BDA The Association of UK Dietitians
THE OFFICIAL JOURNAL OF THE BRITISH DIETETIC ASSOCIATION
Journal of Human Nutrition and Dietetics

TYPE 2 DIABETES

Substitution among milk and yogurt products and the risk of incident type 2 diabetes in the EPIC-NL cohort

J. M. Stuber,[1,2] L. E. T. Vissers,[1] W. M. M. Verschuren,[1,3] J. M. A. Boer,[3] Y. T. van der Schouw[1] & I. Sluijs[1]

[1]Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands
[2]Department of Epidemiology and Biostatistics, Amsterdam Public Health Research Institute, Amsterdam UMC, VU University, Amsterdam, The Netherlands
[3]National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands

Keywords
cohort studies, dairy products, diabetes, milk, substitution models, yogurt.

Correspondence
Y. T. van der Schouw, Julius Centre for Health Sciences and Primary Care, University Medical Centre Utrecht, STRT 6.131, PO Box 85500, 3508GA Utrecht, The Netherlands.
Tel.: +31 88 755 9301
E-mail: y.t.vanderschouw@umcutrecht.nl

How to cite this article
Stuber J.M., Vissers L.E.T., Verschuren W.M.M, Boer J.M.A., van der Schouw Y.T. & Sluijs I. (2021) Substitution among milk and yogurt products and the risk of incident type 2 diabetes in the EPIC-NL cohort. *J Hum Nutr Diet*. **34**, 54–63
https://doi.org/10.1111/jhn.12767

[Correction added on 2 June 2020: The author name "M. W. M. Verschuren" has been amended to "W. M. M. Verschuren".]

Abstract

Background: Higher dairy consumption has been associated with lower type 2 diabetes (T2D) risk, whereas dairy product subtypes appear to differ in their T2D risk association. We investigated whether replacing one type of milk or yogurt product with another is associated with T2D incidence.
Methods: Participants of the European Prospective Investigation into Cancer and Nutrition-Netherlands (EPIC-NL) cohort (n = 35 982) were included in the present study. Information on milk and yogurt consumption at baseline was obtained by a validated food frequency questionnaire. T2D cases were identified by self-report or linkage to the hospital discharge registry, and validated by consulting the general practitioner. Multivariable Cox proportional hazard models were used to estimate associations.
Results: During a mean of 15 years of follow-up, 1467 indecent T2D cases were validated. Median total milk and yogurt intake was 1.5 servings (25th percentile to 75th percentile: 0.8–2.4). After adjustment for demographic and cardiovascular risk factors, replacement of one serving (200 g) of whole-fat milk [hazard ratio (HR) = 0.93, 95% confidence interval (CI) = 0.60–1.44], buttermilk (HR = 0.88, 95% CI = 0.58-1.34), skimmed milk (HR = 0.87, 95% CI = 0.57–1.32) or skimmed fermented milk (HR = 0.99, 95% CI = 0.63–1.54) with whole-fat yogurt was not associated with T2D risk. Substitutions among other milk and yogurt products were also not associated with T2D risk. Sensitivity analysis investigating T2D risk halfway follow-up suggested a lower risk for substitutions with whole-fat yogurt.
Conclusions: No evidence was found for the association between substitutions among milk and yogurt products and the risk of incident T2D, although we cannot exclude possible attenuation of results as a result of dietary changes over time. This analysis should be repeated in a population with a wider consumption range of whole-fat yogurt.

Introduction

In 2015, 8.8% of the adults worldwide were diagnosed with diabetes and this number is expected to rise to 10.4% by 2040 [(1)]. This chronic disease results in a decreased quality of life and higher risks of morbidity and mortality, placing a major burden on healthcare systems [(2)]. Identifying modifiable risk factors for type 2 diabetes (T2D) is important for improving public health prevention strategies.

Healthy lifestyle behaviours include several dietary factors that have been associated with a lower T2D risk [(3,4)], including a high consumption of total dairy products [(5)]. The mechanisms behind dairy consumption and a

reduced T2D risk are not fully understood, although several potential mechanisms have been proposed [6]. Dairy products are heterogeneous as a result of differences in the amount of water, sodium, fats and added sugar, as well as the level of fermentation. The consumption of fermented dairy products has been shown to result in metabolic health benefits by causing a shift in the gut microbial population [7-9]. Pentadecanoic acid and heptadecanoic acid are recognised as markers for dairy fat consumption [10]. Their amounts in erythrocyte membranes are inversely associated with T2D risk [11] and proportions in plasma phospholipids are inversely associated with fasting insulin and glucose [12,13]. Although these fasts have been associated with decreased T2D risk markers, cause and effect relationships remain uncertain. Whey protein consumption has been linked to postprandial stimulate insulin production and activity [14]. Calcium content might play also a role becuase it possibly has antiobesity bioactivity effects [15,16]. In addition, obesity is a major risk factor for the development of T2D and the ability of milk to assist in appetite control might be another relevant contributing factor to the lower risk of T2D being associated with dairy consumption [17].

A dose–response meta-analysis including 22 prospective cohort studies (±580 000 participants; 43 000 cases) observed that total yogurt consumption was associated with a lower risk of T2D, whereas no associations with T2D were observed for the group of whole-fat dairy products and total milk, skimmed milk and whole-fat milk [5]. This meta-analysis is based on studies that compared T2D risk between individuals with different levels of dairy product consumption, at the same time as keeping energy intake at a constant level. As a result of this, individuals differ not only in dairy product intake, but also in the intake of other unspecified energy-providing foods. Hence, the results cannot be interpreted as a direct comparison between individual dairy products.

Substitution modelling can be used to gain further insight into the differences between dairy products and their association with T2D risk because it can be interpreted as a direct comparison between products [18]. Only one study has examined substitution within the group of dairy products so far, and it was observed that consumption of whole-fat yogurt instead of any other milk and yogurt product (i.e. skimmed milk, whole-fat milk, buttermilk or skimmed yogurt) was associated with a lower risk of T2D in a Danish population [19]. We aimed to replicate this previous work by investigating whether substitutions between skimmed milk, whole-fat milk, buttermilk, skimmed fermented milk products and whole-fat yogurt were associated with changes in the incidence of T2D in a Dutch population.

Materials and methods

Study population

Data were sourced from the Dutch contribution to the European Prospective Investigation Into Cancer and Nutrition (EPIC-NL) study. This prospective cohort emerged out of two large cohorts. First, the Prospect-EPIC cohort (n = 17 357), which invited women aged 49–70 years who participated in the nationwide breast cancer screening programme and were living in the Dutch city of Utrecht or its vicinity. Second, the Monitoring Project on Risk Factors for Chronic Diseases (MORGEN) cohort (n = 22 654), covering a randomly selected population sample of both males and females aged 20–59 years from three Dutch towns (Amsterdam, Doetinchem and Maastricht). Both cohorts combined provided baseline measurements to form the EPIC-NL cohort of 40 011 individuals. All participants were enrolled in the study between 1993 and 1997 and have been followed up for a mean (SD) of 15 (3) years for the occurrence of T2D. Further details on recruitment and design of the EPIC-NL cohort are described elsewhere [20]. Prior to study inclusion, participants provided their written informed consent and both cohorts were approved by the local medical ethic committees: the institutional review board of the University Medical Centre Utrecht for the Prospect-EPIC cohort and the Medical Ethical Committee of TNO Nutrition and Food Research for the MORGEN cohort. The study was performed in accordance with the Declaration of Helsinki.

From the 40 011 study participants, excluded participants withdrew their permission for inclusion in the study (n = 1); were missing informed consent to retrieve data from the general practitioner, municipal register office or linkage to the hospital discharge diagnoses registry (n = 1789); comprised unvalidated potential T2D cases (n = 488) and participants with type 1 and type 2 diabetes at baseline (n = 820); were missing data on milk and yogurt consumption (n = 172); were non-consumers of milk and yogurt products (n = 209); had an unrealistic reported dietary intake (highest and lowest 0.5% based on the ratio of total energy intake to the estimated basal metabolic rate) (n = 327); were missing data on covariate smoking (n = 104) and education (n = 108); and were participants with a negative follow-up time (n = 11). This left 35 982 individuals for the analysis.

Assessment of diet and milk and yogurt consumption

Dietary intake was measured at study enrolment by a self-administered validated food frequency questionnaire (FFQ) [21]. Participants were asked to report the average intake of 79 main food categories (in times per day, per

week, per month or per year, or as never) over the past year [22]. Inconsistencies in the FFQ were checked by a dietitian and, if needed, were resolved by contacting the participant [21].

For all milk and yogurt products, total consumption was calculated in grams per day based on serving sizes of 200 g. Products were categorised into five main groups: (i) skimmed milk, including semi-skimmed milk, skimmed coffee milk and semi-skimmed coffee milk; (ii) whole-fat milk, including powdered milk and whole-fat coffee milk; (iii) buttermilk; (iv) skimmed fermented milk products, including skimmed yogurt, drink yogurt and quark; and (v) whole-fat yogurt. The whole-fat groups contained >3 g fat 100 g^{-1}, the skimmed dairy groups contained <3 g fat 100 g^{-1}, and buttermilk groups contained <1 g fat 100 g^{-1}. Other milk products, such as custard, chocolate milk, ice cream and whipped cream, were not included because these products are significantly different in macronutrient composition. Hence, replacement with these products within a diet would unlikely provide health benefits [23].

The relative validity of the FFQ was assessed by comparing collected data on milk product consumption with 12-monthly 24-h recalls among 121 participants. Spearman's rank correlation coefficients for the group of total milk and milk product consumption were 0.69 and 0.77 for males and females, respectively [22].

Additional data on dietary consumption were also collected by the FFQ. The intake of fruits, vegetables, coffee, red meat, processed meat, sugar-sweetened beverages, alcohol and fibre was assessed. Alcohol consumption was categorised in non-consumers, light drinkers 0.1–10 g day^{-1}, moderate drinkers 10–20 g day^{-1} and heavy drinkers >20 g day^{-1}. Fibre consumption was energy adjusted following the nutrient residual model [24] because variation is strongly related to total energy consumption. Energy intake in kilocalories (kcal) per day was calculated based on the total daily consumption with the use of the Dutch Food Composition Table (1996) [25].

Assessment of covariates

Baseline data on potential risk factors for chronic diseases were collected by questionnaires. Smoking status was categorised into current, former or never. Educational level was classified as low (primary education to intermediate vocational education), average (higher secondary education) or high (higher vocational education or university). Physical activity was measured with the validated EPIC questionnaire as used in all EPIC cohorts [26]. Subsequently, the Cambridge Physical Activity Index (CPAI) score was used to categorise participants into inactive, moderately inactive, moderately active and active [27]. As a result of missing values, CPAI-scores could not be calculated for 14% of the study population. Single linear regression modelling was applied to impute the missing scores (missing value analysis procedure in SPSS; IBM Corp., Armonk, NY, USA).

Height (cm) was measured during physical examination and body weight was measured to the nearest 0.5 kg using a floor scale (Seca, Atlanta, GA, USA), without shoes and in light clothing. The body mass index (BMI) was calculated as weight divided by height squared (kg m^{-2}). Presence of hypertension (yes, no) was defined as a mean diastolic blood pressure of >90 mmHg and/or a mean systolic blood pressure of >140 mmHg measured two times in the supine position on the right arm using a Boso Oscillomat (Bosch & Son, Jungingen, Germany) for Prospect-EPIC participants. For MORGEN-EPIC participants, the left arm was measured using a random zero sphygmomanometer. Presence of hypertension was also defined based on self-reported use of antihypertensive medication or physician-diagnosed existence of hypertension. Total serum cholesterol levels and high-density lipoprotein (HDL) concentrations were measured [20]. Total cholesterol to HDL ratio was calculated by dividing total cholesterol by HDL.

Occurrence of type 2 diabetes

A two-step approach was used for the identification and validation of potential T2D cases. For the identification of potential cases, information was obtained through linkage with the hospital discharge diagnosis registry and from follow-up questionnaires. In the hospital discharge diagnoses registry, information on diagnoses was coded according to the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) [28]. Code 250 and underlying codes were used to identify potential T2D cases. The follow-up questionnaires collected data on self-reported diabetes diagnosis, and were sent out with intervals of three to five years (1998–2002 questionnaire 1; 2003–07, questionnaire 2; 2011–12, questionnaire 3). Prospect-EPIC participants additionally received a urinary glucose strip test with the first questionnaire. They were asked to self-report whether the strip had turned purple after 10 s, for detection of glucosuria.

All potential T2D cases up to 2006 were validated by consulting the general practitioner or the pharmacist [21]. The pharmacist was only used to confirm presence, not absence, of diabetes. For all potential cases identified after 2006, only the general practitioner was used as verification source. The verification source provided the diagnosis year and we set the diagnosis date for all identified cases at 1 January, in the year of diagnosis. Verification information was available for 81% of the identified

potential T2D cases. All non-verified potential cases were excluded from the primary analysis because those participants could not be categorised as a case, nor as a non-case. Follow-up was complete until 31 December 2010.

Descriptive analysis

Baseline characteristics were examined per tertile of milk and yogurt product consumption. Results for continuous variables were described as the mean (SD), or as median with the 25th percentile (P_{25}) and 75th percentile (P_{75}) for variables that were not normally distributed. Categorical variables were described in frequencies and percentages. Spearman's rank correlation coefficients were calculated to explore potential correlations between consumption of the different types of milk and yogurt products.

Main survival analysis

Multivariable Cox proportional hazard regression models were used to estimate the hazard ratio (HR) and 95% confidence interval (CI) for substitution of milk and yogurt products and incident T2D. Age was used as underlying timescale [29]. Follow-up duration was calculated starting from enrolment date, to the year of T2D diagnosis, date of death, date of emigration or end of follow-up.

We modelled substitution of milk and yogurt products in servings, with a serving size of 200 g. The substitution model included a variable representing the total number of servings of milk and yogurt products consumed per day, and servings consumed of individual milk and yogurt products in subgroups, except for the milk or yogurt subgroup that would be replaced (i.e. four out of five groups were included). As a result, the estimated HR and 95% CI can be interpreted as the risk of T2D for one serving higher intake of the subgroups included in the model at the expense of one serving lower intake of the subgroup not in the model.

Four models to adjust for potential confounding were used. Model 1 was adjusted for sex (male, female) and total energy intake (kcal day^{-1}). Model 2 was further adjusted for smoking status (current, former, never), physical activity (inactive, moderately inactive, moderately active, active), education level (low, average, high), hypertension (yes, no) and alcohol consumption (non-consumers, light drinkers, moderate drinkers and heavy drinkers). Model 3 was further adjusted for the consumption of fruit, vegetables, processed meat, red meat, coffee, sugar-sweetened beverages and energy adjusted fibre (g day^{-1}). Model 4 was additionally adjusted for T2D risk factors that were considered as potential mediators, namely BMI [2] (kg m^{-2}) and the cholesterol ratio [30]. Participants with missing values on these potential mediators were excluded from model 4. Because model 4 included the potential mediators, final conclusions were based on adjusted model 3 and therefore the main results section focus on describing results derived from model 3. Study cohort (Prospect or MORGEN) was included as a stratum variable for all analyses.

Assumptions of the Cox proportional hazard regression model were evaluated. First, the proportional hazards assumption was investigated by plotting scaled Schoenfeld residuals against time. Thereafter, independence between the scaled Schoenfeld residuals and time was tested with chi-squared tests for each covariate and for the overall model. Next, deviance residuals were plotted for all included variables, to check for influential observations. We did not detect violation of the proportional hazard assumption. Martingale residuals including fitted lines with lowess function were plotted against the milk and yogurt subgroups to evaluate linearity. No indications for non-linearity were observed. The assumption of independent delayed entry was investigated by including the date of enrolment in the final adjusted model (model 3). We did not find evidence for violation of this assumption.

Hazard ratios in which the 95% CI did not include 1 were considered statistically significant. Analyses were conducted with the R software environment [31], using the '*survival*' package to create the regression models [32].

Sensitivity analysis

Seven sensitivity analyses were performed with adjusted model 3. First, because the substitution in serving sizes also entails some unspecified residual substitution of kcal from other dietary products, we repeated model 3 in an isocaloric milk and yogurt substitution analysis, modelled in 50 kcal day^{-1}, for comparison. All participants were censored after 7 years to evaluate the influence of unobserved dietary changes over time on associations between milk and yogurt substitutions and T2D risk. Differences in risk associations compared to the main analysis could indicate the occurrence of dietary changes over time. The first 2 years of follow-up were excluded to assess the possible influence of reverse causation. We included non-verified potential T2D cases that were excluded from the study population as incident cases ($n = 488$) to evaluate whether lack of validation of these cases has affected our results because likely the majority of participants in this group will actually be incident cases. Prevalent cases of cardiovascular disease, hypertension and participants with an increased cholesterol ratio (>5) were excluded because those conditions could have resulted in changes in dietary habits [33]. Considering the role of hypertension as a

confounder or potential mediator in the relationship between milk and yogurt substitution and T2D incidence can be discussed, the model was explored without adjustment for hypertension. We additionally excluded those participants with missing values on the potential mediators cholesterol ratio ($n = 1435$) and BMI ($n = 18$) to investigate study findings in the similar study population as used in model 4. In addition, the baseline characteristics of participants for the complete included study population (model 1-3) were compared with the population characteristics when the participants with missing values on the cholesterol ratio and BMI were additionally excluded (model 4).

For comparison, we investigated the association between the consumption of all milk and yogurt product subgroups and the risk of incident T2D individually, without specified substitutions and adjusted following model 3.

Results

Population characteristics

At baseline, the median milk and yogurt intake was 1.5 servings (P_{25}–P_{75}: 0.8–2.4) per day. Of this milk and yogurt intake, 0.5 servings (P_{25}–P_{75}: 0.2–1.0) were consumed as skimmed milk, 0.2 servings (P_{25}–P_{75}: 0.1–0.3) as whole-fat milk, 0.4 servings (P_{25}–P_{75}: 0.1–1.0) as buttermilk, 0.2 servings (P_{25}–P_{75}: 0.1–0.4) as skimmed fermented milk products and 0.1 servings (P_{25}–P_{75}: 0.0–0.2) as whole-fat yogurt (Table 1).

Two-thirds of the study population represents women. Light alcohol consumption and being physically active were more frequent among participants with a high milk and yogurt intake. Compared to the other milk and yogurt subgroups, high whole-fat yogurt consumers were more likely to be highly educated, less likely to smoke and consumed more fruit (Table 2; see also Supporting information, Tables S1 and S2). The consumption of whole-fat milk and skimmed milk was moderately correlated ($r = 0.76$), as were consumption of whole-fat yogurt and skimmed fermented milk ($r = 0.48$). Correlations between consumption of other milk and yogurt groups were low (all $r < 0.2$) (see Supporting information, Table S3).

Main results

During a mean follow-up of 15 years, 1467 (4.1%) potential incident cases of T2D were validated as an incident T2D case. After adjustment for demographic and T2D risk factors in the final adjusted model (model 3), replacement of whole-fat milk (HR = 0.93, 95% CI = 0.60-1.44), buttermilk (HR = 0.88, 95% CI = 0.58-1.34), skimmed milk (HR = 0.87, 95% CI = 0.57–1.32) or skimmed fermented milk (HR = 0.99, 95% CI = 0.63–1.54) with whole-fat yogurt was not associated with the risk of T2D (Figure 1; see also Supporting information, Table S4). Furthermore, replacing whole-fat milk (HR = 0.94, 95% CI = 0.77–1.15) or buttermilk (HR = 0.89, 95% CI = 0.77–1.04) with skimmed fermented milk, replacing whole-fat milk (HR = 1.07, 95% CI = 0.89–1.29) or skimmed fermented milk (HR = 1.14, 95% CI = 0.98–1.32) with skimmed milk, or replacing whole-fat milk (HR = 1.06, 95% CI = 0.90–1.24) or skimmed milk (HR = 0.99, 95% CI = 0.90–1.08) with buttermilk was also not associated with the risk of T2D. Additional adjustment for potential mediators did not affect the results (see Supporting information, Table S4).

Sensitivity analysis

The isocaloric substitution analysis did not alter conclusions (see Supporting information, Table S5). When censoring after 7 years of follow-up ($n = 35\,981$; 738 cases), associations for substitution with whole-fat yogurt products and reduced T2D risk strengthened because replacing buttermilk (HR = 0.48, 95% CI = 0.26–0.89), skimmed

Table 1 Total milk and yogurt consumption and consumption per milk and yogurt substitution subgroup* in the EPIC-NL cohort ($n = 35\,982$)

	Number of consumers	Daily consumption (g) Median (P_{25}–P_{75})	Daily consumption (servings)† Median (P_{25}–P_{75})
Total milk and yogurt	35 982	302 (158–482)	1.5 (0.8–2.4)
Milk and yogurt subgroups			
Skimmed milk	35 865	93 (31–199)	0.5 (0.2–1.0)
Whole-fat milk	33 468	34 (14–66)	0.2 (0.1–0.3)
Buttermilk	16 957	86 (14–200)	0.4 (0.1–1.0)
Skimmed fermented milk	34 868	31 (11–73)	0.2 (0.1–0.4)
Whole-fat yogurt	34 684	11 (5–31)	0.1 (0.0–0.2)

P, percentile.

*Applicable for the consumers of the specified milk or yogurt products.

†200 g.

Table 2 Baseline characteristics of participants from the EPIC-NL cohort by tertiles of skimmed milk and whole-fat yogurt consumption (n = 35 982)

	Skimmed milk (g day^{-1})			Whole-fat yogurt (g day^{-1})		
	T_1 18 (8–30)	T_2 93 (65–135)	T_3 295 (198–360)	T_1 2 (1–4)	T_2 11 (8–15)	T_3 41 (30–60)
Number of participants	11 994	11 994	11 994	11 994	11 994	11 994
Female, % (n)	78 (9333)	72 (8606)	74 (8875)	70 (8377)	76 (9079)	78 (9358)
Age at recruitment (years), mean (SD)	50 (11)	48 (12)	49 (12)	48 (12)	49 (12)	50 (12)
Hypertension, % (n)	38 (4592)	35 (4172)	37 (4390)	38 (4562)	36 (4334)	36 (4258)
Total HDL to cholesterol ratio, mean (SD)	4.0 (1.4)	4.1 (1.4)	4.1 (1.5)	4.2 (1.5)	4.0 (1.4)	4.0 (1.4)
CVD, % (n)	1 (153)	1 (138)	1 (154)	2 (184)	1 (146)	1 (115)
High education*, % (n)	22 (2,573)	20 (2,374)	20 (2,438)	17 (2,048)	20 (2,407)	24 (2,930)
Physically active, % (n)	39 (4724)	42 (5035)	45 (5352)	39 (4651)	42 (5075)	45 (5385)
Current smoker, % (n)	30 (3636)	32 (3882)	29 (3415)	40 (4,797)	28 (3,320)	24 (2,816)
Former smoker, % (n)	33 (4010)	30 (3609)	31 (3669)	29 (3492)	32 (3863)	33 (3933)
Light alcohol drinkers[†], % (n)	59 (7115)	63 (7533)	66 (7961)	56 (6742)	65 (7778)	67 (8089)
Body mass index (kg m^{-2}), mean (SD)	25.6 (4.1)	25.5 (3.8)	25.7 (3.9)	26.0 (4.1)	25.7 (4.0)	25.2 (3.7)
Total energy intake (kcal day^{-1}), mean (SD)	1889 (554)	2091 (611)	2183 (616)	1981 (618)	2050 (602)	2132 (591)
Milk and yogurt intake (g day^{-1}), median (P_{25}–P_{75})						
Buttermilk	0 (0–200)	0 (0–57)	0 (0–33)	0 (0–20)	0 (0–86)	7 (0–143)
Whole-fat milk	5 (1–15)	29 (17–49)	63 (43–91)	25 (6–58)	32 (12–63)	36 (14–70)
Skimmed milk	18 (8–30)	93 (65–135)	295 (198–360)	63 (17–165)	105 (35–209)	117 (42–227)
Skimmed fermented milk	22 (6–70)	29 (10–68)	36 (14–77)	6 (2–25)	49 (19–92)	44 (21–83)
Whole-fat yogurt	8 (2–24)	11 (4–30)	15 (6–35)	2 (1–4)	11 (8–15)	41 (30–60)
Total milk and yogurt	136 (56–300)	249 (174–386)	479 (369–641)	197 (70–376)	314 (180–491)	383 (260–551)
Other dietary intake (g day^{-1}), median (P_{25}–P_{75})						
Vegetables	131 (100–170)	129 (101–164)	132 (102–167)	125 (94–163)	131 (102–167)	135 (107–170)
Fruits	238 (129–360)	229 (133–341)	251(150–363)	193 (110–316)	247 (146–359)	260 (163–373)
Coffee	450 (180–625)	450 (225–675)	450 (218–562)	450 (210–675)	450 (210–650)	450 (225–562)
Red meat	60 (33–85)	63 (36–86)	59 (34–83)	65 (37–88)	60 (34–84)	57 (33–83)
Processed meat	19 (8–36)	22 (11–38)	20 (10–36)	24 (11–43)	20 (10–35)	19 (9–34)
SSBs	29 (6–83)	33 (8–90)	30 (7–82)	39 (7–102)	30 (7–82)	26 (7–69)
Fibre[‡]	23 (20–27)	23 (20–26)	23 (20–26)	22 (19– 26)	23 (20–26)	24 (21–27)

CVD, cardiovascular disease; HDL, high-density lipoprotein; P, percentile; SSBs, sugar-sweetened beverages; *T*, tertile.

*Higher vocational education and university.

[†]0.1–10 g alcohol day^{-1}.

[‡]Energy adjusted.

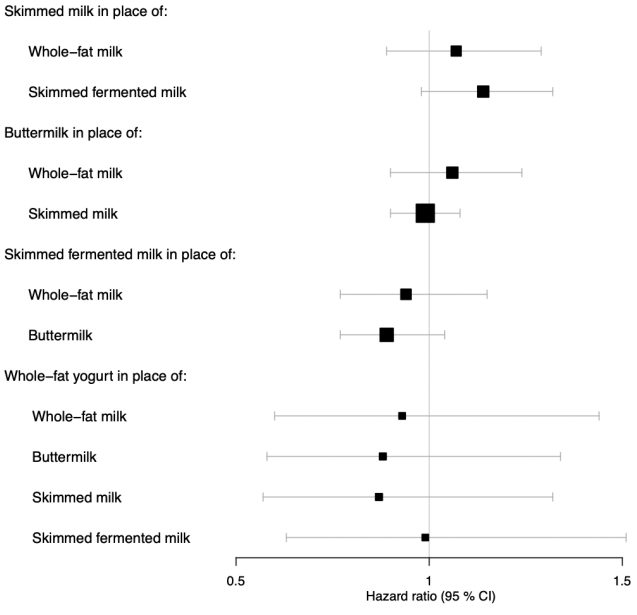

Figure 1 Forest plot of the hazard ratio and 95% confidence interval for substitution of one serving (200 g) of milk or yogurt and the association with type 2 diabetes in the EPIC-NL cohort (n = 35 982; 1467 cases), adjusted for sex (male, female), total energy intake (kcal/day; continuous), smoking status (current, former, never), physical activity (inactive, moderately inactive, moderately active and active), education level (low, average, high), alcohol intake (non-consumer, light, moderate, heavy), hypertension (yes, no) and the dietary intake of fruits, vegetables, processed meat, red meat, coffee, sugar-sweetened beverages and energy adjusted fibre (g day^{-1}; continuous) (model 3).

milk (HR = 0.44, 95% CI = 0.24–0.81) or skimmed fermented milk (HR = 0.48, 95% CI = 0.25–0.91) with whole-fat yogurt was associated with a lower T2D risk (see Supporting information, Table S6). Excluding the first 2 years of follow-up did not change the conclusions, whereas including potential but not verified T2D cases as incident case (n = 488) suggested a potential lower T2D risk for replacement with whole-fat yogurt, although the confidence intervals were wide and the results were not statistically significant (see Supporting information, Table S6). Excluding prevalent cases of cardiovascular disease, hypertension and participants with an increased cholesterol ratio (n = 18 104; 291 cases) resulted in estimates indicating a higher T2D risk for all substitutions, except the replacement of skimmed fermented milk with skimmed milk. However, the results were not statistically significant and the low number of cases and wide confidence intervals suggested low statistical power (see Supporting information, Table S6). When repeating model 3 which adjusted for demographic and T2D risk factors without adjustment for hypertension, and when subjects with missing values on the potential mediators were excluded, the results remained similar (see Supporting information, Table S7). The baseline characteristics after excluding participants with missing values on the potential mediators did not differ from the main study population (see Supporting information, Table S8).

Without specifying substitution, a higher consumption of skimmed milk (HR = 1.11, 95% CI = 1.04–1.19) and buttermilk (HR = 1.09, 95% CI = 1.02–1.16) was associated with increased T2D risk, whereas an increase in the consumption of whole-fat milk (HR = 1.09, 95% CI = 0.95–1.26), skimmed fermented milk (HR = 0.98, 95% CI = 0.86–1.12) and whole-fat yogurt (HR = 1.00, 95% CI = 0.67–1.51) was not (see Supporting information, Table S9).

Discussion

The present study investigated the association between the replacement of milk and yogurt products and the risk of incident T2D among a Dutch study population including 35 982 participants. During follow-up, 1467 (4.1%) validated T2D cases were identified. No evidence was found for an association between the replacement of milk and yogurt products and the risk of incident T2D in the main analysis. However, a lower risk of T2D was suggested when servings of buttermilk, skimmed milk and skimmed fermented milk were replaced by whole-fat yogurt when censoring the follow-up duration to 7 years.

One previous study among a Danish population investigated milk and yogurt product substitutions. In line with the present study, the Danish study did not observe associations with T2D risk when whole-fat milk replaced buttermilk, or skimmed milk replaced whole-fat milk and buttermilk, or skimmed fermented milk replaced skimmed milk, whole-fat milk or buttermilk. Yet, a lower risk of T2D was observed when one serving of whole-fat yogurt was used to replace one serving of skimmed milk (HR = 0.89, 95% CI = 0.83–0.96), whole-fat milk (HR = 0.89, 95% CI = 0.82–0.96), buttermilk (HR = 0.89, 95% CI 0.81–0.97) or skimmed fermented milk (HR = 0.83, 95% CI = 0.71–0.94) [19]. Although our effect estimates are similar to these previous findings, the confidence

intervals were wider and the results were not statistically significant.

The discrepancies between the study findings are not easily explained because there are no evident sources of heterogeneity between the two studies. One explanation for the discrepancy in findings regarding the whole-fat yogurt may be the lower statistical power in the present study as a result of a smaller study population, as well as a more stringent case definition resulting in a lower number of cases. Another explanation might be the potential influence of dietary changes regarding milk and yogurt product intake in our study population during the 15 years of follow-up. As a result of the use of a single FFQ, only baseline information on dietary intake was available. Furthermore, the sensitivity analysis where we censored after 7 years of follow-up suggested a lower T2D risk for the replacement of buttermilk, skimmed milk and skimmed fermented milk with whole-fat yogurt.

The results from observational studies investigating the association between whole-fat yogurt and T2D risk also report inconclusive results. Some individual cohort studies suggest an inverse association for higher whole-fat yogurt consumption and the risk of T2D [33-36], whereas a previous analysis in a Dutch cohort [37] and the current EPIC-NL analysis suggest a neutral association. It is possible that these discrepancies are driven by differences in adjustment for confounding factors, as well as differences between populations with respect to the food that is consumed instead of whole-fat yogurt. Furthermore, current results from randomised controlled trials investigating dairy product consumption and T2D risk markers often comprise short-term studies conducted in mostly overweight and obese participants, suggesting a null effect or small inverse effects [6]. Long-term experimental studies investigating causal effects between milk and yogurt consumption on intermediate risk markers for T2D (such as fasting glucose levels or insulin response) are necessary.

The present study has several strengths. First, we used a large study population with a long follow-up time with a small degree of loss to follow-up (1.7%). Baseline data collection was extensive, resulting in availability of a wide range of potential confounders. Also, we modelled the substitution of both servings and kilocalories. Finally, we were able to examine the consumption of different types of milk and yogurt because these were measured by the FFQ, and the Dutch population has a relatively high intake of various milk and yogurt products.

There are limitations to consider as well. First, using a FFQ to assess milk and yogurt product intake may have led to misclassification [38,39], although we have no reason to assume that this misclassification is differential and, when comparing collected FFQ data with 12-monthly 24-h recalls, reasonable correlation coefficients were found for the consumption of the total group of milk and milk products [22]. However, the consumption of our specific milk and yogurt subgroups has not been validated against 24-h recalls. Regarding the T2D ascertainment, we did not use the golden standard for diagnosing T2D (i.e. multiple tests of fasting plasma glucose levels) [40]. As an alternative, we used verification information from the general practitioner, who has a complete overview of the medical records, and from the pharmacist, who has information on T2D medication, which is very specific. Furthermore, the substitution model takes a mathematical approach to compare participants at various levels of milk and yogurt product intake, which is not the same as a within-person comparison over time. Repeated measurements of dietary intake would have provided the opportunity to examine milk and yogurt substitution within the same person, although this information is not available. Finally, although we adjusted for a wide range of potential confounders, the possibility of residual confounding cannot be excluded because participants with a higher intake of milk and yogurt products (especially whole-fat yogurt) showed healthier lifestyle behaviours.

In conclusion, we did not find evidence for an association between substitutions within the group of milk and yogurt products and the risk of incident T2D among a Dutch population. Our results therefore indicate that there is no difference between milk and yogurt consumption and the development of T2D. However, we cannot exclude possible attenuation of our results as a result of dietary changes over time. To further clarify the association of milk and yogurt products and T2D risk, this analysis should be repeated in a population with a wider consumption range of whole-fat yogurt to improve the generalisability of the study findings, including follow-up data on the dietary intake. Whole-fat yogurt appears to be particularly relevant in affecting T2D risk and our current analyses were limited by the small intake range of whole-fat yogurt. Swedish or French prospective cohorts may be eligible because these populations have a higher overall milk and yogurt consumption compared to our Dutch population [41]. Furthermore, long-term experimental studies investigating causal effects between milk and yogurt consumption on intermediate risk markers for T2D are needed.

Conflict of interests, source of funding and authorship

The authors declare that they have no conflicts of interest.

The EPIC-NL study was funded by 'European Commission: Public Health and Consumer Protection Directorate 1993–2004; Research Directory-General 2005'; Dutch Ministry of Public Health, Welfare and Sports (WVS), Netherlands Cancer Registry (NKR), LK Research

Funds, Dutch Prevention Funds, Dutch ZON (Zorg Onderzoek Nederland), and World Cancer Research Fund (WCRF) (The Netherlands). I. Sluijs was supported by a personal Dr Dekker postdoctoral grant (2015T019) from the Netherlands Heart Foundation.
IS and LV formulated the research question. JS, LV and IS designed the study, carried it out, analysed the data and wrote the article. All authors provided critical feedback and helped to shape the research and analysis, as well as the final manuscript.

Transparency Declaration

The lead author affirms that this manuscript is an honest, accurate and transparent account of the study being reported. The reporting of this work is compliant with STROBE guidelines. The lead author affirms that no important aspects of the study have been omitted and that any discrepancies from the study as planned have been explained.

REFERENCES

1. Ogurtsova K, da Rocha Fernandes JD, Huang Y *et al.* (2017) IDF diabetes Atlas: Global estimates for the prevalence of diabetes for 2015 and 2040. *Diabetes Res Clin Pract* **128**, 40-50.
2. Chatterjee S, Khunti K & Davies MJ (2017) Type 2 diabetes. *Lancet* **389**, 2239-2251.
3. Ley SH, Hamdy O, Mohan V *et al.* (2014) Prevention and management of type 2 diabetes: dietary components and nutritional strategies. *Lancet* **383**, 1999-2007.
4. Walker KZ, O'Dea K, Gomez M *et al.* (2010) Diet and exercise in the prevention of diabetes. *J Hum Nutr Diet* **23**, 344-352.
5. Gijsbers L, Ding EL, Malik VS *et al.* (2016) Consumption of dairy foods and diabetes incidence: a dose-response meta-analysis of observational studies. *Am J Clin Nutr* **103**, 1111-1124.
6. Guo J, Givens DI, Astrup A *et al.* (2019) The impact of dairy products in the development of type 2 diabetes: where does the evidence stand in 2019? *Adv Nutr* **10**, 1066-1075.
7. Fernandez MA, Panahi S, Daniel N *et al.* (2017) Yogurt and cardiometabolic diseases: a critical review of potential mechanisms. *Adv Nutr* **8**, 812-829.
8. Veiga P, Pons N, Agrawal A *et al.* (2014) Changes of the human gut microbiome induced by a fermented milk product. *Sci Rep* **4**, 6328.
9. Layden BT, Yalamanchi SK, Wolever TM *et al.* (2012) Negative association of acetate with visceral adipose tissue and insulin levels. *Diabetes Metab Syndr Obes* **5**, 49-55.
10. Smedman AE, Gustafsson IB, Berglund LG *et al.* (1999) Pentadecanoic acid in serum as a marker for intake of milk fat: relations between intake of milk fat and metabolic risk factors. *Am J Clin Nutr* **69**, 22-29.
11. Krachler B, Norberg M, Eriksson JW *et al.* (2008) Fatty acid profile of the erythrocyte membrane preceding development of Type 2 diabetes mellitus. *Nutr Metab Cardiovasc Dis* **18**, 503-510.
12. Warensjo E, Jansson JH, Berglund L *et al.* (2004) Estimated intake of milk fat is negatively associated with cardiovascular risk factors and does not increase the risk of a first acute myocardial infarction. A prospective case-control study. *Br J Nutr* **91**, 635-642.
13. Warensjo E, Jansson JH, Cederholm T *et al.* (2010) Biomarkers of milk fat and the risk of myocardial infarction in men and women: a prospective, matched case-control study. *Am J Clin Nutr* **92**, 194-202.
14. Zemel MB & Zhao F (2009) Role of whey protein and whey components in weight management and energy metabolism. *Wei Sheng Yan Jiu* **38**, 114-117.
15. Bowen J, Noakes M & Clifton PM (2005) Effect of calcium and dairy foods in high protein, energy-restricted diets on weight loss and metabolic parameters in overweight adults. *Int J Obes (Lond)* **29**, 957-965.
16. Jones KW, Eller LK, Parnell JA *et al.* (2013) Effect of a dairy- and calcium-rich diet on weight loss and appetite during energy restriction in overweight and obese adults: a randomized trial. *Eur J Clin Nutr* **67**, 371-376.
17. Gilbert JA, Joanisse DR, Chaput JP *et al.* (2011) Milk supplementation facilitates appetite control in obese women during weight loss: a randomised, single-blind, placebo-controlled trial. *Br J Nutr* **105**, 133-143.
18. Boeing H (2013) Nutritional epidemiology: new perspectives for understanding the diet-disease relationship? *Eur J Clin Nutr* **67**, 424-429.
19. Ibsen DB, Laursen ASD, Lauritzen L *et al.* (2017) Substitutions between dairy product subgroups and risk of type 2 diabetes: the danish diet, cancer and health cohort. *Br J Nutr* **118**, 989-997.
20. Beulens JW, Monninkhof EM, Verschuren WM *et al.* (2010) Cohort profile: the EPIC-NL study. *Int J Epidemiol* **39**, 1170-1178.
21. Sluijs I, Van der AD, Beulens JW *et al.* (2010) Ascertainment and verification of diabetes in the EPIC-NL study. *Neth J Med* **68**, 333-339.
22. Ocke MC, Bueno-de-Mesquita HB, Goddijn HE *et al.* (1997) The Dutch EPIC food frequency questionnaire. I. Description of the questionnaire, and relative validity and reproducibility for food groups. *Int J Epidemiol* **26 Suppl 1**, S37-S48.
23. Kromhout D, Spaaij CJK, de Goede J *et al.* (2016) The 2015 Dutch food-based dietary guidelines. *Eur J Clin Nutr* **70**, 869-878.
24. Willett WC, Howe GR & Kushi LH (1229S) Adjustment for total energy intake in epidemiologic studies. *Am J Clin Nutr* **65**, 1220S-1228S, discussion 1229S-1231S.
25. Van der Heijden KH LJM, Langius JAE, Oosten HMV *et al.* (1996) *NEVO tabel: Nederlands*

voedingsstoffenbestand. Den Haag: Voorlichtingsbureau voor de voeding.
26. Pols MA, Peeters PH, Ocke MC *et al.* (1997) Estimation of reproducibility and relative validity of the questions included in the EPIC physical activity questionnaire. *Int J Epidemiol* **26**(Suppl 1), S181-189.
27. Wareham NJ, Jakes RW, Rennie KL *et al.* (2003) Validity and repeatability of a simple index derived from the short physical activity questionnaire used in the European Prospective Investigation into Cancer and Nutrition (EPIC) study. *Public Health Nutr* **6**, 407-413.
28. World Health Organizaton: International Classification of Diseases (1977) *Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death*. 1975 Revision (ICD-9). Geneva: International Classification of Diseases.
29. Thiebaut AC & Benichou J (2004) Choice of time-scale in Cox's model analysis of epidemiologic cohort data: a simulation study. *Stat Med* **23**, 3803-3820.
30. Zhang L, Qiao Q, Tuomilehto J *et al.* (2008) Blood lipid levels in relation to glucose status in European men and women without a prior history of diabetes: the DECODE Study. *Diabetes Res Clin Pract* **82**, 364-377.
31. RCoreTeam (2018) R: A language and environment for statistical computing. https://www.R-project.org/ (accessed March 2018).
32. Therneau T (2015) A Package for Survival Analysis in S. https://CRAN.R-project.org/package=survival (accessed March 2018).
33. Harrington JM, Dahly DL, Fitzgerald AP *et al.* (2014) Capturing changes in dietary patterns among older adults: a latent class analysis of an ageing Irish cohort. *Public Health Nutr* **17**, 2674-2686.
34. Hruby A, Ma J, Rogers G *et al.* (2017) Associations of dairy intake with incident prediabetes or diabetes in middle-aged adults vary by both dairy type and glycemic status. *J Nutr* **147**, 1764-1775.
35. Guasch-Ferre M, Becerra-Tomas N, Ruiz-Canela M *et al.* (2017) Total and subtypes of dietary fat intake and risk of type 2 diabetes mellitus in the Prevencion con Dieta Mediterranea (PREDIMED) study. *Am J Clin Nutr* **105**, 723-735.
36. Ericson U, Hellstrand S, Brunkwall L *et al.* (2015) Food sources of fat may clarify the inconsistent role of dietary fat intake for incidence of type 2 diabetes. *Am J Clin Nutr* **101**, 1065-1080.
37. Brouwer-Brolsma EM, van Woudenbergh GJ, Oude Elferink SJ *et al.* (2016) Intake of different types of dairy and its prospective association with risk of type 2 diabetes: the rotterdam study. *Nutr Metab Cardiovasc Dis* **26**, 987-995.
38. Kristal AR & Potter JD (2006) Not the time to abandon the food frequency questionnaire: counterpoint. *Cancer Epidemiol Biomarkers Prev* **15**, 1759-1760.
39. Kipnis V, Midthune D, Freedman L *et al.* (2002) Bias in dietary-report instruments and its implications for nutritional epidemiology. *Public Health Nutr* **5**, 915-923.
40. American Diabetes A (2017) 2. Classification and diagnosis of diabetes. *Diabetes Care* **40**, S11-S24.
41. Hjartaker A, Lagiou A, Slimani N *et al.* (2002) Consumption of dairy products in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort: data from 35 955 24-hour dietary recalls in 10 European countries. *Public Health Nutr* **5**, 1259-1271.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Baseline characteristics of participants from the EPIC-NL cohort by tertiles of whole-fat milk and skimmed fermented milk consumption (n = 35 982).
Table S2. Baseline characteristics of participants from the EPIC-NL cohort by tertiles of buttermilk consumption (n = 35 982).
Table S3. Correlation between milk and yogurt consumption in the EPIC-NL cohort (n = 35 982).
Table S4. Substitution among milk and yogurt products per one servinga and the associationb with T2D in the EPIC-NL cohort (n = 35 982; 1467 cases).
Table S5. Sensitivity analyses with isocaloric substitution among milk and yogurt products and the association[a] with T2D in the EPIC-NL cohort (n = 35 982; 1467 cases).
Table S6. Sensitivity analyses for the substitution among milk and yogurt products per one serving[a] and the associationb with T2D in the EPIC-NL cohort.
Table S7. Sensitivity analyses for the substitution among milk and yogurt products per one serving[a] and the association[b] with T2D in the EPIC-NL cohort.
Table S8. Baseline characteristics of participants from the EPIC-NL cohort for all participants of the included study population (model 1–3) and when additionally excluding participants with missing values on the potential mediators[a] (model 4).
Table S9. Milk and yogurt consumption in servings[a] per day and the association[b] with T2D in the EPIC-NL cohort (n = 35 982; 1467 cases).