-differential misclassification bias resulting in attenuation of associations [63]. In the AusDiab and Lifelines study (**Chapter 4 and 5**), we repeated the analyses using the ADA cutoff levels for prediabetes [89]. This also resulted in attenuation of associations. Much more prediabetes cases are identified based on the ADA cut-offs compared to the WHO-ICE cut-offs, including a lower-risk group with a better cardiometabolic risk profile. Lower cut-offs reduce homogeneity in the prediabetes outcome group and increase the risk of misclassification bias, resulting in attenuation of associations.

In the meta-regression analysis in **Chapter 7**, the prediabetes definition explained some heterogeneity in the associations for high-fat milk and high-fat yogurt with prediabetes. Specifically, non-significant positive associations were found in cohorts using the WHO-IEC definition (HS and Fenland study), but non-significant inverse associations were found with FPG (FHS-OC) or FPG/NPG (RS) definition. Nevertheless, the limited number of studies for each definition hinders a comprehensive elucidation of this potential source of heterogeneity.

In the Fenland study **(Chapter 6**), we compared the participant selection at baseline and associations with dairy intake using four prediabetes definitions differing in the included glycaemic markers. The prediabetes incidence ranged from 1.5% based on FPG only to 4.4% based on the WHO-IEC definition. The different definitions identify varying numbers of participants with prevalent and incident prediabetes and T2D. For most dairy types, the effect estimates were largely consistent across the different prediabetes definitions. Some associations were somewhat stronger when using the FPG only based definition compared to those using the WHO-IEC definition. While the meta-regression in **Chapter 7** revealed heterogeneity in associations between high-fat milk and prediabetes depending on the prediabetes definition used, the Fenland study showed highly comparable effect estimates for high-fat milk across different prediabetes definitions, with the strongest association observed when including HbA1c

in the definition. In accordance with the meta-regression in **Chapter 7**, however, the association of high-fat yogurt was inverse with the FPG based definition while positive with the WHO-IEC definition. This was also observed for low-fat cheese intake. For yogurt intake, strong inverse associations were seen with FPG, but associations were more attenuated or positive with the other definitions. In line, distinct associations of dairy intake on continuously assessed FPG, 2hPG and HbA1c were evident when analysing repeated measures with linear mixed models (**Chapter 6**). Overall, the combination of glycaemic markers to define prediabetes may explain some inconsistencies in associations between cohorts, particularly for specific high-fat and low-fat dairy types. This variability could arise from distinct associations with each glycaemic marker or differences in participant selection, particularly impacting associations with specific dairy types with fewer consumers.

Substitution analyses

This thesis presents a study investigating the associations of substituting high-fat cheese with other dairy types on prediabetes risk using leave-one-out substitution models (**Chapter 2**) [90]. This compares persons with different dietary habits, assuming the validity of baseline dietary measurement. Substitution analyses within the dairy food groups were not repeated in other cohorts, as the non-significant effect estimates from this analysis aligned with the main findings. Additionally, substitutions within the dairy food group present challenging interpretation due to contextual differences in consumption. The substituted dairy items are often consumed in diverse contexts, for example as a beverage, combined with cereals, spread on bread, or eaten as a snack versus as part of a main meal. Prospective cohort studies often lack comprehensive data on contextual factors that influence dietary choices and substitutions, limiting the assessment, the feasibility, and real-world impact of these substitutions. For these substitutions to be effective as health recommendations, they must align with typical dietary patterns and be practical for everyday eating habits. Another methodological consideration is that substitutions with total servings of intake mirrors real-life consumption quantities and may hold greater relevance for public health considerations, yet substitutions are not completely iso-caloric [91]. This means that total energy intake varies as different components are substituted, and thus the problem of unspecified substitutions is reintroduced (i.e., as is present in main analysis) [50]. We aimed to mitigate this issue by adjustment for total energy intake. Differences in intake ranges or misclassification of the dietary exposure may further contribute to this issue. Instead of using mixed units as in our model, isocaloric and equal-mass substitutions (i.e., all included variables are in the same unit, for example E%) aligning with the compositional nature of dietary data could provide more reliable effect estimates and are therefore recommended in future research [92].

Parallel change analysis compared to linear mixed models

In both cross-sectional studies (**Chapter 1**) and prospective studies (**Chapter 7**) a diverse array of statistical approaches was used to analyse associations of dairy intake and continuous glycaemic markers. We conducted a parallel change analysis in the Fenland study, examining changes in dairy intake in relation to changes in glycaemic markers (**Chapter 6**). This parallel change approach was based on a study by Smith *et al.* (2015) relating dietary factors to anthropometric markers showing that the results of the parallel change analysis were most robust, consistent, and biologically plausible, compared to analyses of baseline diet and changes in weight (change-score), or changes in diet with changes in weight in a later time period (lagged-change) [93]. One explanation given by the authors pertains specifically to the physiology of weight-loss, as changes in diet or energy expenditure may be more relevant to the psychological adaptations during weight loss. Applied to glycaemic markers as the main study outcome, recent dietary changes are likely to be of more relevance for concurrent changes than changes in a later time period. Lastly, confounding might be more likely when using the baseline dietary intake compared to changes in dietary intake. The research group evaluating the parallel change approach had previously shown that baseline dietary factors were more strongly interrelated compared to dietary changes [94]. This suggests that individuals make changes in their diet relatively independently of other factors and thereby these parallel change analyses might be less prone to confounding. Exceptions were a positive correlation between changes in fruits and vegetables and negative correlation between changes in high-fat and low-fat dairy [94]. Thus, for dairy types, dependency of changes and associated confounding might be more of an issue.

Building on this study by Smith *et al.* (2015), a simulation study has affirmed that, parallel change analyses resulted in the most credible estimates compared to change-score or lagged change-change analyses [95]. However, this simulation study also highlighted that a causal relationship between baseline glycaemic marker levels and subsequent levels of these markers could potentially violate the assumptions of these analyses (i.e., no correlated error terms of regressor values), warranting careful conduction and interpretation of these analyses.

The associations of dairy types and glycaemic markers in the Fenland study differed between the parallel change analysis and the linear mixed models (**Chapter 6**). Only the positive association between low-fat milk intake and FPG was found using both analytical strategies. Some additional associations were found with the linear mixed model approach. Linear mixed models have the advantage that it accounts for individual variability by incorporating a random intercept and slope. The interaction term shows whether the association of dairy intake and glycaemic outcomes changes over the follow-up period. This allows for a nuanced examination of this relationship over time.

Moreover, parallel change analyses are affected by regression to the mean, as extreme values at either the highest or lowest intake levels tend to have the highest change in intake. This may contribute to the inconsistencies between baseline and parallel change analysis. As in our study, the study by Smith *et al.* (2015) showed that several associations from the parallel change analysis were in opposite direction compared to the change-score analysis [93]. For example, changes in high-fat dairy foods were associated with weight gain (0.08 kg, 95%CI 0.05; 0.11) while higher baseline high-fat dairy foods were non-significantly associated with less weight gain (-0.02 kg, 95%CI -0.03; 0.00). To reduce the impact of extreme values, we adjusted for baseline intake levels of the dairy type. Additionally, exclusion of outliers might reduce the impact of these extreme values, nevertheless, excluding outliers that represent meaningful changes might also introduce bias and result in associations that are overly weakened, or attenuated. Advantages and disadvantages of linear mixed models and parallel change-on-change analyses are summarized in **Table 1**.

Table 1. Advantages and disadvantages of two approaches to analyse associations of repeated exposure measures with repeated outcome measures

	Advantages	Disadvantages
Linear mixed models	• Suitable for correlated repeated measures • Allow for inclusion of random intercept and slope for each participant accounting for individual variability	• Assumption of linearity • Computational complexity • Difficult interpretability of interactions and random effects
Parallel change-on-change analyses	• Less assumption of linearity • Computational efficiency • Simple and intuitive interpretation	• Limited handling of correlated repeated measures • Ignoring individual variability • Reduced power • Affected by regression to the mean

Multiple testing

The studies described in this thesis encompass an extensive replication of standardized analysis in multiple prospective cohort studies of the general population. Examination of multiple dairy types with prediabetes and continuous glycaemic outcomes increase the likelihood of chance findings. At $\alpha = 0.05$, there is a 5% probability of rejecting a correct null hypothesis (type I error), i.e., no association between the exposure and outcome. We applied no statistical correction for multiple testing to reduce type I error (e.g., Bonferroni or Benjamini-Hochberg correction) for several reasons. Firstly, most exposures were correlated, and corrections may have resulted in a type II error [96, 97]. Secondly, statistical correction for multiple testing assumes that findings are primarily explained by chance, while this is inadequate in the context of empirical research [96]. Thirdly, many diet-disease hypotheses are evaluated based on data from these extensive observational cohorts, published in multiple manuscripts, making correction

of multiple testing inherently arbitrary and incomplete. Instead, significant findings in single studies provide targets for replication research, while penalization with multiple testing corrections would have disregarded that finding.

Meta-analytical approach

While individual studies provide insights into specific populations and settings, meta-analyses offer a comprehensive overview by synthesizing data from multiple studies. We conceive our meta-analyses (**Chapter 7**) to be of high-quality as we adhered to guidelines for protocol development, carefully investigated heterogeneity, and investigated risk of bias in both individual cohorts and the meta-analyses [98]. The selection of similar cohorts and standardization of analyses likely contributed to limited heterogeneity observed in the subsequent meta-analysis. Due to the small number of included cohorts and limited variation within cohorts, the exploration of moderation and subgroup analyses by age, sex and BMI was not possible. Nevertheless, in individual cohorts, these moderator analyses are of explorative nature, and meta-regression on aggregate data does not necessarily reflect true associations between participant-level characteristics and outcomes (i.e., ecological fallacy or aggregation bias) [99].

A limitation was that the variation of intake was rather constrained, primarily due to the extraction of relative risks with corresponding median values from intake range categories. This resulted in no zero-intake category for most dairy types, as well as a much lower maximum intake as compared to individual studies. For prospective studies with an adequate number of cases, extraction of estimates from additional intake categories with narrower intake intervals would enable fitting of a more precise dose-response curve. Fitting of a continuous nonlinear curve in each individual cohort would have been feasible but is constrained by the limitations of FFQs in accurately measuring precise individual intake levels.

In meta-analyses, pooling data from multiple studies often results in more precise estimates compared to individual studies, especially with considerable inconsistencies in the estimates from each study. This pitfall of meta-analyses is evident for the observed inverse linear association for ice cream intake with prediabetes. Although the single studies present inconsistent associations with extremely wide confidence intervals, the pooled estimate was much more precise due to larger sample sizes and the ability to account for between-study variability.

Compared to our approach, an individual participant data (IPD) meta-analyses could have potentially resulted in higher precision and provide flexibility in analytical techniques. Nevertheless, pooling of individual data was not possible due to limited possibilities for data sharing - three cohorts were analysed within a protected server and one cohort was not allowed to be transferred from the institute server. Furthermore, major differences in the study protocols, for example baseline measurement period and prediabetes

assessment methods, and differences in the number of items, food group definitions and validity of dietary assessment instruments, pose challenges for IPD meta-analyses as they may hamper data harmonization.

Suggestions for future research

Improving observational evidence

The results of this thesis have provided evidence base for the association of dairy intake and prediabetes risk in multiple cohort studies primarily in Western and predominantly affluent populations. Investigation of the associations between dairy intake and prediabetes risk in non-Western samples and samples with lower SEP are currently lacking and are needed in future. Studies utilizing multiple measurements of FPG and 2hPG to diagnose prediabetes are recommended to limit impact of short-term intraindividual variation [100, 101] and thereby lower non-differential misclassification in the outcome. Furthermore, as we did in the Rotterdam Study (**Chapter 3**) and the Fenland study (**Chapter 6**) studies should incorporate repeated measures of dairy intake to assess within-person variability and dietary changes over time. Moreover, further investigation is warranted to examine the specific within-person replacements of dairy within the diet and their relationship to prediabetes and T2D. This could involve specific assessments of replacement choices within the FFQ. Future prospective studies also need to investigate additional hypotheses that may clarify the potential of various biases including residual confounding and reverse causation. The sensitivity of associations to the presence of unmeasured confounders can be assessed in simulation studies [102]. With extensive repeated measures, it can be explored how changes in dairy consumption patterns - driven by factors such as weight and health goals, perceived risk, medical diagnoses, or psychosocial factors - may contribute to the associations observed between dairy intake and health outcomes. Factors that could contribute to individual variability in health effects associated with dairy intake could be explored in more detail, including population characteristics and background diet, genetic predisposition, lactose intolerance, gut microbiota composition and health factors such as obesity, insulin sensitivity and dyslipidaemia. Integrating data on epigenetics, inflammatory markers, circulating metabolites and gut microbiota data is needed to provide mechanistic evidence [80]. Repeated measures of intermediate markers allow for exploration of moderating or mediating mechanisms such as weight regulation, insulin sensitivity, inflammation, gut microbiota, blood pressure and lipid homeostasis with joint modelling or advanced mediation analyses. Time-varying versions of MGM to assess within-person auto-regressive and cross-lagged associations between dairy intake, health behaviours and intermediate markers could provide more insights into temporality of associations and the potential for reverse causation [56]. A holistic approach with integration of psychological factors such as health awareness, attitudes, coping mechanisms, personality (e.g., extraversion and neuroticism) and mental health could further advance understanding of the complex associations between diet and

disease. With these complex datasets, network models including MGM, clustering and machine learning techniques could be valuable tools for identifying and prioritizing confounding variables.

Innovative designs and statistical methods have the potential to enhance causal inference from observational data. Target trail emulation is a technique used to design and analyse observational studies to closely resemble a hypothetical RCT, thereby strengthening causal inference and reducing influence of confounders, by setting specific inclusion and exclusion criteria, treatment assignment rules, and outcome assignment protocols [103, 104]. Propensity score matching may be used to create comparable groups by matching participants on their probability of being exposed based on many covariates, which can then be related to health outcomes [105]. This mimics the random assignment of exposure in experimental studies, thus improving the validity of causal inferences in observational studies [106]. Stratification of samples who are less likely to have changed their dietary habits, for example by a low diabetes risk score, no desire to lose weight or no diagnosis, can help to create more comparable exposure groups, thereby addressing confounding.

Furthermore, an interesting area for future usage of network models is confirmatory hypothesis testing of certain edges with Bayesian Gaussian Graphical Models [107, 108]. This allows for testing of a priori expectations for the (independence) structure based on theory and clinical evidence. The hypothesis to be evaluated is that a set of partial correlations is stronger compared to another set of partial correlations. These hypotheses can be directly formulated from nutritional guidelines and prior meta-analyses. For example, most guidelines emphasize low-fat dairy types to lower total SFA intake. With Bayesian network models, it can be evaluated if partial correlations with cardiometabolic risk factors are larger for low-fat dairy than for high-fat dairy.

These innovative designs and statistical methods do not account for unmeasured confounding and measurement error, emphasizing the need for more accurate and detailed collection on dairy intake and dietary behaviours. Prospective cohorts should incorporate smartphone-based recalls with possible shorter recall periods reducing underreporting [109] or image-assisted dietary assessment to alleviate the burden of self-reporting. For the latter, food image recognition employing deep-learning approaches have been extensively developed, although not yet implemented in large scale research initiatives. Furthermore, advancements in objective monitoring of physical activity, sleep and stress, or ecological momentary assessment of health behaviours and contextual factors could provide a more accurate measurement and improve the precision of confounder adjustment. With sophisticated statistical techniques, such as regression calibration, these extensive methods could be applied to a subset of the population and used to reduce measurement error and other sources of bias in the exposure variables and covariates. By incorporating standardized (additional) measurements and validation

methods across various prospective cohorts, the variables across cohorts could be made more consistent. This could enhance the ability to compare and analyse data across different studies.

The need for an RCT

Observational evidence showing consistent dose-response trends with biological plausibility is sufficient to inform public health recommendations [110, 111]. Nevertheless, as concluded in prior chapters, to fully elucidate the causal effect of dairy intake on hyperglycaemia and prediabetes incidence, a well-designed RCT is needed. RCTs should focus on alignment of research protocols with observational studies to improve agreement in findings, including a focus on specific dairy types, appropriate selection of the population, and consistent outcome assessment [112].

Multiple large RCTs are warranted to determine effects of specific high-fat or regular versus low-fat dairy products (i.e., milk, yogurt, and cheese), substitution effects of each dairy product versus plant-based alternatives and other logical replacement options (i.e., fortified beverages, nuts, seeds, or other sources of protein and calcium) on metabolic health parameters, including insulin sensitivity and glucose metabolism. We propose that the control group should include a non-dairy consuming group, compared to a moderate dairy consumption group (i.e., 1 serving per day of a certain dairy product) and a dairy group consuming 2-3 servings per day of a certain dairy product in accordance with dietary guidelines. This would ensure sufficient exposure ranges and examination of a dose-response effect. The dairy types consumed should be as homogeneous as possible with considerable efforts done to measure and report compliance. As the effect of dairy varies in trials using isocaloric arms versus free-living non-energy restricted arms [18], a careful consideration between these two settings must be made, considering potential effects of higher energy intake, compensatory dietary changes, and generalizability of findings. Isocaloric weight maintenance diets are recommended for estimating effects of dairy independent of increased energy intake and body weight gain [113]. When comparing high- versus low-fat dairy types, matching energy- and macronutrient intake in intervention arms is needed to prevent compensatory effects, such as increased consumption of carbohydrate-rich foods, which could confound the results. Selection of healthy subjects is recommended to mitigate confounding by individuals' baseline glycaemic or lipid profiles and to gain insights into the role of dairy in maintaining metabolic healthy states and early prevention of T2D. To examine prediabetes incidence as outcome, a considerable follow-up is needed (e.g., at least 12 months, preferably ≥ 24 months). However, given practical and financial constraints, shorter trials with metabolic parameters will be more feasible. In that case, we advocate for a cross-over design lasting at least 16 weeks to capture potential long-term effects [23, 26]. The primary endpoint should be changes in glycaemic markers or insulin sensitivity to ensure appropriateness of study design and sample size calculation.

An assessment of the impact of high-fat compared to low-fat dairy products on energy balance, dietary choices and health behaviours is needed, providing insights into practically made substitutions. Furthermore, studies are needed to evaluate the hypothesis that high-fat dairy foods induce satiety, while low-fat options may lead to overconsumption of for example refined carbohydrates.

More research is needed to explore the heterogeneity of effects attributable to the composition and structure of different dairy types, including their specific SFA content, presence of vitamin K, different probiotic strains, as well as variations in treatment and fermentation methods. Detailed metabolic phenotyping of study participants is needed to understand glycaemic responses to the dietary intervention, with measures of insulin resistance [114, 115], body fat distribution and adipose tissue function [116], lactase persistence [14], and gene expression and regulation.

Future experimental studies are also needed to investigate biological pathways, including the insulinotropic effect of dairy foods, i.e., the repeated postprandial hyperinsulinemia due to habitual dairy intake, on insulin sensitivity [23]. This would require extensive repeated measurements of postprandial insulin responses in an RCT in a healthy population. Furthermore, experimental studies are needed on the effects of dairy intake on alterations in the gut microbiome and subsequent metabolite production, particularly impacting the synthesis of BCAA and SCFA such as butyrate, and on modulation of inflammatory pathways [14, 117].

Implications for public health

The findings of this thesis may offer some insights relevant for public health recommendations regarding the role of dairy in the early-stage prevention of T2D. Considering the large inconsistencies in associations of dairy with continuous glycaemic markers and with prediabetes risk, as well as low to moderate grading of evidence based on the meta-analysis (**Chapter 7**), the findings in this thesis are not sufficient to justify changes to dietary guidelines. Furthermore, no single food may reduce the risk of a disease interlinked to the overall dietary patterns, health behaviours, genetic, psychosocial, and environmental factors. Moreover, we only considered prediabetes as outcome, while dietary guidelines are based on evidence for multiple diseases outcomes and intermediate markers, often addressing T2D but not prediabetes specifically. Public health efforts to improve overall dietary quality and health behaviours should remain priority.

Our results for a nonlinear inverse association of total dairy and prediabetes support the current food-based dietary guidelines that promote the intake of 2-3 servings per day. This optimal consumption level of dairy provides the necessary nutrients with higher intakes resulting in potential overconsumption of calories, sugars and SFA. However, prominent researchers have challenged this recommendation, suggesting that an acceptable intake

of 0 to 2 servings per day suffices [118, 119]. This considers that the optimal intake will depend on overall diet quality as many of the nutrients found in dairy can also be obtained from a variety of other food sources, as well as environmental concerns.

Our findings do not support the recommendation to focus on low-fat dairy products nor further specification of an optimal distribution between low-fat and high-fat dairy types. This recommendation is mainly done to limit intake of total SFA to less than 10 E%, based on the causal link between SFA and plasma concentrations of LDL cholesterol, a key risk factor for atherosclerotic CVD [120]. The dietary fat quality determines LDL cholesterol, as evident from feeding trials, RCTs and mechanistic studies. The evidence base for substitution of foods containing SFA, trans fatty acids and cholesterol to foods containing monounsaturated FAs and polyunsaturated FAs to prevent or improve dyslipidaemia is substantial. Nevertheless, observational studies do not show harmful associations of high-fat dairy intake in relation to CVD incidence and mortality [121, 122], likely as dairy fat presents a complex mixture of SFAs with counteracting effects of other nutrients. Also, trial evidence does not support this recommendation. For example, a RCT comparing 0.5 whole milk for 3 weeks compared to skim milk in healthy individuals did not increase LDL cholesterol, while the whole milk diet had 14.4 E% from SFA compared to 11.3 E% in the skimmed milk diet [21]. Beneficial effects of dairy fat (i.e., specific fatty acids, milk fat globule membrane (MFGM) as well as compensating higher energy intake by lowering carbohydrate intake might underline this finding. The dietary recommendation to limit SFA in the diet without considering benefits of specific food sources including high-fat dairy might inadvertently lead to dietary patterns rich in refined starch and sugar. Therefore, dietary guidelines would benefit from shifting this focus to more effective food-based recommendations that emphasize the consumption of minimally processed foods.

We showed that high-fat cheese intake, mostly comprising hard cheeses in these Western cohorts, is moderately beneficially associated with prediabetes risk. This finding for high-fat cheese contradicts with many dietary guidelines including those in the Netherlands, emphasizing low-fat cheeses (i.e., in Dutch guidelines: 10+, 20+, 30+, cottage cheese, mozzarella and fresh goat cheese) [123, 124]. Prior evidence also shows moderately beneficial associations of total cheese with T2D, all-cause mortality, CVD mortality, and incident CVD, CHD and stroke, and null associations with cancer mortality, hypertension, and prostate cancer [1]. For low-fat cheese, drawing a firm conclusion is limited by low intake levels in each cohort and the high potential for reverse causation. Based on literature, evidence of more favourable associations of low-fat cheese with cardiometabolic outcomes compared to high-fat cheese is lacking [125]. If the beneficial associations for cheese intake are attributed to specific fatty acids or fat-soluble vitamins such as vitamin K2, advocating for low-fat cheese may not be advisable. The direct translation of the current evidence into dietary recommendations is limited as many complex associations of cheese intake and health outcomes have not been validated in

a sufficient number of well-designed studies, and the proposed underlying mechanisms are yet to be fully understood.

Our null findings for total yogurt intake in relation with prediabetes, irrespective of fat content, is not in line with current guidelines stating that yogurt intake results in a lower risk of T2D. In the beginning of 2024, the US Food and Drug Administration has qualified the health claim that eating at least 2 cups per week of yogurt may reduce the risk of T2D based on limited scientific evidence [126]. This claim applies to all types of yogurts with varying fat and sugar content, with effects relating to the protein and micronutrient content, as well as a higher diet quality associated with yogurt consumption. The long-term effects of this claim in the US are yet to be determined. Nonetheless, the focus of recommendations should be on plain yogurts, discouraging excessive consumption of yogurt with added sugar or with unhealthy components (e.g. cookies, refined cereals).

Moderately positive associations at low intake levels of ice cream indicate no detrimental role in prediabetes development when consumed in moderation. However, higher intake of ice cream is likely to lead to excessive consumption of sugar and fat, potentially elevating prediabetes risk. No specific recommendation for cream intake is made considering the low intake levels in observational cohort studies.

In the included cohorts, the average intake of calcium in the lowest intake category of total dairy was approximately 600-700 mg/day across all cohorts, falling below the Dutch recommended levels of 950 to 1200 mg/day depending on age and sex. The Dutch recommended levels are comparable with worldwide recommendations but are higher than in for example the UK (700 mg/day). Low calcium intake among people not consuming recommended levels of dairy foods is not directly an area of concern, as many other foods contain calcium and absorption is upregulated when dietary calcium intake is low. The main foundation for this recommendation is promotion of bone health, while the overall evidence does not support that high dairy diets relate to a reduction of fractures in general populations [118]. For bone health, public health efforts to promote physical activity, being outdoors, and overall dietary quality to ensure sufficient vitamin D synthesis might be more relevant. The intake of calcium, vitamin A and B2 is low among adults in the Netherlands, but there are no indications for concern regarding these low intakes from a public health perspective [127]. Follow-up research into nutritional status or prevalence of clinical symptoms due to low intakes is desirable.

Although guidelines for the prevention of chronic diseases have been well defined, changing diet is extremely challenging. Previous approaches focusing on modification of individual factors such as cognitive abilities (e.g. decision making based on risk and benefits), skills (e.g. nutrition literacy), and eating behaviours (i.e., snacking) have proven to be unsuccessful on a population level, given the continuing raise of obesity and cardiometabolic diseases. This approach mainly involved the provision of information on

healthy diets to the public, tailored to specific subgroups. Nevertheless, making healthy dietary choices is extremely difficult for individuals given that high levels of motivation, self-efficacy and behavioural capability is needed to withstand the many temptations in our obesogenic environment [128]. Thus, modifications of our sociocultural, physical, and economic environments are pivotal to achieve healthier food choices. This requires effective governmental policies. Currently, in the Netherlands, a 'consumption' tax on non-alcoholic SSBs was set in place in January 2024. As for April 2024, dairy drinks are exempt from this tax due to their beneficial protein content. However, implementing this tax could serve as a nudge to encourage consumers to choose dairy foods without added sugar. A nutrient profiling tax (e.g. sugar tax) targeting a wider range of unhealthy foods may further aid healthy food choice [129] and stimulate industries to reformulate their products to reduce sugar, fat, and salt content. Furthermore, marketing of sweets aimed at children is banned, as well as the legal tools for municipalities to ban new fast-food restaurants near schools or neighbourhood where there are too many fast-food restaurants already. The effects of these governmental policies are yet to be investigated.

The Dutch 'Nationaal Preventieakkoord' [National Prevention Agreement] aimed at improving Health in the Netherlands in 2040 includes several actions to improve healthy eating [130]. In supermarkets, these include increasing supply of products included in the Dutch dietary recommendations, improving product formulations, nudging and a food choice logo. Front-of-pack nutrition labelling has a high potential to effectively inform consumers about nutritional value and encourage manufactures to improve nutritional quality. In 2022, a front-of-pack nutrition labelling called the 'Nutri-Score' was implemented in the Netherlands to clearly communicate reliable information on product composition based on scientific evidence. Since January 2024, it is the official food choice logo in the Netherlands. The Nutri-Score is a food choice logo implemented by the government ranging from A (dark green, healthier) to E (dark orange, unhealthier) providing information on healthy composition of food items. The Nutri-Score is also implemented in other European countries. In online purchase experiments, the Nutri-Score assisted with identification of healthy products and increased purchase intentions for healthy products [131]. However, concerns have been raised as the Nutri-Score algorithm results in major discrepancies between the Nutri-Score and dietary guidelines [132]. The comparisons are made within a certain food group and positive ingredients (protein, fibre, vegetables, fruit, legumes, nuts, and certain oils) may counteract negative ingredients (energy, SFA, sugar and salt). Thereby, for example, a pizza with a cauliflower base, or vegetable chips with reduced fat are scored with an A and B respectively, while semi-skimmed milk, full-fat milk or reduced fat cheese are scored with a B, C and D, respectively. Thus, low-fat dairy foods and cheeses included in the Dutch dietary guidelines ['Schijf van Vijf' (Wheel of Five)] are scored worse than pizza and chips which are not recommended. These discrepancies between the Nutri-Score and dietary guidelines need to be resolved to avoid confusion among consumers. Providing clear explanations to the public on these scores has proven to be challenging considering

these discrepancies. The Netherlands Nutrition Centre ['Voedingscentrum'] therefore advises to only use the Nutri-Score only for products not in the dietary guidelines. Nevertheless, 80% of the products in supermarkets are not included in the Dutch dietary guidelines. Another concern is that food choice logos may be more useful for highly educated individuals with above average income [133], and therefore have limited impact on resolving health disparities. Perceived influence of industry involvement in formulations of guidelines and the NutriScore to benefit profit over public health considerations may affect public trust and credibility in the NutriScore. Independency of implementation and scientific evaluation, and clear communication towards the public can help to enhance the usability and effectiveness of the NutriScore. Overall, the NutriScore has a high potential to empower consumers to make healthier choices, given that future educational campaigns are effective. Moreover, priority should be on transforming the food supply to offer more healthier, sustainable, and affordable foods. This requires continuous innovations from the food industry with a focus on improving health rather than maximizing profits.

Regarding diagnosing prediabetes, there is a lack of capacity to screen and treat prediabetes in many countries including the Netherlands. In primary care setting, systematic case-findings among individuals as part of cardiovascular risk assessment or with high waist circumference could be further implemented. Guidelines for treatment of T2D may change over time to emphasize early detection and interventions for prediabetes, especially as awareness of the significance of early interventions for prevention of T2D and complications grows [134]. However, earlier screening may also result in unnecessary drug prescriptions with possible side effects and higher burden on healthcare systems. Also on an individual level, psychological and financial burden may hamper effective treatment.

Conclusion

To conclude, the findings presented in this thesis suggest that overall, dairy intake does not increase the risk of prediabetes. Specifically, moderately beneficial associations were observed for intake of total dairy, total cheese, and high-fat cheese and prediabetes risk. Milk and yogurt were not significantly associated with prediabetes, irrespective of their fat content. A critical appraisal of current literature is needed to inform the dietary guidelines regarding high-fat cheese intake and dairy consumption overall in relation to cardiometabolic health. Continued research efforts are warranted to elucidate the complex relationships between different dairy matrices on health outcomes, as well as to further investigate individual differences. Improvements in study design and analytical strategies to mitigate biases, especially those related to reverse causation, are also essential to advance the field of nutritional epidemiology. Public health efforts to improve overall dietary quality and health behaviours should remain priority.

References

1. Zhang, M., et al., *Cheese consumption and multiple health outcomes: an umbrella review and updated meta-analysis of prospective studies.* Adv Nutr, 2023. **14**(5): p. 1170-1186.
2. Gijsbers, L., et al., *Consumption of dairy foods and diabetes incidence: a dose-response meta-analysis of observational studies.* Am J Clin Nutr, 2016. **103**(4): p. 1111-24.
3. Chen, M., et al., *Dairy consumption and risk of type 2 diabetes: 3 cohorts of US adults and an updated meta-analysis.* BMC Med, 2014. **12**: p. 215.
4. Soedamah-Muthu, S.S. and J. de Goede, *Dairy Consumption and Cardiometabolic Diseases: Systematic Review and Updated Meta-Analyses of Prospective Cohort Studies.* Curr Nutr Rep, 2018. **7**(4): p. 171-182.
5. Jensen, C.F., M. Timofeeva, and G. Berg-Beckhoff, *Milk consumption and the risk of type 2 diabetes: A systematic review of Mendelian randomization studies.* Nutr Metab Cardiovasc Dis, 2023. **33**(7): p. 1316-1322.
6. Mendelian Randomization of Dairy Consumption Working, G. and C. consortium, *Dairy Intake and Body Composition and Cardiometabolic Traits among Adults: Mendelian Randomization Analysis of 182041 Individuals from 18 Studies.* Clin Chem, 2019. **65**(6): p. 751-760.
7. Fan, M., et al., *Dietary Protein Consumption and the Risk of Type 2 Diabetes: ADose-Response Meta-Analysis of Prospective Studies.* Nutrients, 2019. **11**(11): p. 2783.
8. Zhang, K., P. Bai, and Z. Deng, *Dose-Dependent Effect of Intake of Fermented Dairy Foods on the Risk of Diabetes: Results From a Meta-analysis.* Can J Diabetes, 2022. **46**(3): p. 307-312.
9. Alvarez-Bueno, C., et al., *Effects of Milk and Dairy Product Consumption on Type 2 Diabetes: Overview of Systematic Reviews and Meta-Analyses.* Adv Nutr, 2019. **10**(suppl_2): p. S154-S163.
10. *USDA Economic Research Service calculations using data from USDA, National Agricultural Statistics Service; USDA, Farm Service Agency; USDA, Foreign Agricultural Service; USDA, Agricultural Marketing Service; U.S. Department of Commerce, Bureau of the Census; and California Department of Food and Agriculture.*
11. *Centraal Bureau voor de Statistiek, Den haag/Heerlen.*
12. Margolis, K.L., et al., *A Diet High in Low-Fat Dairy Products Lowers Diabetes Risk in Postmenopausal Women12.* The Journal of Nutrition, 2011. **141**(11): p. 1969-1974.
13. Hruby, A., et al., *Associations of Dairy Intake with Incident Prediabetes or Diabetes in Middle-Aged Adults Vary by Both Dairy Type and Glycemic Status*. J Nutr, 2017. **147**(9): p. 1764-1775.
14. Luo, K., et al., *Variant of the lactase LCT gene explains association between milk intake and incident type 2 diabetes.* Nature Metabolism, 2024: p. 1-18.
15. Guo, J., et al., *The Impact of Dairy Products in the Development of Type 2 Diabetes: Where Does the Evidence Stand in 2019?* Adv Nutr, 2019. **10**(6): p. 1066-1075.
16. Hu, M.J., et al., *Effect of Cheese Intake on Cardiovascular Diseases and Cardiovascular Biomarkers.* Nutrients, 2022. **14**(14).
17. Benatar, J.R., K. Sidhu, and R.A. Stewart, *Effects of high and low fat dairy food on cardio-metabolic risk factors: a meta-analysis of randomized studies.* PLoS One, 2013. **8**(10): p. e76480.
18. O'Connor, S., et al., *Increased Dairy Product Intake Modifies Plasma Glucose Concentrations and Glycated Hemoglobin: A Systematic Review and Meta-Analysis of Randomized Controlled Trials.* Adv Nutr, 2019. **10**(2): p. 262-279.
19. Sochol, K.M., et al., *The Effects of Dairy Intake on Insulin Resistance: A Systematic Review and Meta-Analysis of Randomized Clinical Trials.* Nutrients, 2019. **11**(9).
20. Kiesswetter, E., et al., *Effects of Dairy Intake on Markers of Cardiometabolic Health in Adults: A Systematic Review with Network Meta-Analysis.* Adv Nutr, 2023. **14**(3): p. 438-450.
21. Engel, S., M. Elhauge, and T. Tholstrup, *Effect of whole milk compared with skimmed milk on fasting blood lipids in healthy adults: a 3-week randomized crossover study.* Eur J Clin Nutr, 2018. **72**(2): p. 249-254.
22. Eelderink, C., et al., *The effect of high compared with low dairy consumption on glucose metabolism, insulin sensitivity, and metabolic flexibility in overweight adults: a randomized crossover trial.* Am J Clin Nutr, 2019. **109**(6): p. 1555-1568.

23. Schmidt, K.A., et al., *The impact of diets rich in low-fat or full-fat dairy on glucose tolerance and its determinants: a randomized controlled trial.* Am J Clin Nutr, 2021. **113**(3): p. 534-547.
24. Nilsen, R., et al., *Effect of a high intake of cheese on cholesterol and metabolic syndrome: results of a randomized trial.* Food Nutr Res, 2015. **59**: p. 27651.
25. Raziani, F., et al., *High intake of regular-fat cheese compared with reduced-fat cheese does not affect LDL cholesterol or risk markers of the metabolic syndrome: a randomized controlled trial.* Am J Clin Nutr, 2016. **104**(4): p. 973-981.
26. Turner, K.M., J.B. Keogh, and P.M. Clifton, *Dairy consumption and insulin sensitivity: a systematic review of short- and long-term intervention studies.* Nutr Metab Cardiovasc Dis, 2015. **25**(1): p. 3-8.
27. Onvani, S., et al., *Dairy products, satiety and food intake: A meta-analysis of clinical trials.* Clin Nutr, 2017. **36**(2): p. 389-398.
28. Chen, M., et al., *Effects of dairy intake on body weight and fat: a meta-analysis of randomized controlled trials.* Am J Clin Nutr, 2012. **96**(4): p. 735-47.
29. Pradeilles, R., et al., *Effect of Isoenergetic Substitution of Cheese with Other Dairy Products on Blood Lipid Markers in the Fasted and Postprandial State: An Updated and Extended Systematic Review and Meta-Analysis of Randomized Controlled Trials in Adults.* Adv Nutr, 2023. **14**(6): p. 1579-1595.
30. Aslam, H., et al., *The effects of dairy and dairy derivatives on the gut microbiota: a systematic literature review.* Gut Microbes, 2020. **12**(1): p. 1799533.
31. JanssenDuijghuijsen, L., et al., *Changes in gut microbiota and lactose intolerance symptoms before and after daily lactose supplementation in individuals with the lactase nonpersistent genotype.* Am J Clin Nutr, 2024. **119**(3): p. 702-710.
32. Kim, D.H., et al., *Kefir alleviates obesity and hepatic steatosis in high-fat diet-fed mice by modulation of gut microbiota and mycobiota: targeted and untargeted community analysis with correlation of biomarkers.* J Nutr Biochem, 2017. **44**: p. 35-43.
33. Mozaffarian, D. and J.H.Y. Wu, *Flavonoids, Dairy Foods, and Cardiovascular and Metabolic Health: A Review of Emerging Biologic Pathways.* Circ Res, 2018. **122**(2): p. 369-384.
34. Nikbakht, E., et al., *Effect of probiotics and synbiotics on blood glucose: a systematic review and meta-analysis of controlled trials.* Eur J Nutr, 2018. **57**(1): p. 95-106.
35. Huppertz, T., et al., *Dairy Matrix Effects: Physicochemical Properties Underlying a Multifaceted Paradigm.* Nutrients, 2024. **16**(7): p. 943.
36. Christensen, R., et al., *Effect of calcium from dairy and dietary supplements on faecal fat excretion: a meta-analysis of randomized controlled trials.* Obes Rev, 2009. **10**(4): p. 475-86.
37. Jacobsen, R., et al., *Effect of short-term high dietary calcium intake on 24-h energy expenditure, fat oxidation, and fecal fat excretion.* Int J Obes (Lond), 2005. **29**(3): p. 292-301.
38. Hussein, A.G., et al., *Vitamin K(2) alleviates type 2 diabetes in rats by induction of osteocalcin gene expression.* Nutrition, 2018. **47**: p. 33-38.
39. Iwamoto, J., et al., *Vitamin K(2) prevents hyperglycemia and cancellous osteopenia in rats with streptozotocin-induced type 1 diabetes.* Calcif Tissue Int, 2011. **88**(2): p. 162-8.
40. Rahimi Sakak, F., et al., *Glycemic control improvement in individuals with type 2 diabetes with vitamin K(2) supplementation: a randomized controlled trial.* Eur J Nutr, 2021. **60**(5): p. 2495-2506.
41. Dunstan, D.W., et al., *The Australian Diabetes, Obesity and Lifestyle Study (AusDiab)--methods and response rates.* Diabetes Res Clin Pract, 2002. **57**(2): p. 119-29.
42. Maguire, E.R., et al., *Does exposure to the food environment differ by socioeconomic position? Comparing area-based and person-centred metrics in the Fenland study, UK.* International Journal of Health Geographics, 2017. **16**(1): p. 33.
43. Rutters, F., et al., *Cohort Profile: The Hoorn Studies.* Int J Epidemiol, 2018. **47**(2): p. 396-396j.
44. Klijs, B., et al., *Representativeness of the LifeLines Cohort Study.* PLoS One, 2015. **10**(9): p. e0137203.
45. Nohr, E.A., et al., *Does low participation in cohort studies induce bias?* Epidemiology, 2006. **17**(4): p. 413-418.
46. Nohr, E.A. and Z. Liew, *How to investigate and adjust for selection bias in cohort studies.* Acta obstetricia et gynecologica Scandinavica, 2018. **97**(4): p. 407-416.
47. Boshuizen, H.C., et al., *Non-response in a survey of cardiovascular risk factors in the Dutch population: determinants and resulting biases.* Public health, 2006. **120**(4): p. 297-308.

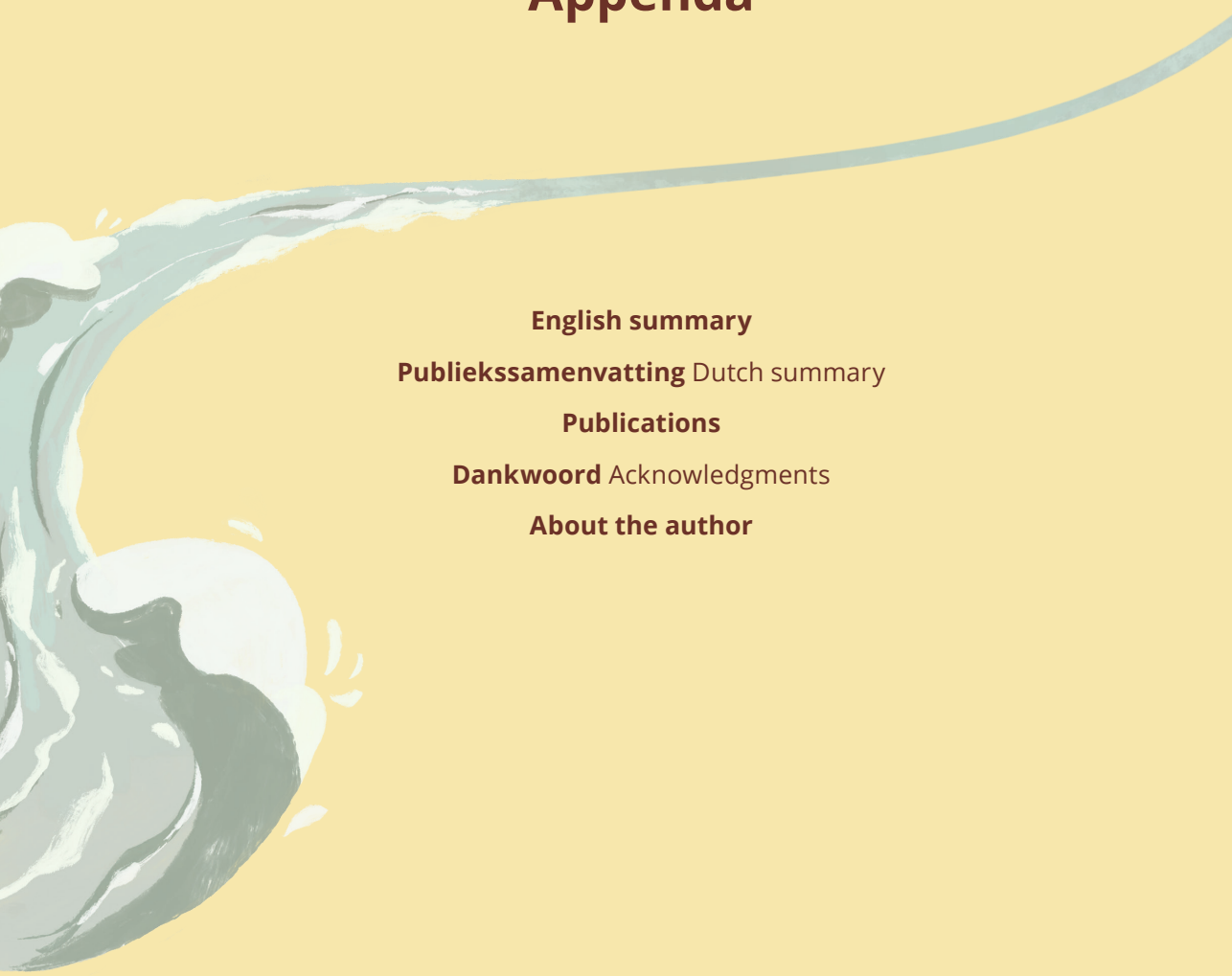
48. Kristman, V., M. Manno, and P. Côté, *Loss to follow-up in cohort studies: how much is too much?* Eur J Epidemiol, 2004. **19**(8): p. 751-60.
49. Bhavadharini, B., et al., *Association of dairy consumption with metabolic syndrome, hypertension and diabetes in 147 812 individuals from 21 countries.* BMJ Open Diabetes Res Care, 2020. **8**(1): p. e000826.
50. Willett, W.C., G.R. Howe, and L.H. Kushi, *Adjustment for total energy intake in epidemiologic studies.* Am J Clin Nutr, 1997. **65**(4 Suppl): p. 1220S-1228S; discussion 1229S-1231S.
51. Greenland, S. and N. Pearce, *Statistical foundations for model-based adjustments.* Annu Rev Public Health, 2015. **36**: p. 89-108.
52. VanderWeele, T.J., *Principles of confounder selection.* Eur J Epidemiol, 2019. **34**(3): p. 211-219.
53. Goossens, G.H., *The role of adipose tissue dysfunction in the pathogenesis of obesity-related insulin resistance.* Physiol Behav, 2008. **94**(2): p. 206-18.
54. Prince, S.A., et al., *A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review.* Int J Behav Nutr Phys Act, 2008. **5**(1): p. 56.
55. Epskamp, S., D. Borsboom, and E.I. Fried, *Estimating psychological networks and their accuracy: A tutorial paper.* Behav Res Methods, 2018. **50**(1): p. 195-212.
56. Haslbeck, J.M.B. and L.J. Waldorp, *mgm: Estimating Time-Varying Mixed Graphical Models in High-Dimensional Data.* Journal of Statistical Software, 2020. **93**(8): p. 1 - 46.
57. Dalege, J., et al., *Toward a formalized account of attitudes: The Causal Attitude Network (CAN) model.* Psychol Rev, 2016. **123**(1): p. 2-22.
58. Borsboom, D. and A.O. Cramer, *Network analysis: an integrative approach to the structure of psychopathology.* Annu Rev Clin Psychol, 2013. **9**: p. 91-121.
59. Epskamp, S. and E.I. Fried, *A tutorial on regularized partial correlation networks.* Psychol Methods, 2018. **23**(4): p. 617-634.
60. Haslbeck, J. and L.J. Waldorp, *How well do network models predict observations? On the importance of predictability in network models.* Behavior research methods, 2018. **50**(2): p. 853-861.
61. Willett, W. and E. Lenart, *Reproducibility and validity of food frequency questionnaires.* Nutritional epidemiology, 2013. **3**: p. 96-141.
62. Schwingshackl, L., et al., *Perspective: NutriGrade: A Scoring System to Assess and Judge the Meta-Evidence of Randomized Controlled Trials and Cohort Studies in Nutrition Research.* Adv Nutr, 2016. **7**(6): p. 994-1004.
63. Flegal, K.M., P.M. Keyl, and F.J. Nieto, *Differential misclassification arising from nondifferential errors in exposure measurement.* American journal of epidemiology, 1991. **134**(10): p. 1233-1246.
64. Trijsburg, L., 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 nutrition, 2017. **20**(4): p. 598-607.
65. Hebert, J.R., et al., *Gender Differences in Social Desirability and Social Approval Bias in Dietary Self-report.* American Journal of Epidemiology, 1997. **146**(12): p. 1046-1055.
66. Black, A.E., *Critical evaluation of energy intake using the Goldberg cut-off for energy intake:basal metabolic rate. A practical guide to its calculation, use and limitations.* Int J Obes Relat Metab Disord, 2000. **24**(9): p. 1119-30.
67. Schofield, W.N., *Predicting basal metabolic rate, new standards and review of previous work.* Hum Nutr Clin Nutr, 1985. **39 Suppl 1**: p. 5-41.
68. Tooze, J.A., et al., *The accuracy of the Goldberg method for classifying misreporters of energy intake on a food frequency questionnaire and 24-h recalls: comparison with doubly labeled water.* Eur J Clin Nutr, 2012. **66**(5): p. 569-76.
69. Jenkins, B., J.A. West, and A. Koulman, *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): p. 2425-44.
70. Jaudszus, A., et al., *trans Palmitoleic acid arises endogenously from dietary vaccenic acid.* Am J Clin Nutr, 2014. **99**(3): p. 431-5.
71. Imamura, F., 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. **15**(10): p. e1002670.

72. Wolk, A., M. Furuheim, and B. Vessby, *Fatty acid composition of adipose tissue and serum lipids are valid biological markers of dairy fat intake in men.* J Nutr, 2001. **131**(3): p. 828-33.
73. Brevik, A., et al., *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): p. 1417-22.
74. Wolk, A., et al., *Evaluation of a biological marker of dairy fat intake.* Am J Clin Nutr, 1998. **68**(2): p. 291-5.
75. Abdullah, M.M., 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): p. 435-44.
76. Golley, R.K. and G.A. Hendrie, *Evaluation of the relative concentration of serum fatty acids C14: 0, C15: 0 and C17: 0 as markers of children's dairy fat intake.* Annals of nutrition and metabolism, 2014. **65**(4): p. 310-316.
77. O'Connor, S., et al., *Increased dairy product intake alters serum metabolite profiles in subjects at risk of developing type 2 diabetes.* Molecular nutrition & food research, 2019. **63**(19): p. 1900126.
78. Vissers, L.E., et al., *Consumption of a diet high in dairy leads to higher 15: 0 in cholesteryl esters of healthy people when compared to diets high in meat and grain.* Nutrition, Metabolism and Cardiovascular Diseases, 2020. **30**(5): p. 804-809.
79. Pranger, I.G., et al., *Fatty acids as biomarkers of total dairy and dairy fat intakes: a systematic review and meta-analysis.* Nutr Rev, 2019. **77**(1): p. 46-63.
80. Trichia, E., et al., *Plasma Metabolites Related to the Consumption of Different Types of Dairy Products and Their Association with New-Onset Type 2 Diabetes: Analyses in the Fenland and EPIC-Norfolk Studies, United Kingdom.* Mol Nutr Food Res, 2024. **68**(1): p. e2300154.
81. Pranger, I.G., et al., *Fatty acids as biomarkers of total dairy and dairy fat intakes: a systematic review and meta-analysis.* Nutrition reviews, 2019. **77**(1): p. 46-63.
82. Huang, L., et al., *Circulating Saturated Fatty Acids and Incident Type 2 Diabetes: A Systematic Review and Meta-Analysis.* Nutrients, 2019. **11**(5): p. 998.
83. Drouin-Chartier, J.P., et al., *Dairy consumption, plasma metabolites, and risk of type 2 diabetes.* Am J Clin Nutr, 2021. **114**(1): p. 163-174.
84. Mozaffarian, D., *Dairy foods and type 2 diabetes: profiling our metabolites and health.* Am J Clin Nutr, 2021. **114**(1): p. 5-6.
85. Sellem, L., et al., *Can individual fatty acids be used as functional biomarkers of dairy fat consumption in relation to cardiometabolic health? A narrative review.* British Journal of Nutrition, 2022. **128**(12): p. 2373-2386.
86. Piller, C., *Dubious diagnosis.* Science, 2019. **363**(6431): p. 1026-1031.
87. Echouffo-Tcheugui, J.B. and E. Selvin, *Prediabetes and what it means: the epidemiological evidence.* Annual review of public health, 2021. **42**: p. 59-77.
88. Rooney, M.R., et al., *Global Prevalence of Prediabetes.* Diabetes Care, 2023: p. dc222376.
89. American Diabetes, A., *2. Classification and Diagnosis of Diabetes.* Diabetes Care, 2017. **40**(Suppl 1): p. S11-S24.
90. Laursen, A.S.D., et al., *Substitutions of dairy product intake and risk of stroke: a Danish cohort study.* Eur J Epidemiol, 2018. **33**(2): p. 201-212.
91. Ibsen, D.B., et al., *Food substitution models for nutritional epidemiology.* Am J Clin Nutr, 2021. **113**(2): p. 294-303.
92. Tomova, G.D., M.S. Gilthorpe, and P.W. Tennant, *Theory and performance of substitution models for estimating relative causal effects in nutritional epidemiology.* Am J Clin Nutr, 2022. **116**(5): p. 1379-1388.
93. Smith, J.D., et al., *A Comparison of Different Methods for Evaluating Diet, Physical Activity, and Long-Term Weight Gain in 3 Prospective Cohort Studies.* J Nutr, 2015. **145**(11): p. 2527-34.
94. Mozaffarian, D., et al., *Changes in diet and lifestyle and long-term weight gain in women and men.* N Engl J Med, 2011. **364**(25): p. 2392-404.
95. Tang, D., et al., *Change analysis for intermediate disease markers in nutritional epidemiology: a causal inference perspective.* BMC Med Res Methodol, 2024. **24**(1): p. 49.
96. Rothman, K.J., *No adjustments are needed for multiple comparisons.* Epidemiology, 1990. **1**(1): p. 43-6.

97. Groenwold, R.H., et al., *Multiple testing: when is many too much?* 2021, Oxford University Press. p. E11-E14.
98. Tobias, D.K., et al., *A Primer on Systematic Review and Meta-analysis in Diabetes Research.* Diabetes Care, 2023. **46**(11): p. 1882-1893.
99. Geissbuhler, M., et al., *Most published meta-regression analyses based on aggregate data suffer from methodological pitfalls: a meta-epidemiological study.* BMC Med Res Methodol, 2021. **21**(1): p. 123.
100. Kuzuya, T., et al., *Report of the Committee on the classification and diagnostic criteria of diabetes mellitus.* Diabetes Res Clin Pract, 2002. **55**(1): p. 65-85.
101. Selvin, E., et al., *Short-term variability in measures of glycemia and implications for the classification of diabetes.* Arch Intern Med, 2007. **167**(14): p. 1545-51.
102. Groenwold, R.H.H., et al., *Sensitivity analyses to estimate the potential impact of unmeasured confounding in causal research.* International Journal of Epidemiology, 2009. **39**(1): p. 107-117.
103. Hernán, M.A. and J.M. Robins, *Using big data to emulate a target trial when a randomized trial is not available.* American journal of epidemiology, 2016. **183**(8): p. 758-764.
104. Katsoulis, M., et al., *On the estimation of the effect of weight change on a health outcome using observational data, by utilising the target trial emulation framework.* Int J Obes (Lond), 2023. **47**(12): p. 1309-1317.
105. Ali, M.S., R.H. Groenwold, and O.H. Klungel, *Best (but oft-forgotten) practices: propensity score methods in clinical nutrition research.* Am J Clin Nutr, 2016. **104**(2): p. 247-58.
106. Garcia-Huidobro, D. and J. Michael Oakes, *Squeezing observational data for better causal inference: Methods and examples for prevention research.* Int J Psychol, 2017. **52**(2): p. 96-105.
107. Williams, D.R. and J. Mulder, *Bayesian hypothesis testing for Gaussian graphical models: Conditional independence and order constraints.* Journal of Mathematical Psychology, 2020. **99**: p. 102441.
108. Williams, D.R., *Bayesian estimation for Gaussian graphical models: Structure learning, predictability, and network comparisons.* Multivariate Behavioral Research, 2021. **56**(2): p. 336-352.
109. Lucassen, D.A., et al., *Validation of the smartphone-based dietary assessment tool "Traqq" for assessing actual dietary intake by repeated 2-h recalls in adults: comparison with 24-h recalls and urinary biomarkers.* The American Journal of Clinical Nutrition, 2023. **117**(6): p. 1278-1287.
110. Rothman, K.J. and S. Greenland, *Causation and causal inference in epidemiology.* American journal of public health, 2005. **95**(S1): p. S144-S150.
111. Dutch Health Council (Gezondheidsraad), *Dutch dietary guidelines 2015 (Richtlijnen goede voeding 2015). Publication nr. 2015/24. ISBN 978-94-6281-089-1. The Hague.* 2015.
112. Schwingshackl, L., et al., *Evaluating agreement between bodies of evidence from randomised controlled trials and cohort studies in nutrition research: meta-epidemiological study.* BMJ, 2021. **374**: p. n1864.
113. Geng, T., L. Qi, and T. Huang, *Effects of dairy products consumption on body weight and body composition among adults: an updated meta-analysis of 37 randomized control trials.* Molecular nutrition & food research, 2018. **62**(1): p. 1700410.
114. Trouwborst, I., et al., *Cardiometabolic health improvements upon dietary intervention are driven by tissue-specific insulin resistance phenotype: a precision nutrition trial.* Cell Metabolism, 2023. **35**(1): p. 71-83. e5.
115. Blanco-Rojo, R., et al., *The insulin resistance phenotype (muscle or liver) interacts with the type of diet to determine changes in disposition index after 2 years of intervention: the CORDIOPREV-DIAB randomised clinical trial.* Diabetologia, 2016. **59**(1): p. 67-76.
116. Goossens, G.H., *The Metabolic Phenotype in Obesity: Fat Mass, Body Fat Distribution, and Adipose Tissue Function.* Obes Facts, 2017. **10**(3): p. 207-215.
117. Brown, A.W., et al., *Science dialogue mapping of knowledge and knowledge gaps related to the effects of dairy intake on human cardiovascular health and disease.* Crit Rev Food Sci Nutr, 2021. **61**(2): p. 179-195.
118. Willett, W.C. and D.S. Ludwig, *Milk and Health.* N Engl J Med, 2020. **382**(7): p. 644-654.
119. Willett, W., et al., *Food in the Anthropocene: the EAT-Lancet Commission on healthy diets from sustainable food systems.* Lancet, 2019. **393**(10170): p. 447-492.

120. Christensen, J.J., et al., *Dietary fat quality, plasma atherogenic lipoproteins, and atherosclerotic cardiovascular disease: An overview of the rationale for dietary recommendations for fat intake.* Atherosclerosis, 2024. **389**: p. 117433.
121. Guo, J., et al., *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. **32**(4): p. 269-287.
122. Bechthold, A., et al., *Food groups and risk of coronary heart disease, stroke and heart failure: A systematic review and dose-response meta-analysis of prospective studies.* Crit Rev Food Sci Nutr, 2019. **59**(7): p. 1071-1090.
123. Comerford, K.B., et al., *Global Review of Dairy Recommendations in Food-Based Dietary Guidelines.* Front Nutr, 2021. **8**: p. 671999.
124. Soedamah-Muthu, S. and J. Guo, *Dairy consumption and cardiometabolic diseases: Evidence from prospective studies. In D. I. Givens (Ed.), Milk and dairy foods: Their functionality in human health and disease (pp. 1-28). Academic Press.* 2020.
125. Drouin-Chartier, J.P., et al., *Systematic Review of the Association between Dairy Product Consumption and Risk of Cardiovascular-Related Clinical Outcomes.* Adv Nutr, 2016. **7**(6): p. 1026-1040.
126. Lordan, R., *A new era for food in health? The FDA announces a qualified health claim for yogurt intake and type II diabetes mellitus risk reduction.* Diabetes & Metabolic Syndrome: Clinical Research & Reviews, 2024. **18**(4): p. 103006.
127. van Rossum, C., et al., *The diet of the Dutch. Results of the Dutch National Food Consumption Survey 2019-2021 on food consumption and evaluation with dietary guidelines.* 2023.
128. Story, M., et al., *Creating healthy food and eating environments: policy and environmental approaches.* Annu Rev Public Health, 2008. **29**: p. 253-72.
129. Eykelenboom, M., et al., *The effects of a sugar-sweetened beverage tax and a nutrient profiling tax based on Nutri-Score on consumer food purchases in a virtual supermarket: a randomised controlled trial.* Public Health Nutr, 2022. **25**(4): p. 1105-1117.
130. *Nationaal Preventieakkoord. November 2018. Uitgave van het ministerie van Volksgezondheid, Welzijn en Sport. https://www.vzinfo.nl/preventie/preventieakkoord.*
131. De Temmerman, J., et al., *The impact of the Nutri-Score nutrition label on perceived healthiness and purchase intentions.* Appetite, 2021. **157**: p. 104995.
132. Gerritsen, J., H. Verhagen, and S. Peters. *The Updated Algorithm of Front-of-Pack Label Nutri-Score Is Not in Line with Dutch Food-Based Dietary Guidelines: Results of Calculations with Dutch Food Composition Database*. in *Proceedings*. 2024. MDPI.
133. van Duist, L., *Kennis en houding ten aanzien van het Nutri-Score logo onder consumenten: Onderzoeksrapportage. Dienst Publiek en Communicatie. Ministerie van Algemene Zaken.* 2022.
134. Jonas, D.E., et al., *Screening for prediabetes and type 2 diabetes: updated evidence report and systematic review for the US Preventive Services Task Force.* Jama, 2021. **326**(8): p. 744-760.

Appenda



English summary

Dairy foods are widely consumed and recommended in guidelines worldwide as part of a healthy diet. Several dairy foods can be important for maintaining cardiometabolic health, being rich in protein, odd-chained fatty acids, calcium, magnesium, potassium, vitamins A, D, B2 and B12. However, dairy foods can also be relatively high in saturated fat, sodium, and sugar which may be detrimental for cardiometabolic health. The Dutch Health Council advises daily consumption of dairy products. This advice is supported by extensive meta-analyses, which reported associations of higher total dairy intake—especially low-fat dairy and yogurt—with a lower risk of type 2 diabetes.

Prediabetes is an intermediate stage between normoglycemia and type 2 diabetes, marked by insulin resistance and beta-cell dysfunction. Plasma glucose levels in prediabetes exceed the normal range but remain below the threshold for diabetes diagnosis. The prevalence of prediabetes is rising rapidly, particularly among people with obesity and older adults. Up to 50% of individuals with prediabetes may progress to diabetes within five years, and they are already at increased risk for microvascular and macrovascular complications. This growing prevalence highlights the need to identify modifiable risk factors to prevent or reverse prediabetes. Given its wide reach, even small shifts in risk could have a significant public health impact. Lifestyle modification is the recommended strategy for preventing and managing prediabetes, with its effectiveness demonstrated in many randomized controlled trials. Additionally, individuals with prediabetes may reverse to normoglycemia through dietary and lifestyle changes.

Although the link between dairy intake and type 2 diabetes is well-researched, its potential role in preventing earlier stages, including prediabetes, remains less explored. Understanding this relationship could offer valuable insights for early intervention. However, establishing clear associations between diet and specific health outcomes is particularly challenging due to individual variability in dietary responses, the complex and multifactorial nature of diseases, and the long latency periods involved.

Observational studies provide valuable insights into potential associations between diet and long-term health outcomes in large populations. Among these, prospective cohort studies are regarded as the highest quality of evidence, as they allow for the measurement of exposures before disease onset. When observational evidence is consistent and well-executed, with adequate adjustment for confounders, and supported by mechanistic studies that demonstrate biological plausibility, these associations provide a foundation for inferring causal relationships.

Aims

In this thesis, we aim to study the relation between the intake of total dairy and various dairy types and prediabetes in prospective cohort studies including general populations. We

focused on assessing this relation in prospective observational cohort studies in adults with normoglycaemia at baseline from the Netherlands, Australia and the United Kingdom. Each cohort collected continuous glycaemic measures at baseline and follow-up assessments, used a validated food frequency questionnaire to measure intake of a wide range of dairy foods and included a comprehensive set of sociodemographic and health risk factors.

We followed a standardized analysis plan to ensure a robust evidence base, providing high-quality input for synthesizing all available evidence from prospective cohort studies in a meta-analysis. We employed regression analysis adjusting for various confounding factors, including sociodemographic factors, health behaviours, cardiometabolic risk factors, to offer insights in the independent association of dairy intake and prediabetes risk. Additionally, we utilized a novel approach by applying network models to capture the holistic interrelationships between dietary characteristics, sociodemographic factors, health behaviours, cardiometabolic risk factors and prediabetes. This approach aimed to determine whether the heterogeneity in associations between dairy intake and prediabetes risk could be (partly) explained by covariation of health behaviours and food intake across the different types of dairy foods.

Finally, we conducted a systematic review and meta-analysis to identify, summarize, and evaluate all available evidence on the associations between any dairy intake and continuous glycaemic markers and the incidence of prediabetes. We assessed the quality of evidence for each included study and evaluated the confidence in the findings derived from the meta-analysis.

Main findings

Most types of dairy were not associated with the risk of developing prediabetes in the five analysed prospective cohort studies. However, the associations between specific dairy types and prediabetes risk varied across the individual cohorts. In the Dutch Hoorn Studies **(Chapter 2)**, high-fat fermented dairy and cheese, particularly high-fat cheese, were associated with a lower risk of prediabetes, while no significant associations were found when substituting high-fat cheese with other dairy types. In the Dutch Rotterdam Study **(Chapter 3)**, higher intakes of high-fat milk and high-fat yogurt were associated with a lower prediabetes risk and longitudinal insulin resistance, with weaker positive associations for low-fat dairy. In the Australian Diabetes, Obesity, and Lifestyle Study (AusDiab) (**Chapter 4**), high-fat dairy, high-fat milk and total cheese were associated with a lower prediabetes risk, with a nonlinear association for low-fat milk. In the Dutch Lifelines study (**Chapter 5**), the highest intakes of plain and low-fat milk intake were associated with a higher risk of prediabetes, though no dose-response relationships were observed. Lastly, in the UK Fenland study (**Chapter 6**), high-fat dairy was associated with a higher prediabetes risk, while low-fat milk intake was associated with a lower prediabetes risk. The opposite was shown in this study for changes in dairy intake over follow-up. Increased high-fat milk intake during follow-up was associated with lower fasting plasma glucose and a lower risk of progressing to prediabetes or type 2 diabetes.

In the Lifelines study (**Chapter 5**), we found that reverse causation may influence the results, even with a prospective design. In nutritional epidemiological research, reverse causation occurs when individuals alter their eating habits due to awareness of disease risk. In our study, awareness of diabetes risk, indicated by a diabetes risk score or a desire to lose weight, was linked to a higher intake of low-fat dairy and a lower intake of high-fat dairy. This suggests that individuals without disease risk (e.g., those without obesity) might choose their diets hedonically rather than based on health considerations, potentially leading to a lower prediabetes risk associated with dairy types not included in dietary recommendations, such as high-fat dairy and ice cream.

In the Lifelines and Fenland study (**Chapters 6 and 7)**, the networks of dairy intake, dietary patterns, sociodemographic factors, health behaviours, and cardiometabolic risk factors, including prediabetes, revealed clusters of health-conscious behaviours, such as higher intake of vegetables, fruit, and physical activity, alongside distinct preferences for high-fat or low-fat dairy. Both networks also revealed a cluster of energy-dense food groups and highlighted the bridging role of waist circumference between prediabetes and sociodemographic characteristics. Nevertheless, the findings from regression and network analyses showed limited alignment. While some high-fat and low-fat dairy types were similarly positioned within the networks, their associations with prediabetes risk pointed in opposing directions in regression analyses due to inverse correlations between their intake levels. In contrast, similar associations in regression analyses corresponded to distinct placements of certain dairy types in the networks. Furthermore, dairy types that showed associations in regression analyses were not necessarily reflected by a more prominent role in the networks. Thus, the findings suggest that covariation of dietary and health behaviours, as captured by the networks, does not fully explain the heterogeneity in associations between dairy intake and prediabetes risk. The findings point to the complexity of these relationships, emphasizing the need for a multifaceted approach to fully understand these nuanced connections.

The systematic review of 14 prospective cohort studies (**Chapter 7**) revealed a mix of inverse, positive, and non-significant associations between dairy intake and continuous glycaemic measurements. Thus, the existing body of evidence remains inconclusive regarding the relationship between subtypes of dairy and glycaemic outcomes. In the meta-analysis of six studies across nine cohorts, total dairy intake was nonlinearly inversely associated with prediabetes risk, with the lowest risk of prediabetes observed at approximately 3.4 servings per day, although no clear trends were found for high-fat versus low-fat dairy. Both total and high-fat cheese showed nonlinear inverse associations, with optimal intake levels around 2.1 servings per day; however, intakes exceeding 4 servings per day were associated with positive risk. These results are somewhat consistent with evidence for type 2 diabetes, though previous studies have not established clear dose-response relationships. Potential beneficial effects on hyperglycaemia may stem from the presence of specific saturated fatty acids, the milk fat globule membrane, calcium, and vitamin K within the

cheese matrix. Consequently, dietary guidelines that focus solely on low-fat cheese may lack adequate support from the current literature. Ice cream intake exhibited an inverse linear association with prediabetes risk, consistent with findings related to type 2 diabetes; however caution is warranted due to low median intake levels, difficulty in accurately assessing intake due to high seasonal variation, and a high potential for reverse causation. No associations were found for total, high-fat, or low-fat milk and cream, consistent with previous literature. Additionally, despite earlier studies suggesting that yogurt intake might reduce type 2 diabetes risk, our meta-analysis found no associations between yogurt intake and prediabetes. The previously observed inverse associations with type 2 diabetes may be attributed to low intake levels and health behaviours associated with yogurt consumption. Inconsistencies in findings may also arise from the variety of yogurt types consumed, which can differ in probiotic strains and sugar content.

The findings provide evidence base for the association between dairy intake and prediabetes risk derived from multiple cohort studies primarily in Western and predominantly affluent populations. However, the confidence in the estimates was moderate for total dairy and ice cream, very low for cream, and low for other dairy types. There is a high likelihood of residual confounding from background diet and health behaviors, as well as reverse causation related to health or risk awareness. Enhancing the accuracy and detail of dairy intake and dietary behavior data may help address these confounding factors in future studies. Refining study design and analytical strategies to mitigate biases—particularly those associated with reverse causation—will be essential for advancing the field of nutritional epidemiology. Observational evidence supports the need for well-designed RCTs to better understand the causal relationship between dairy intake and hyperglycemia. Future studies should focus on specific dairy types, assess the substitution effects of dairy versus plant-based alternatives, and consider variations in dietary behaviors. In terms of mechanisms, exploring the impact of dairy on metabolic health, gut microbiome alterations, and individual differences is essential for informing public health guidelines and dietary recommendations.

Conclusion

In conclusion, the analysis of associations between dairy intake and prediabetes risk across multiple prospective cohort studies reveals a nuanced picture. While most dairy types were not associated with risk of prediabetes, specific associations varied between cohorts. Overall dairy intake was not associated with higher risk of prediabetes; rather, moderate consumption, particularly of cheese, may be beneficial for lowering this risk. Milk and yogurt were not associated with prediabetes, regardless of their fat content. Network analyses illustrated the interconnectedness of dairy intake, dietary patterns, and sociodemographic, health and cardiometabolic factors. Overall, the findings call for a revaluation of the potential benefits of high-fat dairy, emphasizing the need for comprehensive research to inform public health recommendations.

Publiekssamenvatting / Dutch summary

Zuivelproducten worden wereldwijd veel geconsumeerd en aanbevolen in richtlijnen als onderdeel van een gezond voedingspatroon. Verschillende zuivelproducten kunnen belangrijk zijn voor het behoud van de cardiometabole gezondheid, omdat ze rijk zijn aan eiwitten, oneven-keten vetzuren, calcium, magnesium, kalium, en de vitamines A, D, B2 en B12. Echter, zuivelproducten kunnen ook relatief hoog zijn in verzadigd vet, natrium en suiker, wat schadelijk kan zijn voor de cardiometabole gezondheid. De Nederlandse Gezondheidsraad adviseert dagelijks gebruik van zuivelproducten. Dit advies wordt ondersteund door uitgebreide meta-analyses, die associaties rapporteerden tussen hogere totale zuivelinname—met name magere zuivel en yoghurt—en een lager risico op type 2 diabetes.

Prediabetes is een tussenfase van normoglykemie en type 2 diabetes, gekenmerkt door insulineresistentie en betacel-dysfunctie. Bij prediabetes zijn de plasmaglucoseniveaus hoger dan normaal, maar blijven ze onder de drempel voor de diagnose van type 2 diabetes. De prevalentie van prediabetes neemt toe, met name onder mensen met obesitas en oudere volwassenen. Tot 50% van de mensen met prediabetes kan binnen vijf jaar diabetes ontwikkelen, en ze lopen al een verhoogd risico op microvasculaire en macrovasculaire complicaties. De toenemende prevalentie benadrukt de noodzaak om beïnvloedbare risicofactoren te identificeren om prediabetes te voorkomen of om te keren. Gezien de brede prevalentie kunnen kleine verschuivingen in het risico op prediabetes een aanzienlijke impact op de volksgezondheid hebben. Leefstijlaanpassing is de aanbevolen strategie voor het voorkomen en behandelen van prediabetes, waarvan de effectiviteit is aangetoond in gerandomiseerde gecontroleerde onderzoeken. Bovendien kunnen individuen met prediabetes terugkeren naar normoglykemie door voeding- en leefstijlverandering.

Hoewel de link tussen zuivelinname en type 2 diabetes goed is onderzocht, is de potentiële rol ervan in het voorkomen van eerdere stadia, waaronder prediabetes, minder onderzocht. Inzicht in deze relatie kan waardevolle informatie bieden voor vroege interventie. Het vaststellen van duidelijke associaties tussen voeding en specifieke gezondheidsuitkomsten is echter bijzonder uitdagend vanwege individuele metabole variabiliteit, de complexe en multifactoriële aard van ziekten, en de lange latentieperiodes van ziekten.

Observationele studies bieden waardevolle inzichten in de mogelijke associaties tussen voeding en gezondheidsuitkomsten in grote populaties. Van observationele studies worden prospectieve cohortstudies beschouwd als het hoogste niveau van bewijs, omdat ze blootstellingen meten vóór het optreden van de ziekte. Wanneer observationeel bewijs consistent en goed uitgevoerd is, met voldoende aanpassing voor confounders, en

ondersteund wordt door mechanistische studies die biologische plausibiliteit aantonen, bieden deze associaties een fundament voor het aantonen van causale relaties.

Doel van dit proefschrift

In dit proefschrift wordt de relatie tussen de inname van totale zuivel en verschillende zuiveltypes en prediabetes bestudeerd in prospectieve cohortstudies met algemene populaties. De focus ligt op het analyseren van deze relatie in prospectieve observationele cohortstudies bij volwassenen met normoglykemie uit Nederland, Australië en het Verenigd Koninkrijk. Elk cohort verzamelde continue glykemische metingen bij de baseline en follow-up metingen, gebruikte een gevalideerde voedsel-frequentievragenlijst om de inname van een breed scala aan zuivelproducten te meten en omvatte een uitgebreide set sociodemografische en gezondheidsrisicofactoren.

We volgden een gestandaardiseerd analyseplan om een robuuste bewijsbasis te waarborgen, die hoogwaardige input levert voor het synthetiseren van alle beschikbare gegevens uit prospectieve cohortstudies in een meta-analyse. Om inzicht te krijgen in de onafhankelijke associaties tussen zuivelinname en het risico op prediabetes voerden we regressieanalyses uit waarbij we corrigeerden voor verschillende confounders, waaronder sociodemografische factoren, gezondheidsgedragingen en cardiometabole risicofactoren. Bovendien hebben we een nieuwe benadering toegepast met netwerkanalyses om de holistische onderlinge relaties tussen voeding, sociodemografische factoren, gezondheidsgedragingen, cardiometabole risicofactoren en prediabetes in kaart te brengen. Deze aanpak had tot doel om te bepalen of heterogeniteit in de associaties tussen zuivelinname en het risico op prediabetes (deels) kon worden verklaard door covariatie van voedselinname en gezondheidsgedragingen tussen de verschillende soorten zuivel.

Tot slot voerden we een systematische review en meta-analyse uit om alle beschikbare wetenschappelijke literatuur te identificeren, samen te vatten en te evalueren met betrekking tot de associaties tussen zuivelinname, continue glykemische markers en de incidentie van prediabetes. We beoordeelden de kwaliteit van het bewijs voor elke geïncludeerde studie en evalueerden het vertrouwen in de bevindingen die uit de meta-analyse voortkwamen.

Belangrijkste resultaten

De meeste zuiveltypes waren niet geassocieerd met het risico op het ontwikkelen van prediabetes in de vijf geanalyseerde prospectieve cohortstudies. Echter, de associaties tussen specifieke zuiveltypes en het risico op prediabetes varieerden tussen de individuele cohorten. In de Nederlandse Hoorn Studies (**Hoofdstuk 2**) waren vette gefermenteerde zuivelproducten en kaas, met name vette kaas, geassocieerd met een lager risico op prediabetes, terwijl er geen significante associaties werden gevonden bij het vervangen van vette kaas door andere zuiveltypes. In de Nederlandse Rotterdam

Study (**Hoofdstuk 3**) waren hogere innames van volle melk en volle yoghurt geassocieerd met een lager risico op prediabetes en longitudinale insulineresistentie, met zwakkere positieve associaties voor magere zuivel. In de Australian Diabetes, Obesity, and Lifestyle Study (AusDiab) (**Hoofdstuk 4**) waren vette zuivel, volle melk en totale kaas geassocieerd met een lager risico op prediabetes, met een niet-lineaire associatie voor magere melk. In de Nederlandse Lifelines studie (**Hoofdstuk 5**) was de hoogste inname van gewone en magere melk geassocieerd met een hoger risico op prediabetes, hoewel er geen dosis-responsrelaties werden waargenomen. Ten slotte was in de Britse Fenland studie (**Hoofdstuk 6**) vette zuivel geassocieerd met een hoger risico op prediabetes, terwijl de inname van magere melk geassocieerd was met een lager risico op prediabetes. Het tegenovergestelde werd aangetoond in deze studie voor veranderingen in zuivelinname tijdens de follow-up. Een verhoogde inname van volle melk tijdens de follow-up was geassocieerd met lagere nuchtere plasma glucose en een lager risico op het ontwikkelen van prediabetes en type 2 diabetes.

In de Lifelines studie (**Hoofdstuk 5**) vonden we dat omgekeerde causaliteit de resultaten kan beïnvloeden, zelfs met een prospectief studiedesign. In voedings-epidemiologisch onderzoek komt omgekeerde causaliteit voor wanneer individuen hun eetgewoonten veranderen als reactie op hun bewustzijn van ziekterisico. In onze studie was het bewustzijn van diabetesrisico, gemeten aan de hand van een risicoscore of een wens om af te vallen, geassocieerd met een hogere inname van magere zuivel en een lagere inname van vette zuivel. Dit suggereert dat individuen zonder ziekterisico (bijvoorbeeld degenen zonder obesitas) mogelijk hun voeding kiezen op basis van hedonische voorkeuren in plaats van om gezondheidsredenen. Dit kan leiden tot een lager risico op prediabetes, vooral voor zuiveltypes die niet zijn opgenomen in de voedingsrichtlijnen, zoals vette zuivel en ijs.

In de netwerken van zuivelinname, voedingspatronen, sociodemografische factoren, gezondheidsgedragingen en cardiometabole risicofactoren, en prediabetes, in de Lifelines en Fenland studie (**Hoofdstukken 6 en 7**), clusterden gezondheidsbewuste gedragingen, zoals inname van groenten en fruit en fysieke activiteit. Verschillende voorkeuren voor vette of magere zuivel waren zichtbaar. In beide netwerken clusterden energiedichte voedselgroepen en had tailleomtrek een verbindende rol tussen prediabetes en sociodemografische kenmerken. Desondanks toonden de resultaten van de regressie- en netwerkanalyses beperkte overeenstemming. Hoewel sommige vette en magere zuiveltypes vergelijkbaar gepositioneerd waren binnen de netwerken, waren hun associaties met het risico op prediabetes in regressieanalyses in tegengestelde richtingen, vanwege inverse correlaties tussen hun inname. Omgekeerd waren zuiveltypes met vergelijkbare associaties in regressieanalyses verschillend gepositioneerd in de netwerken. Bovendien hadden zuiveltypes die associaties vertoonden in regressieanalyses niet per se een prominente rol binnen de netwerken. Dit suggereert dat covariantie tussen voeding en gezondheidsgedrag, zoals vastgelegd

door de netwerken, niet volledig de heterogeniteit in associaties tussen zuivelinname en het risico op prediabetes verklaart. De bevindingen wijzen op de complexiteit van de relatie tussen voeding en gezondheid en benadrukken het belang van een holistische benadering voor het begrijpen van deze associaties.

De systematische review van 14 prospectieve cohortstudies (**Hoofdstuk 7**) liet een mix van inverse, positieve en niet-significante associaties zien tussen zuivelinname en continue glykemische metingen. De huidige bewijslast blijft dus onduidelijk over de relatie tussen subtypes van zuivel en glykemische uitkomsten. In de meta-analyse van zes studies met negen cohorten was een hogere inname van totale zuivel niet-lineair geassocieerd met een lager risico op prediabetes, waarbij het laagste risico werd waargenomen bij ongeveer 3,4 porties per dag, waarbij er geen duidelijke trends werden gevonden voor volle versus magere zuivel. Zowel een hogere inname van totale als volle kaas waren niet-lineair geassocieerd met lager risico op prediabetes, met optimale innamenniveaus rond 2,1 porties per dag; echter, innames hoger dan 4 porties per dag waren geassocieerd met een verhoogd risico. Deze resultaten zijn enigszins consistent met bewijs voor type 2 diabetes, hoewel eerdere studies geen duidelijke dosis-responsrelaties hebben vastgesteld. Potentieel gunstige effecten van kaas op hyperglykemie kunnen voortkomen uit de aanwezigheid van specifieke verzadigde vetzuren, het membraan om melkvet, calcium en vitamine K in de kaasmatrix. Voedingsrichtlijnen die uitsluitend magere kaas aanraden, zijn dus niet voldoende onderbouwd door de huidige literatuur. Een hogere inname van ijs was geassocieerd met een lager risico op prediabetes, wat consistent is met bevindingen met betrekking tot type 2 diabetes. Voorzichtigheid bij de interpretatie van dit resultaat is echter geboden vanwege een lage inname van ijs, de moeilijkheid om de inname nauwkeurig te meten vanwege hoge seizoensvariatie, en een hoge potentieel voor omgekeerde causaliteit. Er werden geen associaties gevonden voor totale, volle of magere melk en room, wat consistent is met eerdere literatuur. Ondanks eerdere studies die suggereren dat een hogere inname van yoghurt het risico op type 2 diabetes kan verlagen, vond onze meta-analyse geen associaties tussen inname van yoghurt en het risico op prediabetes. De eerder waargenomen associaties met type 2 diabetes kunnen worden toegeschreven aan lage innames in die studies en gezondheidsgedragingen die verband houden met de consumptie van yoghurt. Inconsistenties in de bevindingen kunnen ook het gevolg zijn van de verscheidenheid aan geconsumeerde yoghurttypes, die variëren in probiotische stammen en suikergehalte.

De bevindingen bieden een bewijsbasis voor de associatie tussen zuivelinname en het risico op prediabetes, afgeleid uit meerdere cohortstudies, voornamelijk in westerse en overwegend welvarende populaties. Het vertrouwen in de schattingen was echter gematigd voor totale zuivel en ijs, zeer laag voor room, en laag voor andere zuiveltypes. Er is een grote kans op confounding van associaties door voedingspatronen en gezondheidsgedrag waarvoor we niet volledig konden corrigeren, evenals

omgekeerde causaliteit gerelateerd aan bewustzijn op gezondheidsrisicio's. Het verbeteren van de nauwkeurigheid en details van de gegevens over zuivelinname en voedingsgedrag kan helpen om de invloed van deze verstorende factoren in toekomstige studies te verminderen. Het verfijnen van de onderzoeksopzet en analytische strategieën om onderzoekbias te verminderen—met name die verband houden met omgekeerde causaliteit—zal essentieel zijn voor de vooruitgang op het gebied van voedingsepidemiologie. Observationeel bewijs onderstreept de noodzaak voor goed ontworpen gerandomiseerde gecontroleerde onderzoeken om de causale relatie tussen zuivelinname en hyperglykemie vast te stellen. Toekomstige studies moeten zich richten op specifieke zuiveltypes, de substitutie-effecten van zuivel versus plantaardige alternatieven, rekening houdend met de variaties in voedingsgedrag. Wat betreft mechanismen is het essentieel om de impact van zuivel op de metabole gezondheid, veranderingen in de darmmicrobioom en individuele verschillen hierin te onderzoeken om zo publieke gezondheidsrichtlijnen en voedingsaanbevelingen te informeren.

Conclusie

Concluderend laat de analyse van de associaties tussen zuivelinname en het risico op prediabetes in meerdere prospectieve cohortstudies een genuanceerd beeld zien. Terwijl de meeste zuiveltypes niet geassocieerd waren met het risico op prediabetes, varieerden specifieke associaties tussen de cohorten. De totale zuivelinname was niet geassocieerd met een hoger risico op prediabetes; integendeel, een gematigde consumptie, vooral van kaas, kan gunstig zijn voor het verlagen van het risico op prediabetes. Inname van melk en yoghurt was niet geassocieerd met prediabetes, ongeacht hun vetgehalte. De netwerkanalyses illustreerden de onderlinge connecties tussen zuivelinname, voedingspatronen en sociodemografische, gezondheids- en cardiometabole factoren. Samenvattend vragen de bevindingen om een revaluatie van de mogelijke voordelen van volle zuivel en benadrukken ze de noodzaak van uitgebreider onderzoek ter ondersteuning van publieke gezondheidsaanbevelingen.

List of publications

Westerhof I, De Boer A, Lupattelli A, **Slurink IAL**, *et al.* & Rozhnova G. A comparative analysis of risk classification for severe COVID-19 based on chronic medical conditions. Submitted.

Calame W, **Slurink IAL**, Budelli, A. New Standards for Nutrition Science, Concepts and Methods—Novel Approach to Substantiate Cause- and -Effect Relationships in Nutritional Science by Ranking Studies and Subsequent Statistical Modelling. Submitted.

Van den Houdt S, Mertens G & **Slurink IAL**. Unraveling the Complex Web of Long COVID symptoms: A Network Analysis Approach. Submitted.

Slurink IAL, Vogtschmidt YD, Brummel B, Kupper N, Smeets T, & Soedamah-Muthu SS. Dairy intake in relation to prediabetes and continuous glycaemic outcomes: a systematic review and dose-response meta-analysis of prospective cohort studies. Current Developments in Nutrition, 2024.

Slurink IAL, Kupper N, Smeets T, & Soedamah-Muthu SS. Dairy consumption and risk of prediabetes and type 2 diabetes in the Fenland study. Clinical Nutrition, Volume 43, Issue 11, 69 - 79.

Slurink IAL, van den Houdt SCM, Mertens G. Who develops long COVID? Longitudinal pre-pandemic predictors of long COVID and symptom clusters in a representative Dutch population. Int J Infect Dis. 2024 Jul;144:107048.

Slurink IAL, Nyklíček I, Kint R, Tak D, Schiffer AA, Langenhoff B, Ouwens MA & Soedamah-Muthu SS. Longitudinal trajectories and psychological predictors of weight loss and quality of life until 3 years after bariatric surgery. Journal of Psychosomatic Research, 2024.

Van den Houdt S, **Slurink IAL**, Mertens G. Long COVID is not a uniform syndrome: Evidence from person-level symptom clusters using latent class analysis. Journal of Infection and Public Health, 2023.

Slurink IAL, Corpeleijn E, Bakker SJL, Jongerling J, Kupper N, Smeets T & Soedamah-Muthu SS. Dairy consumption and incident prediabetes: prospective associations and network models in the large population-based Lifelines study. The American journal of clinical nutrition, 2023; 118.6: 1077-1090.

Slurink IAL, Chen L, Magliano DJ, Kupper N, Smeets T & Soedamah-Muthu SS. Dairy product consumption and incident prediabetes in the Australian Diabetes, Obesity and Lifestyle Study with 12 years follow up. The Journal of Nutrition, 2023; 153(6), 1742-1752.

Habibovic M, Douma E, **Slurink IAL**, Kop WJ & Soedamah-Muthu SS. Positive Psychological Constructs and Lifestyle Behaviours in a Community-Based Sample. Psychology and Behavioral Sciences, 2022; 11(6), 185.

Lifestyle4Health Onderzoek. Meer aandacht voor leefstijlbehoeften van mensen met meerdere chronische aandoeningen. 24 november, 2022 Geschreven door: **Isabel Slurink**, Mirela Habibovic, Meeke Hoedjes en Sabita Soedamah-Muthu.

Slurink IAL, Smaardijk VR, Kop WJ, Kupper N, Mols F, Schoormans D & Soedamah-Muthu SS. Changes in Perceived Stress and Lifestyle Behaviors in Response to the COVID-19 Pandemic in The Netherlands: An Online Longitudinal Survey Study. International Journal of Environmental Research and Public Health, 2022; 19(7), 4375.

Slurink IAL, Voortman T, Ochoa-Rosales C, *et al.* Dairy product consumption in relation to incident prediabetes and longitudinal insulin resistance in the Rotterdam Study. Nutrients, 2022; 14(13), 415.

Slurink IAL, den Braver NR, Rutters F, *et al.* Dairy product consumption and incident prediabetes in Dutch middle-aged adults: The Hoorn Studies prospective cohort. Eur J Nutr 2022; 61, 183–196.

Slurink IAL, van Aar F, Parkkali S, Heijman T, Götz H, Kampman K, van Weert Y, van Benthem BHB, van Laar T & Op de Coul E. Recently acquired HIV infections and associated factors among men who have sex with men diagnosed at Dutch sexual health centres. International Journal of STD & AIDS. 2021;32(10):946-956.

Slurink IAL, Götz HM, van Aar F, van Benthem BHB. Educational level and risk of sexually transmitted infections among clients of Sexual Health Centers: a cross-sectional study using the Dutch national surveillance database. International Journal of STD & AIDS. 2021; 32(11):1004-1013.

Slurink IAL, van de Baan F, van Sighem A, *et al.* Monitoring recently acquired HIV infections in Amsterdam, the Netherlands: the attribution of test locations. 2021. Frontiers in Reproductive Health 2021; 3.

Slurink IAL, den Braver N, Rutters F, Kupper N, Smeets T, Beulens, J & Soedamah-Muthu SS. Dairy product consumption and incident prediabetes in Dutch middle-aged adults: The Hoorn studies. Nederlands tijdschrift voor Diabetologie. 2020. 18, 4, p. 35-36 25. (meeting abstract)

Slurink IAL, van Dam A, Heijne J, van Bergen J & van Benthem BHB. Laboratoriumsurveillance van soa: van toegevoegde waarde? Infectieziektebulletin 2020;31(1).

Slurink IAL & Soedamah-Muthu SS. Dairy consumption and cardiometabolic risk: advocating change on change analyses. Am J Clin Nutr, 2020. 111(5): p. 944-945. (Editorial)

Slurink IAL, Groen K, Götz HM, *et al.* GPs and sexual health centres contribution to STI consultations in five Dutch regions using laboratory data of Chlamydia trachomatis testing. Int J STD AIDS, 2020. 31(6): p. 517-525

Achterbergh RCA, Drückler S, van Rooijen MS, van Aar F, **Slurink IAL**, de Vries HJ and Boyd A. Sex, drugs, and sexually transmitted infections: a latent class analysis among men who have sex with men in Amsterdam and surrounding urban regions, the Netherlands. Drug Alcohol Depend, 2020 Jan 1;206:107526.

Slurink IAL, van Benthem BHB, van Rooijen MS, Achterbergh RCA & van Aar F. Latent classes of sexual risk and corresponding STI and HIV positivity among MSM attending centres for sexual health in the Netherlands. Sex Transm Infect, 2020; 96(1): p. 33-39.

Slurink IAL, van Aar F, Op de Coul ELM, *et al.* Sexually transmitted infections in the Netherlands in 2018. Bilthoven: Centre for Infectious Disease Control - National Institute for Public Health and the Environment (RIVM), 2019.

Slurink IAL, Soedamah-Muthu SS. Dietary Fibre and Cardiovascular Risk in Diabetes Mellitus. J Clin Nutr Diet, 2016; 2:3.

Dankwoord / Acknowledgements

Welkom bij het dankwoord! Lees ook zeker de samenvatting, daar heb ik ook erg mijn best op gedaan. Maar nee, even serieus, vier jaar werk is nu gebundeld in een boekje, en daar ben ik best trots op, en een heleboel fantastische mensen om mij heen heel erg dankbaar voor. Een heerlijk cheesy begin zo!

Allereerst wil ik mijn grote dank uitspreken aan de bijna 100,000 deelnemers van de Hoorn Studies, de Rotterdam Study, de Lifelines study, AusDiab en de Fenland study. Jullie bijdrage aan de wetenschap wordt enorm gewaardeerd, zeker omdat het invullen van de voedingsvragenlijst een behoorlijke tijd kost en soms de nodige frustratie met zich mee kan brengen. Zonder jullie inzet was dit onderzoek niet mogelijk geweest.

Ik wil graag mijn promotoren bedanken voor hun advies, steun, en kritische blik. De samenwerking met jullie was fijn en inspirerend, en ik waardeer de tijd en energie die jullie in mij en in dit onderzoek hebben gestoken.

Dr. Soedamah-Muthu, beste Sabita, een bijzonder dankwoord gaat uit naar jou. Wij kennen elkaar al sinds ik mijn bachelorthesis bij jou heb geschreven en dat vind ik best bijzonder. Je hebt mij toen geleerd om wetenschappelijke informatie kritisch te beoordelen en zelfverzekerd de leiding te nemen tijdens overleggen. Na twee jaar soa-surveillance greep ik dan ook de kans om mijn PhD-project bij jou te doen met twee handen aan. Maar, ik geef het maar gewoon toe, na twee maanden in Tilburg twijfelde of ik hier wel goed aan deed. Ik was net terug in de voedingsepidemiologie en hoopte op een rustig begin om mij in te lezen. Maar bij het eerste overleg kreeg ik meteen de opdracht om een editorial voor AJCN te schrijven. Ik had het kunnen weten, je pusht je studenten altijd tot het uiterste, maar het levert wel prachtige resultaten op, zoals mijn eerste publicatie! Jouw begeleidersstijl is heel bijzonder: betrokken, geïnteresseerd, en altijd met veel tijd voor persoonlijke gesprekken, waardoor ik mij gesteund en gewaardeerd voelde. Je hebt mij altijd aangemoedigd om groot te denken, vragen uit te pluizen, zijpaden te ontdekken en niet genoegen te nemen met wat er is. Jouw mensgerichte werkwijze en je advies op wetenschappelijk, carrière- en persoonlijk gebied zullen mij blijven inspireren in mijn verdere ontwikkeling.

Prof. dr. Smeets, beste Tom, ook jou wil ik bedanken voor onze fijne samenwerking en je begeleiding. Ook al was je meer vanaf de zijlijn betrokken, je inzichten hebben mij geholpen om gedegen onderzoek te doen en mooie manuscripten te schrijven. Jouw steun heeft een waardevolle bijdrage geleverd aan mijn groei als onderzoeker.

Ik wil graag de leden van de promotiecommissie, prof. dr. Ien van de Goor, prof. dr.ir. Ingeborg Brouwer, prof. dr. Marianne Geleijnse, prof. dr. Ronald Mensink en dr. Kirsten Berk, hartelijk danken voor het kritisch beoordelen van mijn proefschrift en voor het opponeren tijdens mijn verdediging.

Ook een grote dank voor dr. Kupper; beste Nina, dank voor je waardevolle bijdrage aan de onderzoeken beschreven in dit proefschrift. Vanuit jouw totaal andere expertise kwamen vaak verrassende inzichten, zeker bij de interpretatie van de psychometrische netwerkmodellen. Jouw perspectief heeft mij geholpen om anders na te denken en nieuwe verbindingen te leggen, niet alleen tussen de variabelen, maar ook in de bredere context van het onderzoek.

Ik wil ook graag alle co-auteurs bedanken voor hun inzet en bijdrage aan de verschillende hoofdstukken in dit proefschrift. In het bijzonder wil ik dr. Trudy Voortman bedanken; de vaardigheden die ik tijdens mijn masterstage bij jou heb geleerd, vormden de basis voor dit proefschrift. Ik kijk er naar uit om te blijven samenwerken! Bijna dr. Vogtschmidt, lieve Kim, ik heb echt heel erg genoten van de DNSG-congressen, en dan bedoel ik natuurlijk de wijntjes in de veel te luxe congreshotels. Superfijn om ervaringen en advies met elkaar te delen, en samen te werken op de meta-analyse. Je bent zo'n gedreven, precieze en hardwerkende collega, en ik hoop dan ook dat we nog veel papers samen gaan schrijven! Dr. Lei Chen, thank you for collaborating with me on the AusDiab paper. It was a pleasure getting to know you, and your ability to balance work and family is truly inspiring. Although it wasn't always convenient to communicate the analysis plan and work with different statistical software, I am very pleased with our final paper! Bo Brummel, heel erg bedankt voor de data-extractie voor de meta-analyse. Jouw energie, vriendelijkheid, en je manier van connectie maken met mensen, is echt bijzonder. Hopelijk maak ik snel weer veel van dat enthousiasme mee op de dansvloer.

Nog leuker dan zelf uitzoeken en leren, is het om anderen enthousiast te maken voor de wetenschap. Ik wil dan ook allereerst Hilde bedanken voor het aangaan van de uitdaging om een PhD-project te starten en zo het onderzoek in dit proefschrift voort te zetten. Ook wil ik Thomas, Jasmijn, Lotte en Jojanne bedanken voor hun waardevolle bijdrage en inzet.

Ik heb ontzettend veel plezier gehad op de afdeling medische en klinische psychologie en veel geleerd van dit andere vakgebied. Ik wil dan ook graag mijn collega's bedanken die mij hebben ondersteund in mijn zijstapje naar de medische psychologie, met onderzoeken naar stress, positieve psychologie, leefstijlgedrag, en kwaliteit van leven. In het bijzonder dr. Ivan Nyklíček, prof. dr. Wijo Kop, en dr. Mirela Habibovic, dank voor de fijne samenwerking, jullie interesse in mij en mijn onderzoek, en de gezellige tijd op de afdeling.

Gaetan, ik wil ook jou heel graag bedanken, door jou waren de lange onderzoeksdagen een stuk minder zwaar. Jouw 'sarcastische, te gemene grappen die altijd ten koste van mij gingen' vormden een heel andere manier van steun en advies dan die van Sabita, maar ik kon ze van tijd tot tijd goed gebruiken. En jouw afleiding heeft toch maar mooi geleid tot een NWO-beurs, een opstapje terug naar de infectieziekte-epidemiologie, en een heel gezellige samenwerking met Sophie erbij. Ook jij Sophie heel erg bedankt dat je

dit project zo succesvol hebt uitgevoerd en voor jouw geduld in de omgang met ons als projectmanagers. Bewonderingswaardig hoe veel jij voor elkaar krijgt in een korte tijd!

Ook een grote dank aan al mijn oud-collega's van MKP; in het bijzonder Lotte, Daniëlle, Emma, Janniek, Dinah, Lianne, Myrthe, Paul, Stefanie, Tom, Bo, Lisa, Manon en Tom. Ik heb enorm genoten van de online activiteiten tijdens de pandemie, en daarna de stapavondjes, aio-uitjes, schrijfweken, carnaval, maar ook vooral van de vele momentjes op kantoor. Ook veel dank aan Abbie, Anouk, Barbara, Carmen, Charlotte, Dounya, Egon, Ellen, Elvire, Esther, Eveline, Floortje, Frederieke, Gubing, Heidi, Inge, Jamie, Jonas, Josine, Kevin, Kitty, Laura, Lotje, Marijn, Marlies, Marlot, Meeke, Marije, Nathan, Nicole, Nina, Noor, Paula, Rosie, Sandra, Saskia, Silke, Simone, Veerle en Wendy voor de gezellige lunches en borrels.

Dan wil ik ook graag mijn favoriete collega bedanken, Thijs; zonder jou was ik het thuiswerken tijdens de lockdowns niet doorgekomen met behoud van mijn mentale gezondheid. De woensdagen samenwerken waren het hoogtepunt van de werkweek, ook al waren we soms meer bezig met het lunchmenu dan met de analyses. Ik bewonder jouw passie voor de ontwikkeling van algoritmes voor onkruidherkenningsrobotten, en ook al neigt het soms naar mansplaining, vind ik het geweldig hoe enthousiast je hierover kan vertellen. Ik hoop dan ook, ondanks je verhuizing naar de andere kant van het land, dat de sportavondjes, roddelen over onze partners, en stoom afblazen over werk blijft. Heel erg bedankt voor onze fijne vriendschap!

Lieve Ottilie, ook jou wil ik van harte bedanken. Je bent een geweldige vriendin al sinds de middelbare school. Jouw fantastische humor, betrokkenheid en kijk op het leven waardeer ik enorm. Bedankt ook dat je alle grammaticale fouten uit mijn proefschrift hebt gepluisd. Dus, beste lezer, mocht je er nog een tegenkomen, dan mag je bij Ottilie aankloppen.

Ik wil ook graag mijn lieve vrienden bedanken voor de fijne vriendschappen. Zeker degenen die voor mij hebben gekozen en al heel lang gezelligheid in mijn leven brengen, maar ook zeker degenen die niet voor mij hebben gekozen en mij gratis bij Han kregen. Nog even in het bijzonder Lena, die de meest bijzondere dag (tot nu toe) uit mijn leven tot in de puntjes verzorgd heeft. Beste lezer, mocht je na het lezen van de conclusie van dit proefschrift opeens heel veel zin hebben in kaas, bezoek dan zeker haar prachtige kaaswinkeltje in Wageningen. Trots op jou!

Ik wil ook graag mijn familie bedanken. Allereerst mijn schoonouders Eric en Jacqueline, jullie zijn fantastisch lieve mensen - altijd geïnteresseerd en verwelkomend. Lieve Oma, bedankt dat je mij hebt laten zien hoe leuk het is om creatief te zijn, en Opa, bedankt dat je een geweldige opa voor mij bent geweest. En dan mijn gezin, mijn meestal leuke broers Boris en Wijnand. Na een lange donderdag is de spaghetti altijd iets om naar uit

te kijken, hopelijk blijft dit nog een lange tijd een traditie. Ik wil jullie bedanken voor onze fijne band, de gezelligheid en het er zijn voor elkaar. Beste Jasper Jan, ik wil je bedanken voor de flinke dosis humor die je brengt in ons gezin, dat je altijd voor ons klaar staat om advies te geven over werk en relaties, en de kansen die je ons geeft. Lieve Jaap, ik weet dat het soms lastig is om een dochter te hebben die een stuk meer weet over voeding, gezondheid, vaccinaties, etc etc, en je onder de tafel praat tijdens discussies, maar dan moet je de wetenschap maar wat serieuzer nemen ;) Ik wil je bedanken voor de normen en waarden die je ons hebt bijgebracht, de grote liefde voor de natuur, en een brede blik op de wereld. Lieve Annemarie, ik bewonder wat jij allemaal hebt bereikt terwijl je voor ons zorgde. Je bent dan ook een grote bron van inspiratie voor mij, en ik heb veel aan jou als sparringpartner. Bedankt voor mijn fijne jeugd, en dat wij altijd jouw prioriteit zijn.

Lieve Han, mijn dankbaarheid voor jou is moeilijk om in een paar zinnen samen te vatten. Al was zonder jou dit proefschrift waarschijnlijk een stuk sneller af geweest, mijn leven zou lang niet zo leuk, relaxed en zinvol zijn. Bij mijn eerste publicatie gaf je mij een orchidee, waarmee je liet zien dat je begreep dat het een echt hoogtepunt voor me was. Het kopen van ons huis, trouwen, promoveren, en nu het volgende avontuur - ik had het niet met iemand anders willen meemaken dan met mijn beste vriend. Bedankt voor jouw onvoorwaardelijke steun en liefde.

About the author

Isabel Slurink was born on 25 April 1994 in Wageningen, the Netherlands. She completed her pre-university education at Pallas Athene College, Ede, and the Junior College Utrecht, Utrecht, in 2011. In 2015, she obtained her Bachelor's (BSc) degree in Nutrition and Health at Wageningen University (WUR). Subsequently, she completed a two-year Master's program (MSc) in Nutrition and Health, specializing in Epidemiology and Public Health. During her Masters, she did a research internship at the Erasmus University in Rotterdam on diet quality in relation to cardiometabolic health. After completing her studies, she joined the National Institute for Public Health and the Environment (RIVM) as a junior researcher at the Center for Infectious Disease Control. There, she conducted epidemiological studies on health behaviour and educational level in relation to the incidence of sexually transmitted infections, and contributed to surveillance efforts. In January 2020, Isabel started her PhD Research at the Center of Research on Psychological disorders and Somatic diseases (CoRPS) of Tilburg University. Her research focused on dairy in relation to prediabetes in large observational cohort studies using epidemiological methods. She also studied the usability of food pictures in nutritional research, weight and quality of life after bariatric surgery, the impact of the COVID-19 pandemic on lifestyle behaviours and stress, and the symptom presentation and predictive factors of post-COVID. She continued her career in epidemiological research as a postdoctoral researcher at the Julius Center for Health Sciences and Primary Care of the University Medical Center Utrecht. Her current research focuses on infectious diseases and vaccination in relation to post-infection sequala and non-communicable diseases, utilizing registry data-linkage approaches. By working across both the infectious disease epidemiology department and the data science department, she leverages innovative epidemiological methods to further her mission of advancing public health.