Eur J Epidemiol
DOI 10.1007/s10654-017-0328-x

MORTALITY

Lactase persistence, milk intake, and mortality in the Danish general population: a Mendelian randomization study

Helle Kirstine Mørup Bergholdt[1,2] · Børge Grønne Nordestgaard[2,3,4] · Anette Varbo[2,3] · Christina Ellervik[2,5,6]

Received: 10 May 2017 / Accepted: 20 October 2017

Abstract Meta-analyses have suggested no association between milk intake and mortality. Since only few studies have been conducted, we investigated the association between the lactase persistent genetic variant *LCT-13910 C/T* (rs4988235), a proxy for long-term low and high intake of milk, and mortality. We used two Danish population-based studies with self-reported intake of milk and genotyping for *LCT-13910 C/T*. We obtained information on all-cause and cause-specific mortality (cardiovascular and cancer) from the national Danish registries. We used multivariable adjusted Cox regression to assess the association between milk intake and mortality in 74,241 individuals, and both logistic and Cox-regression to assess the association between genetic lactase persistence and mortality in 82,964 individuals using a Mendelian randomization design. We applied per T-allele, co-dominant and dominant models. During a mean follow-up of 7 years, 9759 individuals died, 2166 from cardiovascular disease, and 2822 from cancer. Observationally, there was no association between intake of skimmed milk and all-cause or cardiovascular mortality, and we did not find any associations between intake of semi-skimmed or whole milk with all-cause or cause-specific mortality. Intake of skimmed milk was associated with lower cancer mortality with a hazard ratio of 0.97 (95% CI 0.96–1.00) per doubling in milk intake. Per T-allele, milk intake increased with 0.58 (0.50–0.68) glasses/week. Genetically, we found no associations between the lactase persistent *LCT-13910 C/T* genotype and all-cause or cause-specific mortality; per *T*-allele OR (95% CI) for all-cause mortality was 1.02 (0.97–1.06). Our study did not provide strong evidence of observational or genetic associations between milk intake and all-cause or cause-specific mortality.

Keywords Dairy · Milk · Mortality · Lactose · LCT-13910 · Lactase persistence

Electronic supplementary material The online version of this article (doi:10.1007/s10654-017-0328-x) contains supplementary material, which is available to authorized users.

✉ Christina Ellervik
christina@ellervik.dk

[1] Department of Clinical Biochemistry, Naestved Hospital, Naestved, Denmark

[2] Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

[3] Department of Clinical Biochemistry and the Copenhagen General Population Study, Herlev and Gentofte Hospital, Copenhagen University Hospital, Copenhagen, Denmark

[4] The Copenhagen City Heart Study, Frederiksberg Hospital, Copenhagen University Hospital, Copenhagen, Denmark

[5] Department of Research, Nykoebing Falster Hospital, Nykoebing Falster, Denmark

[6] Department of Laboratory Medicine, Boston Children's Hospital, 300 Longwood Avenue, Boston, MA 02115, USA

Introduction

Genetic lactase persistence in Europe is estimated to have co-originated 7500 years ago with dairying among farmers [1]. Milk contains important macronutrients (protein, fat, and carbohydrates) and micronutrients (vitamins and minerals such as calcium). Several meta-analyses and large prospective studies have shown no association of milk intake with all-cause, cardiovascular or cancer mortality in dose–response analyses or with different fat types [2–8]; however, individual studies have been very heterogenous and have shown conflicting results. A null-finding in the

Published online: 25 October 2017

meta-analyses, does not preclude a true association, since observational studies can be confounded, and since the observational studies included were heterogenous. The ideal study to test if milk intake increases or decreases mortality, is a large, preferably randomized and controlled trial. Unfortunately, difficulties of ensuring long-term compliance with the necessary dietary restrictions in a randomized controlled trial makes such a study difficult, if not impossible, to perform. Another way to gain insight into the potential causal relationship between milk intake and mortality is to perform a Mendelian randomization study [9], using a genetic variant as proxy for long-term milk intake.

The *LCT-13910 C/T* genetic variant (rs4988235) is the mutation responsible for lactase persistence among adults of European ancestry [10]. Individuals with the *T*-variant are able to uphold activity of the lactase enzyme (lactase-phlorizin hydrolase), which breaks down lactose, the main carbohydrate in milk, enabling them to consume milk without gastro-intestinal symptoms. High prevalence of lactase persistence, that is *LCT-13910* genotypes *TC* and *TT*, is seen in northern Europe. Individuals with *LCT-13910* genotype *CC* have lower lactase enzyme activity (i.e. are lactase non-persistent) and may need to limit or avoid the intake of milk and dairy products, as maldigestion of lactose can lead to symptoms of lactose intolerance, such as intestinal gas, bloating, and diarrhea.

In a previous study of the *LCT-13910 C/T* genotype, we found a higher milk intake among lactase persistent individuals compared to lactase non-persistent individuals in a sample of the Danish general population [11, 12]. At gamete formation, individuals are essentially genetically randomized and thus using a genetic variant as a proxy for milk intake creates nature's own randomization study to long-term high versus low milk intake. Therefore, investigating the association between the *LCT-13910 C/T* genetic variant and mortality allows us to indirectly assess a potential causal effect of milk intake on mortality, as reverse causation is avoided and most confounding factors will be evenly distributed between individuals with different genotypes. A Swedish study investigating the Northern European variant *LCT-13910 C/T* variant (N = 7404) [3] and an American study investigating the Mediterranean rs3754686 variant [13] (N = 7185) did not find an increased mortality with lactase persistence. In this study, we investigated whether high milk intake is associated with risk of mortality in an observational study of 74,241 individuals from the general Danish population. However, as we cannot make causal inference based on evidence from observational studies, we also performed a Mendelian randomization study [9] in 82,964 Danish individuals, using the genetic variant *LCT-13910* as a proxy for long-term differences in milk intake.

Methods

The Copenhagen general population study and the Copenhagen city heart study

Studies were approved by institutional review boards and Danish ethical committees, and the investigations conform to the principles of the Declaration of Helsinki. Written informed consent was given by all individuals participating in the studies, and all individuals were white and of Danish descent.

We included a total of 82,964 individuals $\geq$ 20 years of age, with 8721 individuals from the 1991–1994 examination of the Copenhagen City Heart Study(CCHS) [14], and 74,243 from the Copenhagen General Population Study(CGPS) [14], recruited from 2003 to 2011. Both population studies had information from a general questionnaire, a health examination, and blood samples including samples for DNA analyses. Supplementary Fig. 1 illustrates the sampling procedure for the present study.

Intake of milk and dairy

Individuals from the CCHS did not report intake of milk or dairy products, so observational analyses including information on milk intake were done for the CGPS alone. In the CGPS, individuals reported their intake of non-fermented milk (average intake; glasses per week) in terms of whole milk (3.5% fat), semi-skimmed milk (0.5–1.5%), and skimmed milk (0.1–0.3%), as well as cheese consumption (times per week) in a general questionnaire. No other information on intake of dairy products were available, and we could therefore not investigate the effects of total dairy intake, nor could we explore differences between fermented and non-fermented milk products. In terms of the response to the questions on milk intake, we observed the following response rates: For whole milk, non-response was 6439/74,243 = 8.7%. For semi-skimmed milk, non-response was 3528/74,243 = 4.6%. For skimmed milk, non-response was 4762/74,243 = 6.4%.

We made a variable for total milk intake by adding information on total non-fermented milk intake (glasses/week of all three types of milk). Using the information on total milk intake, we also created three categorical variables, i.e. milk intake in approximate quintiles, any milk (no, yes), and type of milk (no milk intake, fat free (0.1–0.3%), and fatty (0.5–3.5%) milk). We also investigated the effect of a higher intake of whole milk, semi-skimmed milk, and skimmed milk, respectively, among individuals reporting intake of only that particular type of milk versus no milk intake (individuals reporting an intake

of more than one type of milk were not included). The four continuous variables (total milk intake, whole milk, semi-skimmed milk, and skimmed milk) were all log2-transformed to approach normal distribution, and effect estimates for the log2-transformed milk variables in the analyses thus expresses the effect on the outcome per doubling of milk intake.

Genotypes

All 82,964 individuals from the CCHS and the CGPS had blood samples drawn at the health examination, and were genotyped for the *LCT-13910 C/T*(rs4988235) genetic variant using TaqMan (Applied Biosystems). We sequenced thirty randomly selected samples in order to confirm genotyping. As the genotypes were not in Hardy–Weinberg equilibrium, we additionally genotyped a random subset of individuals (N = 15,880) using real-time PCR, followed by melting curve analysis (LightCycler instrument 1.0(Roche; Applied Science)) as previously described [12]. Results from this method confirmed the disequilibrium. Allele frequencies for the *T* allele and the *C* allele were 76 and 24%, respectively. Call rates were 99.9% due to reruns. We constructed and applied both a per allele genetic model (*CC, TC, TT*; treated as a continuous variable), a co-dominant model (*TT* and *TC* respectively vs. *CC*), and a dominant model (*TT* + *TC* vs. *CC*).

All-cause and cause-specific mortality

Using the unique identification number assigned to all individuals in Denmark with permanent residence (the Central Person Registry-number), we retrieved information on vital status (dead/alive) from the national Danish Civil Registration System [15] and cause of death from the national Danish Causes of Death Registry [16]. We followed each individual from the time of participation in the population study until one of the following events occurred: death, migration, or end of follow-up on November 14, 2014. Information on cause-specific mortalities (cardiovascular and cancer) were based on information on the primary cause of death as well as any underlying or contributing cause of death registered on the death certificate. Causes of deaths occurring before January 1 1994 were classified according to the International Classification of Diseases(ICD) version 8 (i.e. individuals from the Copenhagen City Heart Study), thereafter, ICD-10 codes were used: 400.0-448 and I00-I99 (cardiovascular), and 140.0-239.9 and C00-D48 (cancer).

Baseline covariates

We obtained self-reported information on the following baseline characteristics from the individuals in the CGPS: physical activity at work and in leisure time, alcohol intake, smoking status, education, household income, marital status, family history of myocardial infarction, diabetes, and/or cancer, intake of fruit, vegetables, fish, fast-food, soda-drinks, red meat(beef and veal), processed meat(deli meats), coffee, and tea, as well as use of antidiabetic, antihypertensive, and/or lipid-lowering treatments. Height and weight measures were taken at the health examination, and body mass index (BMI) was calculated as measured weight (kg) divided by measured height (meter) squared. Along with information on sex and age, these variables were included in multivariable adjusted observational analyses of milk intake and mortality, as we considered them potential confounding factors. The adjustments for confounding were based on *a priory* knowledge as opposed to performing a model search.

Response categories for the question regarding physical activity at work was "mainly sitting", "sitting/standing/walking", "walking and some lifting", "heavy bodywork". None-responders were categorized as "not in workforce". Response categories for leisure time physical activity was "mainly passive", "light activity 2–4 h/week", "light/moderate activity > 4 h/week", and "very active > 4 h/week". Alcohol intake was reported as units per week and subsequently divided into approximate quartiles: "0–3", "4–8", "9–15", and " >= 16". Smoking status included the categories "never", "previous", and "current". Fruit, vegetable, and fish intake were categorized as no intake, ≤ 4 times/week, 5–7 times/week, > 7 times/week, intake of fast-food were categorized as no intake, once per week, twice or more per week, and intake of soda drinks were categorized as no intake, < 7 0.5L bottles/week, or ≥ 7 0.5L bottles/week. Coffee and tea intake were reported as the average intake in cups per week, and meat intake was reported as the average times of intake per week, and were included as continuous adjustment variables.

Statistics

Analyses were performed using StataSE 12.0. We used Mann–Whitney U and Pearson Chi square tests in analyses of population characteristics, i.e. associations between population (CCHS and CGPS) and the continuous and categorical baseline variables.

We used Cox-regression for observational analyses of associations between milk intake and all-cause mortality, and for analyses of cause-specific endpoints: cardiovascular and cancer mortality (primary cause of death only). In

the Cox-regression models of the observational association between milk and mortality, we set "origin" (i.e. the time when a subject becomes at risk) to time of birth. We used "entry" as the date of the health examination when the subject was included in the population study and rescaled time to years (365.25 days in 1 year). We used age as time scale, providing automatic adjustment for age. In multivariable adjusted analyses, we adjusted for all baseline variables previously mentioned. In all primary analyses of milk intake, we applied log2 transformation which corresponds to a doubling of milk intake to account for the skewed distribution (Supplementary Fig. 2a and b). We investigated all first-order interactions between covariates and milk intake using likelihood ratio tests and tested the assumption of proportional hazards using Schoenfeld residuals.

In order to assess whether there might be a causal relationship between milk intake and risk of premature death, we investigated the association between the *LCT-13910 C/T* and mortality; thus, performing the most basic version of a Mendelian randomization study as proposed by Katan [17, 18]. It is well known that the *LCT-13910 C/T* genotype is associated with milk intake and thus the variant meets the first assumption of an instrumental variable [10]. In addition, the genetic variant must be independent of the confounding factors that influence the association between the exposure (i.e. milk) and the outcome (i.e. mortality). Therefore, we examined the association between the *LCT-13910 C/T* variant (three categories, i.e. CC, TC, and TT) and the potential confounding factors of the milk-mortality association (mentioned above in the section on baseline covariates) using Kruskal–Wallis test (for continuous covariates) and Chi square test (for categorical covariates). The third assumption that applies to a Mendelian randomization study states that the genetic variant affects the outcome only through the exposure. We consider that this assumption likely is met, as we have no knowledge of other pathways. We examined the association between *LCT-13910 C/T* and mortality using logistic regression and applying a per T-allele, a co-dominant, and a dominant genetic model. We performed unadjusted analyses as well as analyses adjusted for sex, age, population, and height. Height was included in the model to account for potential underlying population stratification bias [21, 22], as both between and within country variations in frequencies of *LCT*-alleles have been reported in European countries [23]. We investigated all first-order interactions using likelihood ratio tests. In case of an interaction ($P < 0.05$) between the *LCT*-genotypes and a covariate, we stratified for the covariate in question.

Sensitivity analyses

In sensitivity analyses, we analyzed the observational association between milk intake (untransformed) and cause-specific mortality (primary cause of death only) in individuals from the Copenhagen General Population Study. In addition, we investigated the observational associations between milk intake in different groupings and mortality (all-cause mortality as well as cardiovascular and cancer mortality (primary cause of death only)). Furthermore, we investigated the observational association between milk intake (continuous as well as in groupings) and cause-specific mortality combining information in both primary and secondary causes of death, thus increasing the number of cases.

In sensitivity analyses of the genetic association, we investigated the association between *LCT-13910 C/T* and mortality in each study separately. We also investigated the association between *LCT-13910 C/T* and mortality using Cox-regression in a per T-allele, a co-dominant, and a dominant model for both studies combined and separately. In Cox-regression, we set "origin" (i.e. the time when a subject becomes at risk) to time of birth. We used "entry" as the date of the health examination when the subject was included in the population study and rescaled time to years (365.25 days in 1 year).

Results

Table 1 shows characteristics of individuals from the CCHS and the CGPS. In both study populations, 55% were women. The median age was 60 years in the CCHS and 57 years in the CGPS. Compared to the CCHS, individuals from the CGPS reported being more active in their leisure time and fewer smoked; however, alcohol intake and BMI were higher in the CGPS compared to the CCHS. The median intake of milk in the CGPS was 5 glasses/week (interquartile range 0-10 glasses/week).

Milk intake

Figure 1 shows milk intake in the CGPS: 20,777 individuals reported no milk intake, 15,675 consumed only skimmed milk, 28,852 reported drinking only semi-skimmed milk, 3852 consumed only whole milk, and 5085 individuals reported drinking more than one type of milk. Among individuals drinking skimmed milk only, the median(IQR) intake was 7 [5–14] glasses/week. For individuals drinking only semi-skimmed or whole milk, the median(IQR) intake was 7 [4–10] and 5 [2–8] glasses/week.

Table 1 Population characteristics in participants from the Copenhagen City Heart Study and the Copenhagen General Population Study

Characteristics	Population			
	CCHS		*CGPS*	
	N	%/ Median (IQR)	N	%/ Median (IQR)
Sex				
Women	4810	55.2	41,109	55.4
Men	3911	44.9	33,134	44.6
Age, years	8721	60 (48–70)	74,243	57 (47–66)
Physical activity in leisure time				
Mainly passive	1015	11.6	4875	6.6
Light activity 2–4 h/week	4696	53.9	32,738	44.1
Light/moderate activity > 4 h/week	2711	31.1	32,079	43.2
Very active > 4 h/week	299	3.4	4551	6.1
Smoking				
Never smoker	2144	24.6	29,852	40.2
Former smoker	2296	26.3	29,774	40.1
Current smoker	4281	49.1	14,617	19.7
Education[a]				
None/student	2046	23.5	8031	10.8
Practical	4106	47.1	25,428	34.3
Short, < 3 year	1096	12.6	9010	12.1
Middle, 3–4 years	621	7.1	19,025	25.6
Long, ≥ 5 years	852	9.8	12,749	17.2
Marital status				
Married/cohabiting	4947	56.7	54,613	73.6
Unmarried	1321	15.2	6067	8.2
Seperated/divorced	1152	13.2	7429	10.0
Widow/widower	1249	14.3	5976	8.1
Unknown status	52	0.6	158	0.2
Household income, DKK				
< 100.000		NA	1311	1.8
100.000–199.000		NA	9465	12.8
200.000–399.000		NA	18,049	24.3
400.000–599.000		NA	15,502	20.9
600.000–799.000		NA	13,247	17.8
≥ 800.000		NA	15,676	21.1
Do not wish to answer			993	1.3
Alcohol intake, units/week	8721	6 (1-13)	74,243	8 (3-15)
0–3	3421	39.2	20,733	27.9
4–8	2020	23.2	19,000	25.6
9–15	1580	18.1	17,173	23.1
≥ 16	1700	19.5	17,337	23.4
Fruit				
No/almost never		NA	2742	3.7
≤ 4 times/week		NA	18,771	25.3
5–7 times per week		NA	25,284	34.1
> 7 times per week		NA	27,446	37.0
Vegetables				
No/almost never		NA	3010	4.1

Table 1 continued

Characteristics	Population			
	CCHS		*CGPS*	
	N	%/ Median (IQR)	N	%/ Median (IQR)
≤ 4 times per week		NA	20,859	28.1
5–7 times per week		NA	29,089	39.2
> 7 times per week		NA	21,285	28.7
Soda drinks		NA	74,243	0 (0–2)
No/almost never		NA	45,919	61.9
< 7 0.5L bottles/week		NA	23,694	31.9
≥ 7 0.5L bottles/week		NA	4630	6.2
Coffee, cups/week		NA	74,243	15 (7–28)
Te, cups/week		NA	74,243	2 (0–10)
Milk, glasses/week		NA	74,243	5 (0–10)
LCT-13910				
CC (lactase non-persistent)	548	6.3	4348	5.9
TC (lactase persistent)	3041	34.9	26,571	35.8
TT (lactase persistent)	5132	58.8	43,324	58.4
Family history of myocardial infarction				
No		NA	52,129	70.2
Yes		NA	19,137	25.8
Unknown		NA	2977	4.0
Family history of diabetes				
No		NA	56,878	76.6
Yes		NA	14,145	19.1
Unknown		NA	3220	4.3
Family history of cancer				
No		NA	34,867	47.0
Yes		NA	36,599	49.3
Unknown		NA	2777	3.7
BMI, kg/m^2	8721	25.1 (22.6–28.1)	74,243	25.6 (23.2–28.5)
Total cholesterol, mmol/L	8721	6.1 (5.3–6.9)	74,243	5.6 (4.9–6.3)
Low-density lipoprotein cholesterol, mmol/L	8721	3.7 (3.0–4.4)	74,243	3.2 (2.6–3.9)
High-density lipoprotein cholesterol, mmol/L	8721	1.5 (1.2–1.9)	74,243	1.6 (1.2–1.9)
Triglyceride, mmol/L	8721	1.5 (1.1–2.2)	74,243	1.4 (1.0–2.1)
Glucose, mmol/L	8721	5.4 (5.0–6.1)	74,243	5.1 (4.7–5.7)
Systolic blood pressure, mmHg	8721	136 (123–153)	74,243	137 (124–151)
Diastolic blood pressure, mmHg	8721	84 (76–92)	74,243	80 (74–89)

BMI body mass index, *CCHS* Copenhagen city heart study, *CGPS* Copenhagen general population study, *IQR* interquartile range, *NA* information not available

[a] The education variable indicates the level of education obtained since the individual left the mandatory 7–9 years of lower and middle school education

Milk intake and mortality

Figure 2 shows results from observational analyses of associations between intake of different types of milk and all-cause, cardiovascular, and cancer mortality in the CGPS with a mean follow-up of 7 years (range 0–11 years). In multivariable adjusted analyses, we found an association between intake of skimmed milk and lower cancer mortality, with a hazard ratio(HR) of 0.97(95% CI 0.96–1.00, P value = 0.002) per doubling in milk intake. However, we found no association between intake of skimmed milk and cardiovascular mortality or all-cause mortality, and we

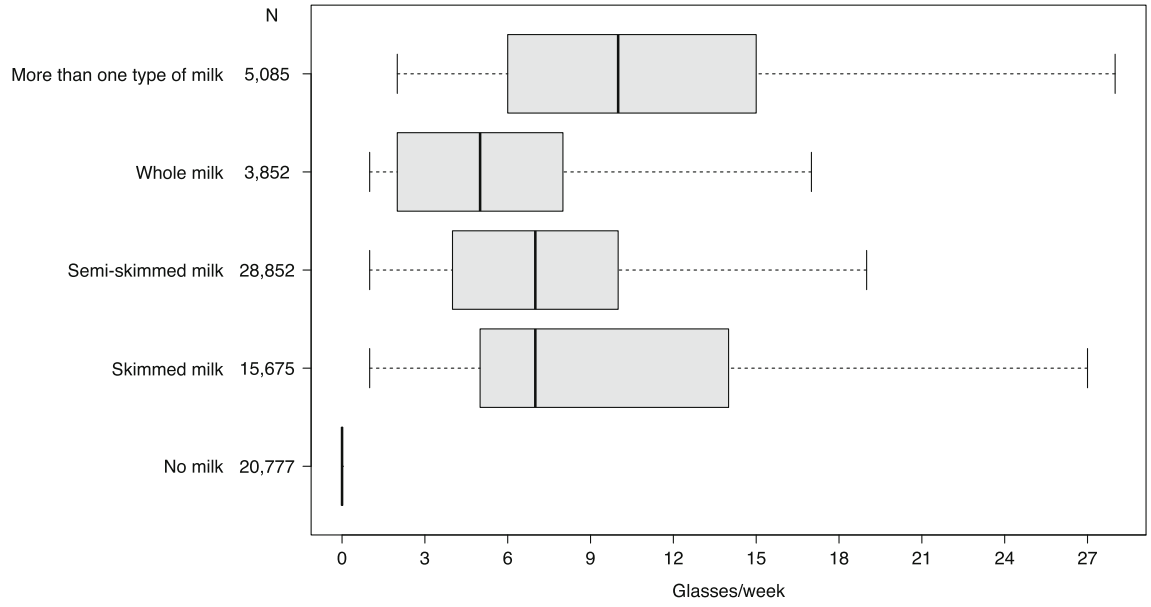

Fig. 1 Milk intake in individuals from the Copenhagen General Population Study. Milk intake (glasses/week) among 74,241 individuals from the Copenhagen General Population Study. The figure shows median intake with interquartile range for individuals reporting no intake, intake of skimmed milk only, intake of semi-skimmed milk only, intake of whole milk only, or intake of > 1 type of milk

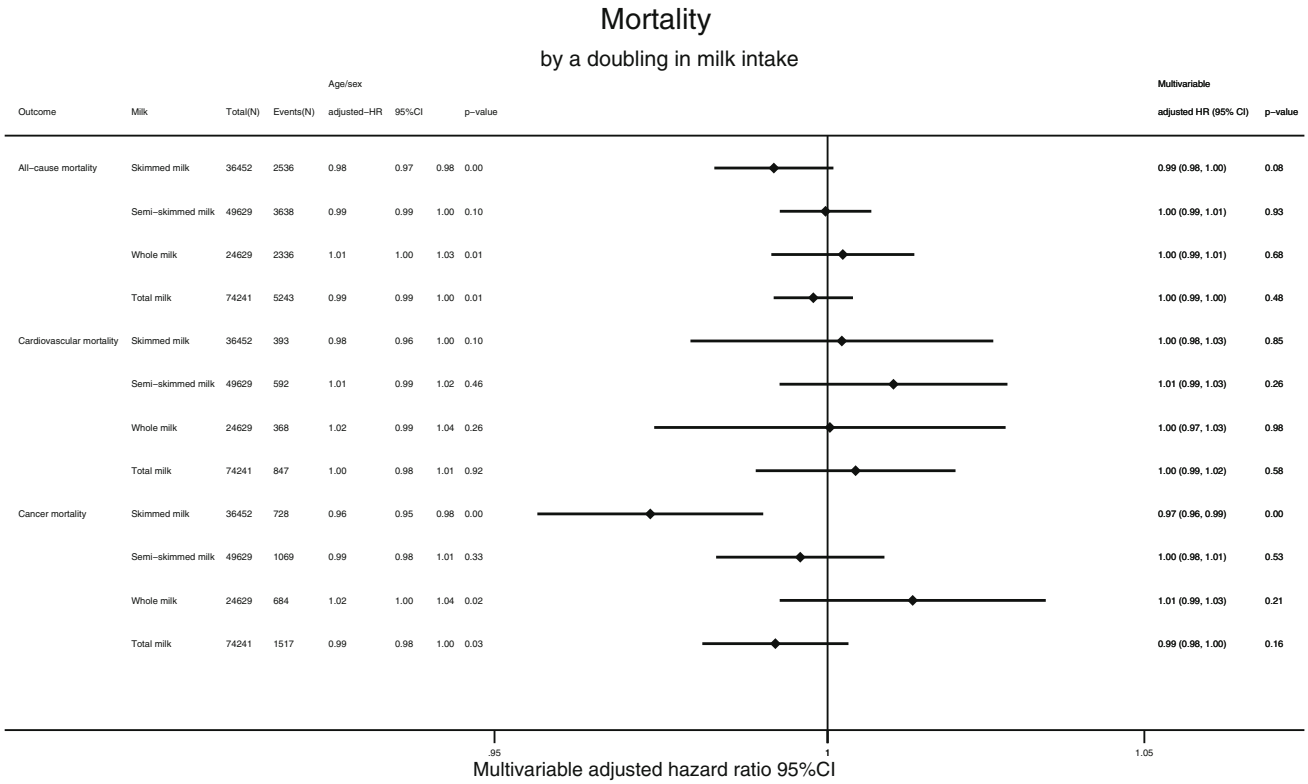

Fig. 2 Associations between milk intake and mortality. Risk of all-cause, cardiovascular, and cancer mortality for higher intake of skimmed, semi-skimmed, whole, and total milk, respectively, in individuals from the Copenhagen General Population Study followed for a mean of 7 years (range 0–11 years). Milk intake was log2-transformed and estimates thus correspond to a doubling of milk intake. Estimates of effect of intake of skimmed, semi-skimmed, and whole milk, respectively, include individuals reporting intake of only that particular type of milk or no milk intake. CI = confidence interval, N = number

did not find any associations between intake of semi-skimmed or whole milk and cancer, cardiovascular or all-cause mortality. Results for analyses adjusted for age and sex only showed similar results.

Sensitivity analyses for observational associations

In Fig. 2, milk intake was log2 transformed, but when using milk intake untransformed, results were similar (Supplementary Fig. 3). Supplementary Table 1 shows the association between milk intake in different groupings (quintiles, no/yes, and type of milk based on fat content) and mortality. We found associations between intake of fat free milk and lower all-cause mortality and cancer mortality compared to individuals drinking no milk.

Supplementary Table 2 shows the association between milk intake (log2 transformed, untransformed, and in groupings) with cardiovascular and cancer mortality when combining information on both primary and secondary cause of death. For cardiovascular mortality, only the multivariable adjusted estimate for total milk (untransformed) had a P value < 0.05 and showed slightly higher risk of cardiovascular mortality with higher intake of total milk with a HR of 1.01 (1.00–1.01, P value $= 0.03$). For cancer mortality, we found an association between higher intakes of skimmed milk (log2 transformed) and lower risk of cancer mortality with a HR of 0.98 (0.96–1.00, P value $= 0.01$), as well as lower risk associated with intake of fat free milk compared to no intake.

Lactase persistence and mortality

Among individuals from the CGPS, milk intake was higher in those with the lactase persistent genotype (median intake(IQR) of 5 (0–10) glasses/week among both *TC* & *TT*) compared to those with the lactase non-persistent genotype (2 (0–7) glasses/week among *CC,* P value $= 4 \times 10^{-70}$). Per T-allele, milk intake increased with 0.58 (0.50–0.68) glasses/week. We found no association between lactase persistence and all-cause, cardiovascular, or cancer mortality using *LCT-13910 C/T* genotypes, regardless of applying a per *T*-allele (OR for all-cause mortality (95% CI) was 1.02 (0.97–1.06, P value 0.51)), a co-dominant (Fig. 3), or a dominant genetic model (Supplementary Table 3). Results were similar when we investigated lactase persistence and association with cause-specific mortality supplementing the respective analysis with cases of underlying/contributing causes of cardiovascular and cancer mortality (Supplementary Table 3 and Supplementary Table 4). Also, no associations were found when we investigated lactase persistence and association with all-cause and cause-specific mortality in

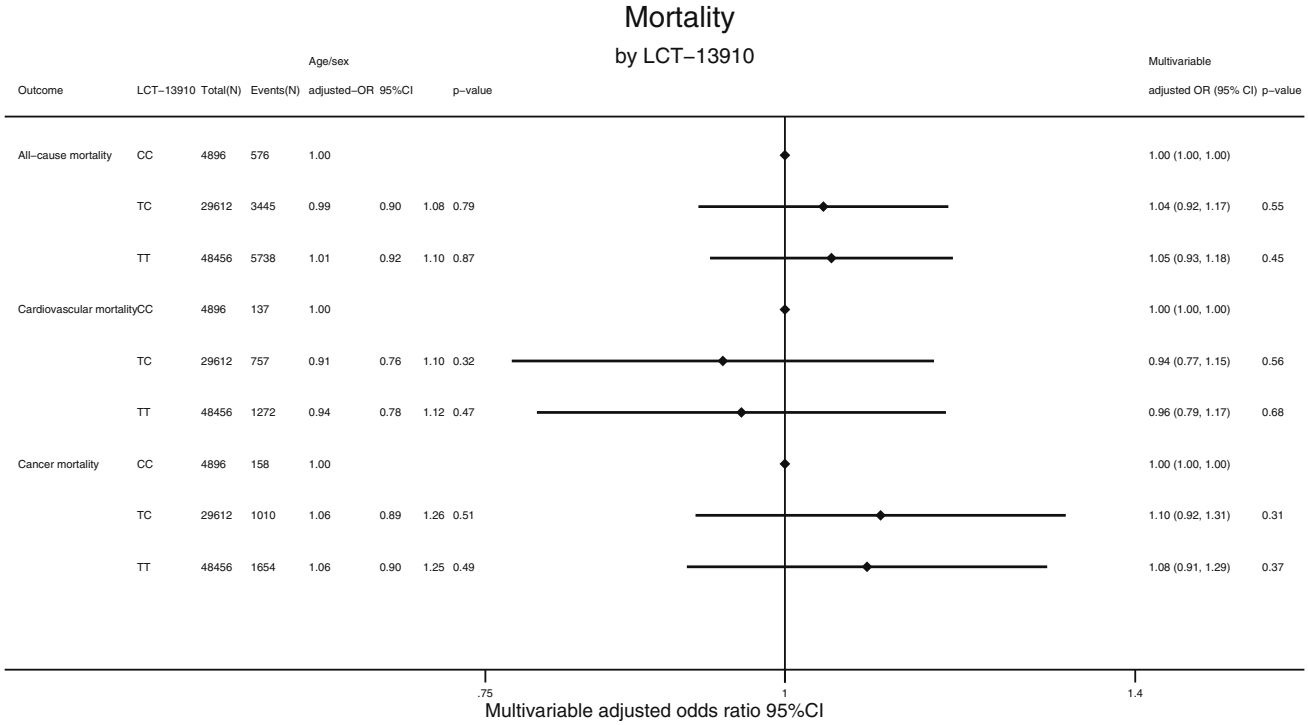

Fig. 3 Association between *LCT-13910 C/T* genotypes and mortality. Odds ratios for association of all-cause, cardiovascular, and cancer mortality with lactase persistent genotypes (*TC* and *TT*), compared to individuals with the lactase non-persistent genotype (*CC*). Individuals from the Copenhagen City Heart Study and the Copenhagen General Population Study were combined. Odds ratios were unadjusted (left panel) or multivariable adjusted for sex, age, population, and height (right panel). CI = confidence interval, N = number

individuals from the CCHS and the CGPS separately (Supplementary Table 3 and Supplementary Table 4). In addition, using Cox-regression analyses also yielded similar results (Supplementary Table 5); a HR (95% CI) of 1.00 (0.97–1.03, *P* value = 0.97) for all-cause mortality per *T*-allele.

Discussion

In this study, observational analyses of 74,241 individuals from the Danish general population demonstrated that consumption of milk was not convincingly associated with all-cause, cardiovascular, or cancer mortality; however, intake of skimmed milk was associated with a 1% lower risk of all-cause mortality and a 3% lower cancer mortality, compared to no intake of milk. Our results cannot rule out a possible small protective effect of intake of skimmed milk on cancer mortality; however, given how close our estimates in observational analyses were to 1.00, the effect, if any, would probably be very small. Also, we found no association between the genetic variant *LCT-13910 C/T* and all-cause, cardiovascular, or cancer mortality, indicating no causal association between milk intake and mortality in our cohorts.

Several meta-analyses and large prospective studies have shown no association of milk intake with all-cause, cardiovascular or cancer mortality [2–8], but individual studies have been very heterogenous and have shown conflicting results. However, meta-analyses [4] as well as recent large prospective studies in Sweden [3, 19] have shown that fermented milk products are associated with decreased risk of all-cause, cardiovascular, and cancer mortality as opposed to non-fermented milk products which were associated with increased risk [3]. It has been speculated that the difference between fermented and non-fermented milk products can be explained by lower intake of D-galactose from fermented milk products compared to non-fermented milk [19], based on evidence of accelerated aging in mice exposed to D-galactose [20]. This explanation has however been questioned [21] and it will likely require an updated investigation from the quoted study from the 1980's [22] into the D-galactose content in different dairy products to confirm the premise of this hypothesis. In accordance with our finding that genetic lactase persistence was not associated with mortality, two previous studies investigating the *LCT-13910 C/T* (rs4988235) variant [3] and the Mediterranean rs3754686 [13] variant, respectively, did not find an increased mortality with lactase persistence. Furthermore, emerging studies of genetic lactase persistence from large prospective studies and consortia studies cannot confirm associations with cardiovascular disease [12], blood pressure [23], Type 2 diabetes [11, 24], glycemic traits [24], or osteoporosis [24]. However, it seems plausible that genetic lactase persistence may increase BMI [11, 24].

In addition to D-galactose, other components in milk have been considered in general speculations about possible biological pathways from milk intake to disease and mortality. The content of saturated fat in milk is relatively high, and concerns have been raised that high milk intake therefore may lead to elevated cholesterol levels and thus increase the risk of cardiovascular disease. However, meta-analyses investigating food sources of saturated fat it was reported that there was no association between high intake of milk and dairy products and risk of cardiovascular-, cancer-, or all-cause mortality compared to individuals with low intake [2, 4]. In addition, we have recently investigated the association between milk intake and risk of ischemic heart disease and myocardial infarction in the same large Danish population as in this study, and we found no association, neither observational nor genetic [12].The content of minerals like calcium, potassium, and magnesium and vitamins such as vitamin B12 and riboflavin in milk may have beneficial health effects. For instance, studies suggest, that high intake of dairy products and dietary calcium has been associated with reduced risk of colon cancer [25].[1] It has been hypothesized that calcium has anti-carcinogenic properties that prevent colonic mutations and exert an anti-proliferative effect on colonic epithelium cells [26]. Interestingly, Dik et al. [26] did not find any effect of pre-diagnostic intake of dietary calcium or total dairy products on colorectal cancer-specific or all-cause death in European patients with colorectal cancer.

A major strength of our study was the high response rate on the milk questions and the large sample size; this study is the largest Mendelian randomization study to investigate associations between milk intake and mortality. We obtained information on self-reported milk intake using a general questionnaire reviewed by an investigator on the day of inclusion. Some misclassification of milk intake may have occurred due to recall bias, and we therefore included genetic analyses to better reflect long term differences in milk intake, avoiding the confounding that usually is a problem in observational analyses. A limitation of our study was the relative short follow-up time; however, our study results were in accordance with previous findings in meta-analyses of observational studies of milk intake and mortality with various follow-up times [2–8], with a Swedish study investigating the Northern European variant *LCT-13910 C/T* in 7404 individuals with 13 years of follow-up [3], and an American study investigating the Mediterranean rs3754686 variant [13] in 7185 individuals with 5 years of follow-up [3, 13]. While the CPR-number, assigned to all Danish citizens, and the Danish registries insures no losses to follow-up, misclassification of cause of

death may occur; however, while the analyses of cause-specific mortality may thus be somewhat biased due to misclassification, this is not the case for all-cause mortality. Our study isa Mendelian randomization study, but the study did not fulfill the requirements for the calculating the Mendelian instrumental variable estimates, as only the gene-milk association was significant but neither the gene-mortality association nor the milk-mortality association were significant. We confirmed however that the genotype was not associated with confounding factors unlike milk intake.

We restricted our analyses to the *LCT-13910 C/T* variant which is specific for populations of Northern European origin [10, 27], as our populations were selected on being of Danish descent. This information is based on Danish registries dating several generations back. Another lactase persistent variant found in populations of Mediterranean descent, rs3754686, and more widespread than the *LCT-13910 C/T* variant, likewise showed no association with mortality [13]. Other studies are still needed to confirm the findings of the *LCT-13910 C/T* and rs3754686 variants, and to investigate other lactase persistent variants with mortality.

We only estimated the *LCT-13910 C/T variant* as this is the variant completely associated with lactase persistence/non-persistence in Northern Europeans whereas the variant at position -22018 does not completely differentiates between lactase persistence/non-persistence [10]. Furthermore, as we restricted our population to only individuals of Danish descent, we did not find it meaningful to include other European or Middle Eastern variants from the MCM6 gene in a gene score [27]. Also, in a previous review, the particular variant *LCT-13910 C/T* is mentioned as ideal to investigate the lifelong effect of milk-exposure with outcomes in Northern Europeans [28]. Furthermore, many other MR articles about the genetically determined milk intake and outcomes are performed using only one variant [11, 12, 23, 24]. Lastly, since we selectively genotyped the *LCT-13910 C/T* variant and we do not have any chip-genotyping or imputation in the CCHS and CGPS, we could not study other variants.

Our genetic results showed no association between lactase persistence (proxy for higher milk intake) and mortality, however, the Mendelian randomization design does not take into account the type of milk consumed. There may be different health effects associated with intake of different types of milk e.g. full fat, fat reduced, lactose free, fermented, and organic milk, and this could be interesting to investigate in other studies.

In conclusion, our study did not provide strong evidence of an association between milk intake and mortality, and genetic analyses suggested no causal relationship. Overall, our results suggest that there is no large effect of milk intake on mortality in the general population, but further investigations into effects of different types of milk could be beneficial.

Acknowledgements We are grateful to all participants and staff in the Copenhagen General Population Study and the Copenhagen City Heart Study for their valuable contributions.

Funding The Copenhagen General Population Study was funded by The Danish Council for Independent Research; Medical Sciences(FSS); Herlev and Gentofte Hospital, Copenhagen University Hospital; Copenhagen County Foundation; and Chief Physician Johan Boserup and Lise Boserup's Fund, Denmark. The Copenhagen City Heart Study was funded by the Danish Heart Foundation. HKMB's PhD project was partly funded by the Danish Dairy Research Foundation; the Research Unit at Naestved Hospital, Denmark; and the Regional Research Unit in Region Zealand, Denmark. None of the funding sources were involved in the design, conduct of study, collection, management, analysis, or interpretation of data, or in preparation of the manuscript, and were not involved in the decision to submit the manuscript.

Authors contribution Authors responsible for this article are Helle KM Bergholdt(HKMB), Børge G Nordestgaard(BGN), Anette Varbo(AV) and Christina Ellervik(CE). BGN and CE conceived and designed the research and acquired the data. Literature search was performed by HKMB. HKMB prepared the data, performed statistical analyses, and drafted the manuscript. All authors undertook critical revisions of the manuscript and contributed intellectually to the development of this paper, as well as performed final approval of the paper. All authors had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data in the analyses, and affirm that the manuscript is an honest, accurate, and transparent account of the study being reported and that no important aspects of the study have been omitted.

Compliance with ethical standards

Conflicts of interest All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author). Dr Bergholdt reports grant from the Danish Dairy Research Foundation during the conduct of the study. Dr Nordestgaard, Dr Varbo, and Dr Ellervik report no conflicts of interest.

References

1. Itan Y, Powell A, Beaumont MA, Burger J, Thomas MG. The origins of lactase persistence in Europe. PLoS Comput Biol. 2009;5(8):e1000491. doi:10.1371/journal.pcbi.1000491.
2. O'Sullivan TA, Hafekost K, Mitrou F, Lawrence D. Food sources of saturated fat and the association with mortality: a meta-analysis. Am. J Public Health. 2013;103(9):e31–42. doi:10.2105/AJPH.2013.301492.
3. Tognon G, Nilsson LM, Shungin D, et al. Nonfermented milk and other dairy products: associations with all-cause mortality. Am J Clin Nutr. 2017;105(6):1502–11. doi:10.3945/ajcn.116.140798.
4. Guo J, Astrup A, Lovegrove JA, Gijsbers L, Givens DI, Soedamah-Muthu SS. Milk and dairy consumption and risk of cardiovascular diseases and all-cause mortality: dose–response meta-analysis of prospective cohort studies. Eur J Epidemiol. 2017;32(4):269–87. doi:10.1007/s10654-017-0243-1.

5. Soedamah-Muthu SS, Ding EL, Al-Delaimy WK, et al. Milk and dairy consumption and incidence of cardiovascular diseases and all-cause mortality: dose–response meta-analysis of prospective cohort studies. Am J Clin Nutr. 2011;93(1):158–71. doi:10.3945/ajcn.2010.29866.
6. Mullie P, Pizot C, Autier P. Daily milk consumption and all-cause mortality, coronary heart disease and stroke: a systematic review and meta-analysis of observational cohort studies. BMC Public Health. 2016;16(1):1236. doi:10.1186/s12889-016-3889-9.
7. Lu W, Chen H, Niu Y, Wu H, Xia D, Wu Y. Dairy products intake and cancer mortality risk: a meta-analysis of 11 population-based cohort studies. Nutr J. 2016;15(1):91. doi:10.1186/s12937-016-0210-9.
8. Larsson SC, Crippa A, Orsini N, Wolk A, Michaëlsson K. Milk consumption and mortality from all causes, cardiovascular disease, and cancer: a systematic review and meta-analysis. Nutrients. 2015;7(9):7749–63. doi:10.3390/nu7095363.
9. Lawlor DA, Harbord RM, Sterne JAC, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. Stat Med. 2008;27(8):1133–63. doi:10.1002/sim.3034.
10. Enattah NS, Sahi T, Savilahti E, Terwilliger JD, Peltonen L, Järvelä I. Identification of a variant associated with adult-type hypolactasia. Nat Genet. 2002;30(2):233–7.
11. Bergholdt HK, Nordestgaard BG, Ellervik C. Milk intake is not associated with low risk of diabetes or overweight-obesity: a Mendelian randomization study in 97,811 Danish individuals. Am J Clin Nutr. 2015;102(2):487–96. doi:10.3945/ajcn.114.105049.
12. Bergholdt HKM, Nordestgaard BG, Varbo A, Ellervik C. Milk intake is not associated with ischaemic heart disease in observational or Mendelian randomization analyses in 98 529 Danish adults. Int J Epidemiol. 2015;44(2):587–603. doi:10.1093/ije/dyv109.
13. Smith CE, Coltell O, Sorli JV, 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;6:33188. doi:10.1038/srep33188.
14. Ellervik C, Tybjaerg-Hansen A, Nordestgaard BG. Total mortality by transferrin saturation levels: two general population studies and a metaanalysis. Clin Chem. 2011;57(3):459–66. doi:10.1373/clinchem.2010.156802.
15. Pedersen CB. The Danish civil registration system. Scand J Public Health. 2011;39(7 Suppl):22–5. doi:10.1177/1403494810387965.
16. Helweg-Larsen K. The Danish register of causes of death. Scand J Public Health. 2011;39(7 Suppl):26–9. doi:10.1177/1403494811399958.
17. Katan MB. Apolipoprotein E isoforms, serum cholesterol, and cancer. Lancet. 1986;1(8479):507–8.
18. VanderWeele TJ, Tchetgen Tchetgen EJ, Cornelis M, Kraft P. Methodological challenges in mendelian randomization. Epidemiology. 2014;25(3):427–35. doi:10.1097/EDE.0000000000000081.
19. Michaëlsson K, Wolk A, Langenskiöld S, et al. Milk intake and risk of mortality and fractures in women and men: cohort studies. BMJ. 2014;349:g6015.
20. Cui X, Wang L, Zuo P, et al. D-galactose-caused life shortening in Drosophila melanogaster and Musca domestica is associated with oxidative stress. Biogerontology. 2004;5(5):317–25. doi:10.1007/s10522-004-2570-3.
21. Hettinga K. Study used wrong assumption about galactose content of fermented dairy products. BMJ. 2014;349:g7000.
22. Alm L. Effect of fermentation on lactose, glucose, and galactose content in milk and suitability of fermented milk products for lactose intolerant individuals. J Dairy Sci. 1982;65(3):346–52. doi:10.3168/jds.S0022-0302(82)82198-X.
23. Ding M, Huang T, Bergholdt HK, et al. Dairy consumption, systolic blood pressure, and risk of hypertension: Mendelian randomization study. BMJ. 2017;356:j1000. doi:10.1136/bmj.j1000.
24. Yang Q, Lin SL, Au Yeung SL, et al. Genetically predicted milk consumption and bone health, ischemic heart disease and type 2 diabetes: a Mendelian randomization study. Eur J Clin Nutr. 2017;71(8):1008–12. doi:10.1038/ejcn.2017.8.
25. Aune D, Lau R, Chan DS, et al. Dairy products and colorectal cancer risk: a systematic review and meta-analysis of cohort studies. Ann Oncol. 2012;23(1):37–45. doi:10.1093/annonc/mdr269.
26. Dik VK, Murphy N, Siersema PD, et al. Prediagnostic intake of dairy products and dietary calcium and colorectal cancer survival–results from the EPIC cohort study. Cancer Epidemiol Biomarkers Prev. 2014;23(9):1813–23. doi:10.1158/1055-9965.EPI-14-0172.
27. Itan Y, Jones BL, Ingram CJ, Swallow DM, Thomas MG. A worldwide correlation of lactase persistence phenotype and genotypes. BMC Evol Biol. 2010;10:36. doi:10.1186/1471-2148-10-36.
28. Schatzkin A, Abnet CC, Cross AJ, et al. Mendelian randomization: how it can-and cannot-help confirm causal relations between nutrition and cancer. Cancer Prev Res (Phila). 2009;2(2):104–13. doi:10.1158/1940-6207.CAPR-08-0070.