nature portfolio

Corresponding author(s): COMMSMED-23-0591-B

Last updated by author(s): Feb 25, 2025

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

At baseline, trained health workers administered laptop-based questionnaires on socio-demographics, lifestyle behaviors, medical history and females' reproductive history, measured height, weight, waist circumference (WC), blood pressure, and collected blood samples. For long-term follow-up (2004-2015), all participants were followed by hospital admission, while disease diagnoses were collected through electronic linkage with local chronic disease registries and the new national health insurance claim databases.
At baseline, females completed an interviewer-administered electronic questionnaire on lactation duration for each live birth using the specific question "Age and length of lactation at each live birth".
During the follow-up (2004-2015), IHD was assessed through hospital admission, established disease registries, and health insurance databases in China.

Data analysis

Statistics and reproducibility
Latent class analysis (LCA) was employed to identify underlying socioeconomic status (SES) classes based on three individual socioeconomic factors: highest education attainment, occupation, and annual household income. Each factor was categorized into three levels for practical interpretation and considering sample size. The selection criteria of various LCA models included theoretical interpretability and fit statistics, such as low absolute Akaike information criterion and Bayesian information criterion values, likelihood ratio statistic, and high scaled relative entropy42-44. The scaled relative entropy measures the certainty of classification, where 0% indicates poor certainty and 100% represents perfect certainty. Finally, two latent SES classes (low, high) were identified according to the item-response probabilities, and the class memberships were determined by the posterior item probabilities (Supplementary Table S1). The study used R package poLCA (v1.6.0) to implement the LCA procedure45.
Baseline characteristics of the included parous postmenopausal females were described in the form of means and standard deviations (SDs) for normally distributed continuous variables, and medians and interquartile ranges (IQRs) for non-normally distributed continuous variables. Differences were examined by using two samples t-test or Wilcoxon rank sum test, accordingly. Numbers (N) and percentages (%) were

presented for categorical variables, with differences calculated using Chi-square test. The incidence density of incident IHD was calculated as the number of events per 100,000 person-years.
Cox proportional hazards regression models were used to estimate the hazard ratios (HRs) and 95% confidence intervals (CIs) of lactation durations (lifetime lactation duration, per-child lactation duration, and first-child lactation duration) with incident IHD during follow-up. The proportional hazards assumption was examined by Schoenfeld residuals plot (Supplementary Figure S2). To addressing the time-varying effects, Cox models used age as the underlying time scale and adopted a strata function of birth cohort (in 1-year intervals). The time to event was determined as the period from the baseline survey (2004 to 2008) to the date of IHD event, death, loss to follow-up, or the end of follow-up (December 31, 2015), whichever came first. Model 1 was adjusted for age at baseline. Model 2 was additionally adjusted for socio-demographics (marital status, residence, highest education attainment, occupation, annual household income), lifestyle factors (smoking status, passive smoking status, alcohol intake, physical activity), anthropometric measurements (BMI, WC), medical history (history of stroke, history of diabetes, history of hypertension, history of anticoagulation therapy, history of hypolipidemic therapy), and reproductive factors (history of taking oral contraceptive pills, parity, age at first live birth, age at menopause). A competing risk model analysis was conducted to account for the influence of death as a competing event on IHD. Restricted cubic splines (RCS) were used to investigate the dose-risk associations of various forms of lactation durations with incident IHD.
To examine potential variations in different subgroups, we conducted stratified analyses based on various factors including SES*residence, SES, residence, SES components, and parity. Furthermore, we performed stratification analysis by menopausal status to assess the variations in premenopausal and perimenopausal females. P for interaction was evaluated by comparing models with and without a multiplicative interaction term, which included one each for lactation and stratification variables. In addition, we conducted two sensitivity analyses to examine the potential confounding effects of these factors on the model performance. Firstly, we excluded participants who were taking cardiovascular drugs such as angiotensin converting enzyme inhibitors, beta-blockers, calcium antagonists, diuretics, aspirin, and statins. Secondly, we excluded participants with cardiovascular diseases (stroke, diabetes, hypertension, rheumatic heart disease, kidney disease).
To provide a population perspective of how scaling up breastfeeding duration practice may impact the number of IHD cases in China, we calculated the number of preventable IHD cases among Chinese females over 40 in 2019 based on three hypothetical scenarios for average lactation duration per child: (I) all females lactating for an optimal duration (7-12 months as indicated in CKB); (II) 50% of non-lactation females switching to lactating for an optimal duration (7-12 months); (III) 50% of females who lactate for less than 7 months or not at all extending their lactation duration to an optimal duration (7-12 months). . In contrast to PAF, which is typically based on an impractical counterfactual scenario assuming a 100% reduction in the risk factor, PIF is estimated based on reasonably achievable theoretical counterfactual scenarios46,47.This study used PAF for scenario I and PIF for scenarios II and III to estimate the preventable IHD cases.
The calculation of PAF was based on the following equation48:
(PAF=(∑_(i=1)^n⊠⊠Pe?_i (?HR?_i-1) ?)/(∑_(i=1)^n⊠⊠Pe?_i (?HR?_i-1) ?+1) #(1))
Where Pei=the fraction of the IHD cases in the lactation duration category i, HRi=the relative hazard ratio for IHD at lactation duration category i, n=the number of lactation duration categories.
The PIF was calculated by the following formula:
(PIF=((Pe-P^' e)*(HR-1))/(Pe(HR-1)+1) #(2))
Where Pe=the fraction of the IHD cases in lactation duration category, P' e=the counterfactual fraction of the IHD cases that lactation duration category adherent to the optimal duration, HR=the relative hazard ratio for IHD at lactation duration category.
The fraction of IHD cases in each category of average lactation duration per child and relative HRs of IHD were derived from CKB. The 95% CI of PAF and PIF were calculated using bootstrap49. We further calculated how many fewer IHD cases could be expected based on PAF or PIF. The preventable IHD cases were estimated based on the number of the incident IHD among Chinese females above 40 years old in 2019 in Global Burden of Disease 20191.
All analyses were performed using R statistical software (version 4.2.2). The statistical significance was set at two-tailed P<0.05. This study followed with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline for cohort studies.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Data availability: This research has been conducted using the China Kadoorie Biobank (CKB) resource (www.ckbiobank.org). The raw China Kadoorie Biobank data underlying this article can be accessed via https://www.ckbiobank.org/CKBDataAccess, following the institution's data-access policies. Preliminary event adjudication data are not publicly available. Publication of results does not require or imply approval by the membership of the CKB Collaborative Group.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	Our study mainly focuses on the associations between lactation duration with incident ischemic heart disease. These characteristics are biological attributes associated with the physical and physiological features of the human that are present in the participant's biological status as a female. Therefore, the sex term "female" should be used in this paper.
Population characteristics	A total of 130,147 parous postmenopausal females were finally included in our study (median age: 58.4 [IQR 54.0-64.8]), of whom 14,268 (11.0%) developed IHD during follow-up, showing an incidence density of 1276.2 per 100,000 person-years. The baseline characteristics of the 130,147 parous postmenopausal females are shown in Supplementary Data 1. Generally, most participants were urban residents (55.4%), had an education level of primary school or below (74.5%), unemployed or retired (53.5%), belonging to middle-income households (51.0%). These participants with incident IHD tended to have

	gravidity of ≥5 (41.2% vs 30.3%), have parity of ≥4 (37.4 vs% vs 26.1%), compared to non-IHD ones. While in non-IHD subgroup, the proportions of participants with age at first live birth between 20 and 24 (53.4%) or with age at menopause between 45 and 54 (84.4%) are larger. Compared with those who did not develop IHD, parous postmenopausal females who developed IHD have significantly longer lifetime lactation duration (44.0 [IQR 24.0-72.0] months vs. 36.0 [IQR 24.0-60.0] months) and per-child lactation duration (13.0 [IQR 12.0-19.5] months vs. 12.2 [IQR 12.0-19.3] months). While for the first-child lactation duration, this difference was not statistically significant. More baseline characteristics of females by SES or residence are shown in Supplementary Data 2-3.
Recruitment	The CKB study is a large-scale prospective cohort study in the Chinese population28-30. The baseline survey took place between June 25, 2004 and July 15, 2008 in ten geographically defined regions (five urban, five rural) across China. 512,726 residents aged 30-79 years (born in 1930-1970) were recruited.
Ethics oversight	Ethics approvals for the CKB study were obtained from the Ethical Review Committee of the Chinese Center for Disease Control and Prevention (Beijing, China) (Approval No. 005/ 2004), and the Oxford Tropical Research Ethics Committee, University of Oxford (UK) (Approval No. 025-04). All participants provided written informed consent for participation.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	130,147 parous postmenopausal females were our study sample for the analysis.
Data exclusions	Among the 512,726 participants, we excluded females with a history of breast lump removal, hysterectomy, ovary removal, cancer, coronary heart disease, or ischemic heart disease at baseline, 272,052 parous females were included in our study (Supplementary Figure S1). We further excluded females with exceptionally long each-child lactation duration (>60 months) or missing data on socio-demographics and lifestyle factors, resulting 271,435 included. Among them, 134,094 were postmenopausal. After excluding those with abnormal menopausal age (<40 or ≥age)31, 130,147 parous postmenopausal females were our study sample for the analysis.
Replication	Findings of this study are generated by the statistical analyses described in the methodology. All data required for the analysis are public available in the China Kadoorie Biobank website. The programming code is available upon reasonable request, The analyses were run repeatedly and sensitivity analyses were conducted to verify the robustness of the main results.
Randomization	This does not apply to this study as this is a prospective cohort study rather than a randomised clinical trial.
Blinding	This does not apply to this study as this is a prospective cohort study rather than a randomised clinical trial.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging