Milk intake, lactase non-persistence, and type 2 diabetes risk in Chinese adults

Maria G. Kakkoura[1], Robin G. Walters[1], Robert Clarke[1], Zhengming Chen[1], Huaidong Du[1]

[1] Clinical Trial Service Unit and Epidemiological Studies Unit (CTSU), Nuffield Department of Population Health, University of Oxford, UK

Address for correspondence:

Dr Huaidong Du, Clinical Trial Service Unit and Epidemiological Studies Unit (CTSU), Nuffield Department of Population Health, University of Oxford, Old Road Campus, Oxford OX3 7LF, UK. Telephone: +44 (0)1865 743831. Email: huaidong.du@ndph.ox.ac.uk,

Arising from Luo K. et al. *Nature Metabolism* https://doi.org/10.1038/s42255-023-00961-1 (2024)

In the 2024 January issue of *Nature Metabolism*, Luo et al.[1] reported an inverse association between milk intake and risk of type 2 diabetes (T2D) among individuals who were lactase non-persistent (LNP), as determined by the lactase (*LCT*) rs4988235 homozygous GG genotype. The authors also reported inverse cross-sectional associations of milk intake with several T2D-related metabolic traits, including body mass index (BMI) and waist circumference (WC) in LNP individuals. However, in their analyses, no adjustments for BMI/WC were done in assessing the association between milk intake and T2D risk, which may lead to biased results. Given the established causal relevance of adiposity for incident T2D[2,3] and the observed inverse associations of milk intake with adiposity in the study population, it is therefore important to establish whether the associations of milk consumption, milk-associated gut bacteria and milk-associated metabolites with T2D risk are confounded, or mediated, by BMI/WC in the paper by Luo et al.

In the China Kadoorie Biobank (CKB) of >0.5 million adults recruited from 10 diverse regions across China in 2004-2008, we previously reported a similar inverse association of dairy (chiefly milk) intake with BMI[4]. In the prospective analyses of 18,306 incident T2D cases in the CKB (i.e. 461,046 participants who were free ofT2D, cardiovascular diseases and cancer at baseline, and followed-up for an average of 11.8 years), we found a significant inverse association between dairy intake and T2D risk, with regular consumption of dairy products (≥4 days/week) being associated with a 7% (hazard ratio [HR]: 0.93; 95% confidence interval [CI]: 0.89-0.98) lower risk of T2D than non-consumers after adjustment for all available covariates in the models as used by Luo et al.[1], plus several others including family history of T2D. However, when BMI was additionally included in the models, the association of dairy intake with T2D risk disappeared (HR: 1.02; 95% CI: 0.97-1.07), indicating strong support for a confounding and/or mediating role of BMI.

Among the >100,000 genotyped CKB participants[5], the proportion of individuals with *LCT*-rs4988235 GG genotype (i.e. LNP individuals) was >99%, with the remainder having AG genotype (i.e. total absent of AA genotype). Interestingly, the frequency of AG genotype (lactase persistent,

LP) varied by 7-fold across the 10 geographical regions covered in CKB, with much higher frequency in the four north regions including Qingdao, Henan, Harbin and Gansu (0.16%-0.38%, mean 0.27%) than in the other six south regions (0.01-0.07%, mean 0.04%). Compared with the LP individuals, the LNP individuals indeed were less likely to be regular dairy consumers (12.8% vs 16.7%). Moreover, LNP individuals were also found to have lower daily milk consumption compared with LP individuals (33 vs 41 ml per day; 24% difference). Such a difference in CKB is highly consistent with the 23% difference reported by Luo et al. (i.e. 0.73 vs 0.90 serving per day in LNP vs LP group, respectively)[1], suggesting that this genetic variation alone does not fully explain the much lower consumption level of dairy products in Chinese than in European ancestry populations (e.g. our previous work identified a six-fold lower milk consumption in CKB than in UKB[6]).

Methods

Study design and population

Details on the study design, methods and participants of the CKB study have been previously described[7]. Briefly, it is a population-based prospective study of over 0.5 million adults recruited from ten diverse regions across China. The baseline survey was conducted during 2004-2008 in five urban and five rural regions, selected from China's nationally representative Disease Surveillance Points (DSP) system to capture a broad spectrum of socioeconomic factors, risk factor exposures and disease patterns. All permanent, non-disabled residents from preselected rural villages or urban committees in each study area were identified from local residential records and invited to participate. In total, 512,724 participants aged 30-79 years were enrolled (overall response rate ~30%). At local study assessment clinics, trained health professionals administered a laptop-based questionnaire recording information on sociodemographic characteristics, personal and family

medical history and lifestyle factors (e.g. diet, physical activity, smoking and alcohol consumption). Participants undertook physical measurements (e.g. blood pressure, anthropometry), a blood sample was collected from each of them for long-term storage and onsite blood tests including random plasma glucose levels (Johnson & Johnson SureStep Plus Meter) were conducted. Genotyping data were available for 100,640 participants and more details on genotyping and quality control have been described elsewhere[5]. Two resurveys of 5-6% of randomly selected surviving participants were conducted using procedures largely identical to those at baseline. Written informed consent was obtained from all participants. Ethical approval was obtained from the Ethical Review Committee of the Chinese Centre for Disease Control and Prevention (Beijing, China, 005/2004) and the Oxford Tropical Research Ethics Committee, University of Oxford (UK, 025-04).

Dietary intake assessment

Information on dietary intake was collected using a validated interviewer-administered laptop-based questionnaire. This questionnaire aimed to record the habitual consumption frequency (five categories including daily, 4-6 days/week, 1-3 days/week, monthly or never/rarely) of 12 major food groups (i.e. rice, wheat products, coarse grain products, red meat, poultry, fish, eggs, total dairy products, fresh vegetables, preserved vegetables, fresh fruit and soybean products) by the participants over the preceding year. The reproducibility and validity of the questionnaire was assessed by a separate validation study that was performed among 432 CKB participants during 2015-2016, with the use of 12-day (four consecutive days from three seasons) -24-hour recalls as reference. The weighted kappa statistics were 0.82 and 0.75 for reproducibility and validity of total dairy consumption frequency, demonstrating a good reproducibility and relative validity[8]. Furthermore, the second resurvey (2013-2014) that included ~25,000 participants collected daily portions of each food group and the intake (both frequency and amount) of three subtypes of dairy products, i.e., cow milk, yoghurt and other dairy products (e.g. cheese and milk powder)[4].

Follow-up for incident type 2 diabetes

The vital status of participants was obtained periodically via China's DSP system[9] (annual active follow-up through local residential records, health insurance data and administrative records actively confirmed by street committees or village administrators). Additionally, information on T2D incidence was collected via linkages with chronic disease registries (ischemic heart disease, stroke, cancer and T2D) and national health insurance claim databases. All events (fatal and non-fatal) were coded with the use of International Classification of Diseases, 10th Revision (ICD-10). Incident T2D for the present study included all recorders cases (ICD-10 E10-E14) that occurred between the ages of 35 and 79 years. By 1 January 2019 (global censoring date), only 3672 (0.8%) participants were lost to follow-up and they were censored in the prospective analyses.

Statistical analysis

Participants with baseline history of T2D, cardiovascular disease or cancer (n = 51,676) or those with missing values for BMI (n = 2) were excluded, leaving 461,046 participants in the main analyses. Cox regression was used to calculate HRs and 95% CIs for T2D incidence in relation to baseline frequency categories of dairy consumption, stratified by age-at-risk (in 5-year intervals), sex and region and adjusted for potential confounders including baseline age, education, annual household income, smoking, alcohol consumption, total physical activity, family history of T2D, consumption of fresh fruit, red meat, poultry, fish and eggs, hypertension (self-reported hypertension or mean systolic blood pressure ≥ 140 mmHg or mean diastolic blood pressure ≥ 90 mmHg or use of antihypertensive medication) and BMI. The proportional hazards assumption was examined by comparing the HRs for the first and second half of the follow-up period, with no violation being observed. To facilitate comparisons between any two exposure groups rather than just with an arbitrarily chosen reference group, the floating absolute risk method was used that

provides variance of log risk for each category (including the reference group)[10]. Analysis was performed for a two-sided hypothesis test. The statistical package SAS (version 9.4, SAS Institute, Cary, NC, USA) was used for performing the analyses.

References

1 Luo, K. *et al.* Variant of the lactase LCT gene explains association between milk intake and incident type 2 diabetes. *Nature Metabolism* **6**, 169-186 (2024). https://doi.org/10.1038/s42255-023-00961-1

2 Yuan, S. & Larsson, S. C. An atlas on risk factors for type 2 diabetes: a wide-angled Mendelian randomisation study. *Diabetologia* **63**, 2359-2371 (2020). https://doi.org/10.1007/s00125-020-05253-x

3 Yuan, S., Merino, J. & Larsson, S. C. Causal factors underlying diabetes risk informed by Mendelian randomisation analysis: evidence, opportunities and challenges. *Diabetologia* **66**, 800-812 (2023). https://doi.org/10.1007/s00125-023-05879-7

4 Kakkoura, M. G. *et al.* Dairy consumption and risks of total and site-specific cancers in Chinese adults: an 11-year prospective study of 0.5 million people. *BMC Medicine* **20**, 134 (2022). https://doi.org/10.1186/s12916-022-02330-3

5 Walters, R. G. *et al.* Genotyping and population characteristics of the China Kadoorie Biobank. *Cell Genom* **3**, 100361 (2023). https://doi.org/10.1016/j.xgen.2023.100361

6 Papier, K. *et al.* Intakes of major food groups in China and UK: results from 100,000 adults in the China Kadoorie biobank and UK biobank. *European Journal of Nutrition* **62**, 819-832 (2023). https://doi.org/10.1007/s00394-022-03031-6

7 Chen, Z. *et al.* China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. *Int J Epidemiol* **40**, 1652-1666 (2011). https://doi.org/10.1093/ije/dyr120

8 Qin, C. *et al.* The Relative Validity and Reproducibility of Food Frequency Questionnaires in the China Kadoorie Biobank Study. *Nutrients* **14** (2022). https://doi.org/10.3390/nu14040794

9 Yang, G. H., Stroup, D. F. & Thacker, S. B. National public health surveillance in China: implications for public health in China and the United States. *Biomed Environ Sci* **10**, 1-13 (1997).

10 Plummer, M. Improved estimates of floating absolute risk. *Stat Med* **23**, 93-104 (2004). https://doi.org/10.1002/sim.1485

Ethical approval and informed consent

The CKB complies with all the required ethical standards for medical research on human subjects. Ethical approvals were granted and have been maintained by the relevant institutional ethical research committees in the UK and China. Informed consent was obtained from all participants included in the study.

Data availability

The CKB is a global resource for the investigation of lifestyle, environmental, blood biochemical and genetic factors as determinants of common diseases. The CKB study group is committed to making the cohort data available to the scientific community in China, the UK and worldwide to

advance knowledge about the causes, prevention and treatment of disease. For detailed information on what data is currently available to open access users and how to apply for it, visit: https://www.ckbiobank.org/data-access. Researchers who are interested in obtaining the raw data from the China Kadoorie Biobank study that underlines this paper should contact ckbaccess@ndph.ox.ac.uk. A research proposal will be requested to ensure that any analysis is performed by bona fide researchers and - where data is not currently available to open access researchers - is restricted to the topic covered in this paper.

Acknowledgements

The chief acknowledgment is to the participants, the project staff, and the China National Centre for Disease Control and Prevention (CDC) and its regional offices for assisting with the fieldwork. We thank Judith Mackay in Hong Kong; Yu Wang, Gonghuan Yang, Zhengfu Qiang, Lin Feng, Maigeng Zhou, Wenhua Zhao, Yan Zhang in China CDC; Lingzhi Kong, Xiucheng Yu, and Kun Li in the Chinese Ministry of Health; and Garry Lancaster, Sarah Clark, Martin Radley, Mike Hill, Hongchao Pan, and Jill Boreham in the CTSU, Oxford, for assisting with the design, planning, organization, and conduct of the study. The CKB baseline survey and the first re-survey were supported by the Kadoorie Charitable Foundation in Hong Kong. The long-term follow-up has been supported by Wellcome grants to Oxford University (212946/Z/18/Z, 202922/Z/16/Z, 104085/Z/14/Z, 088158/Z/09/Z) and grants from the National Natural Science Foundation of China (82192900,82192901, 82192904,) and from the National Key Research and Development Program of China (2016YFC0900500).The UK Medical Research Council (MC_UU_00017/1,MC_UU_12026/2, MC_U137686851), Cancer Research UK (C16077/A29186; C500/A16896) and the British Heart Foundation (CH/1996001/9454), provide core funding to the Clinical Trial Service Unit and Epidemiological Studies Unit at Oxford University for the project. M.G.K was supported by the Wellcome Trust, Our Planet Our Health Programme (Livestock, Environment and People - LEAP) [grant number 205212/Z/16/Z]. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

M.G.K. analysed the data. M.G.K and H.D. drafted the manuscript. R.C., R.G.W. and Z.C. contributed to the writing of the manuscript. All authors critically reviewed the manuscript and approved the final submission.

Competing interests

The authors declare no competing interests.