See corresponding editorial on page 625.

Medical conditions associated with coffee consumption: Disease-trajectory and comorbidity network analyses of a prospective cohort study in UK Biobank

Can Hou,[1,2] *Yu Zeng,*[1,2] *Wenwen Chen,*[3] *Xin Han,*[1,2] *Huazhen Yang,*[1,2] *Zhiye Ying,*[1,2] *Yao Hu,*[1,2] *Yajing Sun,*[1,2] *Yuanyuan Qu,*[1,2] *Fang Fang,*[4] *and Huan Song*[1,2,5]

[1]West China Biomedical Big Data Center, West China Hospital, Sichuan University, Chengdu, China; [2]Med-X Center for Informatics, Sichuan University, Chengdu, China; [3]Division of Nephrology, Kidney Research Institute, State Key Laboratory of Biotherapy and Cancer Center, West China Hospital, Sichuan University, Chengdu, China; [4]Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden; and [5]Center of Public Health Sciences, Faculty of Medicine, University of Iceland, Reykjavík, Iceland

ABSTRACT

Background: Habitual coffee consumption has been associated with multiple health benefits. A comprehensive analysis of disease trajectory and comorbidity networks in relation to coffee consumption is, however, currently lacking.

Objectives: We aimed to comprehensively examine the health outcomes associated with habitual coffee consumption, through clarifying its disease trajectory and comorbidity networks.

Methods: Based on the UK Biobank cohort, we included 395,539 individuals with available information on coffee intake collected at recruitment between 2006 and 2010. These individuals were categorized as having low ($<$1 cup per day), moderate (1–3 cups), and high ($\geq$4 cups) levels of coffee intake, and were followed through 2020 to ascertain 496 medical conditions. Cox regression was used to assess the associations between high-level coffee intake and the risk of medical conditions with a prevalence $\geq$0.5% in the study population, after adjusting for multiple confounders, using low-level coffee intake as the reference. Disease-trajectory and comorbidity network analyses were then applied to visualize the temporal and nontemporal relationships between the medical conditions that had an inverse association with high-level coffee intake.

Results: During a median follow-up of 11.8 years, 31 medical conditions were found to be associated with high-level coffee intake, among which 30 showed an inverse association (HRs ranged from 0.61 to 0.94). The inverse associations were more pronounced for women, compared with men. Disease-trajectory and comorbidity network analyses of these 30 conditions identified 4 major clusters of medical conditions, mainly in the cardiometabolic and gastrointestinal systems, among both men and women; 1 cluster of medical conditions following alcohol-related disorders, primarily among men; as well as a cluster of estrogen-related conditions among women.

Conclusions: Habitual coffee consumption was associated with lower risks of many medical conditions, especially those in the cardiometabolic and gastrointestinal systems and those related to alcohol use and estrogen regulation. *Am J Clin Nutr* 2022;116:730–740.

Keywords: coffee, caffeine, cardiometabolic diseases, gastrointestinal diseases, estrogen-related conditions, alcohol-related disorders

Introduction

Coffee is among the world's most popular beverages, and it is estimated that over 2.25 billion cups of coffee are consumed

This work was supported by 1.3.5 Project for Disciplines of Excellence, West China Hospital, Sichuan University (grant no. ZYYC21005 to HS) and the National Natural Science Foundation of China (grant no. 81971262 to HS).

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Supplemental Tables 1–5, Supplemental Methods, and Supplemental Figures 1–9 are available from the "Supplementary data" link in the online posting of the article at https://academic.oup.com/ajcn/.

Address correspondence to HS (e-mail: songhuan@wchscu.cn).

Abbreviations used: D1, disease 1; D2, disease 2; HRT, hormone-replacement therapy; IBD, inflammatory bowel disease; IBS, irritable bowel syndrome; ICD-10, International Classification of Diseases, Tenth Edition; NHS, National Health Service; q-value, false discovery rate–adjusted *P* value; PheWAS, phenome-wide association analysis.

Received February 16, 2022. Accepted for publication May 14, 2022.

First published online July 15, 2022; doi: https://doi.org/10.1093/ajcn/nqac148.

Am J Clin Nutr 2022;116:730–740. Printed in USA.

daily worldwide (1). Given the high consumption level of coffee, the health consequences of coffee intake have attracted much research attention. Previous studies generally support the notion that long-term coffee consumption is associated with more health benefits than harm; meta-analyses of observational studies have consistently shown an association between coffee consumption and reduced all-cause mortality (2). Other studies have also shown that coffee consumption is associated with lower risks of diabetes mellitus (3), cardiovascular disease (4), dementia (5), Parkinson's diseases (6), and specific types of cancer (7).

Although the existing literature indicates that the health benefits of coffee consumption could extend to a wide range of medical conditions, a comprehensive analysis of disease trajectories and comorbidity networks in relation to coffee consumption is still lacking. Given the complexity of human disease networks, a comprehensive analysis of health outcomes following habitual coffee consumption could be important. The recently proposed disease-trajectory analysis is an ideal approach to achieve this purpose (8). Despite being designed as a method to visualize the networks of disease progression over time (9), a disease-trajectory analysis can also be used to investigate the temporal order of a set of diseases in relation to a predetermined phenotype. However, as the majority of diseases do not demonstrate clear temporal orders, they are discarded automatically in a disease-trajectory analysis. One way to address this limitation is to use a comorbidity network analysis, which investigates the diversity of disease clusters associated with a predetermined phenotype, with an emphasis on the strength of correlations rather than the temporal orders (10), as a complement.

To this end, leveraging the rich information on demographics, lifestyle factors, and health records of the UK Biobank, we comprehensively examined the health benefits associated with habitual coffee consumption, through clarifying the disease trajectories and comorbidity networks among individuals with different levels of coffee intake. The findings of the present study could help to improve our understanding of the associations between coffee consumption and different medical conditions and provide novel insights on the potential underlying mechanisms for such associations.

Methods

Study design

The UK Biobank was a population-based, prospective study conducted across the United Kingdom between 2006 and 2010 (11). In brief, the UK Biobank recruited a total of approximately 500,000 participants aged 40–69 years and, at recruitment, collected information on sociodemographic factors, lifestyle factors, and body measurements. Participants were followed individually from recruitment to ascertain a broad range of medical conditions, as well as mortality, through periodical linkages to several national databases. For participants registered in England, Wales, and Scotland, inpatient hospital data were derived from the Hospital Episode Statistics database, the Patient Episode Database for Wales, and the Scottish Morbidity Record, respectively, which were deemed to cover all UK Biobank participants for the period of January 1997 to November 2020. Primary care data were obtained from multiple general practice data system suppliers, covering approximately 45% of the UK Biobank participants (12). The diagnoses of medical conditions in inpatient hospital data and primary care data have been validated, showing an overall high level of accuracy (13, 14). Nationwide mortality data were updated from the National Health Service (NHS) Digital (England and Wales) and NHS Central Register (Scotland). The UK Biobank was approved by the NHS National Research Ethics Service (reference number: 16/NW/0274). The present study was approved by the biomedical research ethics committee of West China Hospital (reference number: 2019-1171).

In the present study, among the 502,507 UK Biobank participants, we excluded 48 individuals who withdrew their informed consent forms, as well as 52,337 individuals with a history of a severe disease and 52,915 individuals with a preexisting medical condition at recruitment that might have prohibited a high level of coffee intake, including gastroduodenal ulcers, gastroesophageal reflux disease, inflammatory bowel disease (IBD), irritable bowel syndrome (IBS), and glaucoma (15–19), according to both the inpatient hospital care and primary care data. These diseases and their diagnostic codes are listed in **Supplemental Table 1**. Among the remaining 397,207 individuals, 395,539 provided information on coffee intake (cups per day and type of coffee they usually consumed: i.e., regular or decaffeinated coffee) through a dietary touchscreen questionnaire at recruitment, and were included in the final analysis cohort. These individuals were categorized into groups with low (<1 cup per day; 113,995 individuals), moderate (1–3 cups per day; 204,140 individuals), and high (≥4 cups per day; 77,404 individuals) levels of coffee intake. The coffee intake level obtained from the baseline dietary questionnaire agreed well with the measurement based on a 24-hour dietary assessment conducted between 2009 and 2010 (kappa = 0.71) among 70,692 participants with available information from both assessments. Follow-up of the cohort started from recruitment and went until death, loss to follow-up [i.e., no further linkage to health-care records due to different reasons (20)], or the end of the study (30 November 2020), whichever occurred first.

Diagnoses of medical conditions.

Diagnoses of medical conditions during the follow-up were identified based on the main and secondary diagnoses in the UK Biobank inpatient hospital data and underlying cause of death in the mortality data, according to the International Classification of Diseases, Tenth Edition (ICD-10), codes. We excluded the ICD-10 codes related to pregnancy and perinatal conditions (i.e., ICD-10 Chapters 15–17), symptoms or signs (Chapter 18), and factors influencing health status (Chapter 21). We then mapped the remaining ICD-10 codes to "phecodes," a coding system considered more relevant to diseases discussed in clinical settings and widely used in the phenome-wide association analysis (PheWAS) (21). To allow sufficient statistical power in the association analysis, we restricted our analysis to the 496 top-level phecodes (**Supplemental Table 2**) corresponding to 3-digit ICD-10 codes. Due to the hypothesis-generating nature of the study, we made no further selection among the 496 medical conditions. For each medical condition, the date of diagnosis was defined as the date of the first hospital visit with a diagnosis of the corresponding phecode.

Covariates.

Information on sociodemographic and lifestyle factors, including date of birth, sex, household income, tea intake, smoking, alcohol drinking, and physical activity, was collected from all participants at recruitment using touchscreen questionnaires. For women, information on reproductive factors, including menopause status and use of hormone-replacement therapy (HRT), was additionally collected at recruitment. According to the postal codes, each participant was assigned a Townsend deprivation index, with a higher score indicating greater deprivation (22). We created a drinking variable with 6 categories (i.e., low-risk drinking, hazardous drinking, harmful drinking, former drinker, never drinker, and unknown) based on the drinking status and the frequency and intensity of current alcohol consumption. Similarly, a 6-level smoking variable was created based on smoking status, type of tobacco currently smoked, and current smoking frequency and intensity. We calculated total physical activity by summing all time spent in vigorous, moderate, and walking activities per day, weighted by the metabolic equivalent task score of each activity (23). The total fruit and vegetable consumption per day was measured by summing the reported fresh or dried fruit intake and cooked or raw vegetable intake.

BMI was calculated based on the height and weight measured during the medical center visit at recruitment. Blood pressure and heart rate were measured by trained nurses using an electronic sphygmomanometer at rest. In addition, as we found associations between coffee consumption and lower risks of several estrogen-related conditions, we further investigated the influence of coffee intake on estrogen levels. We retrieved data on the circulating estradiol level for all participants, which was measured using a 2-step sandwich immunoassay based on blood samples collected at recruitment.

The existence of gastrointestinal or cardiovascular diseases and their related symptoms might modify the habit of coffee drinking (24, 25); we therefore ascertained the history of these conditions (Supplemental Table 1) at baseline for all study participants, according to the inpatient hospital care and primary care data.

Statistical analysis

PheWAS.

Cox models were used to assess the HR of each medical condition in relation to high-level coffee consumption, compared to low-level coffee consumption, adjusting for age, sex, Townsend deprivation index, household income, BMI, tea intake, smoking status, alcohol drinking status, physical activity, and fruit and vegetable consumption. To ensure statistical power, only medical conditions with a prevalence $\geq$0.5% were included in the corresponding PheWAS. In the analysis of each medical condition, we excluded from the analysis individuals with a history of all medical conditions related to the analyzed condition (Supplemental Table 2), to reduce the possibility that the studied medical condition was more related to these preexisting medical conditions rather than the exposure to high-level coffee intake. To adjust for multiple testing in the PheWAS, we calculated the false discovery rate–adjusted *P* value (hereafter referred to as the q-value) for each test (26), where a q-value < 0.05 was considered to be statistically significant. The proportional hazard assumption was checked by conducting statistical tests and visualizing the scaled Schoenfeld residuals plots (27), and no violation was detected for any of the medical conditions with a q-value < 0.05 in the main PheWAS.

Disease-trajectory and comorbidity network analyses.

As we were interested in the potential health benefits of coffee intake, we included only medical conditions with an HR < 1.0 and q-value < 0.05 according to the PheWAS in the disease-trajectory and comorbidity network analyses. A detailed description of the analysis steps is available in the **Supplemental Methods**. As both disease-trajectory and comorbidity network analyses need to be performed in a population with an enhanced disease risk, we restricted the following analyses among individuals with low-level coffee intake. In step 1, we identified all possible disease 1 (D1) and disease 2 (D2) pairs if the cooccurrence of the 2 diseases was experienced by at least 0.1% of these individuals. RRs of observing a disease pair in the same individual and Pearson's correlations of these 2 diseases (i.e., Φ-correlation) were calculated, according to formulas provided in the Supplemental Methods, as the measures of comorbidity strength. Disease pairs with considerable comorbidity strength—that is, an RR > 1.0 and Φ-correlation > 0—were eligible in the analysis of step 2, where disease-trajectory and comorbidity network analyses were conducted in parallel. Specifically, in the disease-trajectory analysis, we applied a binomial test to identify the disease pairs with a clear temporal order (i.e., significantly more individuals had D2 diagnosed after D1 than vice versa). We then used conditional logistic regression to confirm the directionality after adjusting for confounders. The D1→D2 pairs identified in this analysis (i.e., OR >1.0 and q-value < 0.05) were used to construct a disease-trajectory network. In the comorbidity network analysis, without the consideration of directionality, we selected disease pairs using unconditional logistic regression, after adjusting for confounders. Disease pairs (i.e., D1↔D2) with confirmed comorbidity (i.e., OR > 1.0 and q-value < 0.05) in this analysis were then used for comorbidity network construction to investigate comorbidity patterns (i.e., modules with high intrinsic connectivity) by community detection algorithm Louvain (28).

Subgroup and sensitivity analyses.

To investigate whether the effects of high-level coffee consumption could be modified by sex, for medical conditions with statistically significant results in the main PheWAS, we tested the differences between sex-specific HRs by introducing an interaction term in the models. The subsequent disease-trajectory and comorbidity network analyses were performed separately for men and women. Further, instead of focusing on the high-coffee-consumption group alone, we tested the association between the level of coffee consumption (cups per day) and risks of these diseases. Both linear and categorical, nonlinear models were fitted, and a likelihood ratio test was used to determine whether the dose-response relationship was linear. In addition, we performed the PheWAS separately for a high-level intake of regular coffee (N = 61,784) and a high-level intake of decaffeinated coffee (N = 14,722).

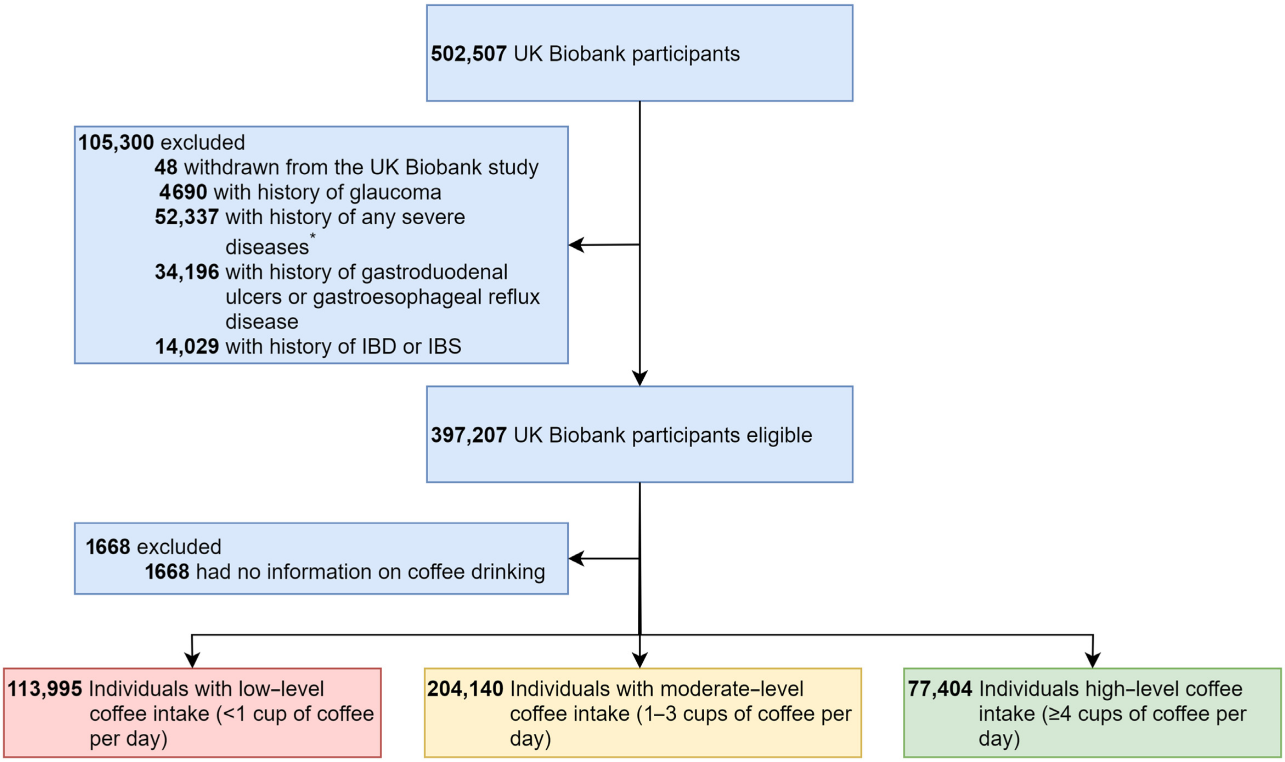
502,507 UK Biobank participants
105,300 excluded
48 withdrawn from the UK Biobank study
4690 with history of glaucoma
52,337 with history of any severe diseases*
34,196 with history of gastroduodenal ulcers or gastroesophageal reflux disease
14,029 with history of IBD or IBS
397,207 UK Biobank participants eligible
1668 excluded
1668 had no information on coffee drinking
113,995 Individuals with low–level coffee intake (<1 cup of coffee per day)
204,140 Individuals with moderate–level coffee intake (1–3 cups of coffee per day)
77,404 Individuals high–level coffee intake (≥4 cups of coffee per day)

FIGURE 1 Flowchart of the study population selection. *Severe diseases include myocardial infarction, congestive heart failure, peripheral vascular disease, cerebrovascular disease, dementia, chronic pulmonary disease, connective tissue disease, liver disease, diabetes mellitus (with or without chronic complications), hemiplegia, moderate or severe renal disease, any tumor, leukemia, lymphoma, and acquired immune deficiency syndrome. IBD, inflammatory bowel disease; IBS, irritable bowel syndrome.

To further reduce the possibility that preexisting health conditions might affect one's choice of coffee consumption, in addition to the exclusion of individuals with gastroduodenal ulcers, gastroesophageal reflux disease, IBS, IBD, or glaucoma at baseline, we performed 2 sensitivity analyses by repeating the PheWAS after: *1*) excluding 90,509 individuals with a diagnosis or symptoms of cardiovascular or gastrointestinal diseases at baseline (Supplemental Table 1) and 137,835 individuals with elevated blood pressure ($\geq$140/90 mm Hg) or heart rate ($\geq$120) at baseline; and *2*) excluding noncoffee drinkers (i.e., individuals with 0 cups of coffee per day) from the low-level consumption group ($n = 85{,}308$), assuming an extreme scenario that stringent noncoffee drinkers were intolerant to coffee. In an additional sensitivity analysis, we used a 2-year lag time since the recruitment date to assess the impact of the temporal order between exposure to coffee intake and subsequent incidences of medical conditions on the results.

All the statistical analyses were conducted using SciPy (version 1.4.1)(29), Statsmodels (version 0.11.1)(30) and Lifelines (version 0.25.2)(31) in Python 3.8.

Results

The final study cohort consisted of 395,539 individuals, among whom 113,995 and 77,404 had low and high levels of coffee consumption, respectively (**Figure 1**). The median ages at cohort entry were 56.0 and 57.0 years for individuals with low and high levels of coffee consumption, respectively, and the median follow-up time was 11.8 years for both (**Table 1**). Compared with individuals with low-level coffee consumption, those with high-level coffee consumption were more likely to be male (51.57% compared with 42.02%, respectively), to be obese (BMI $\geq$ 29.9 kg/m^2, 27.32% compared with 23.38%, respectively), to be a current smoker (17.03% compared with 8.85%, respectively), to be a current heavy alcohol drinker (hazardous or harmful drinking, 42.29% compared with 31.17%, respectively), and to have a better household income ($\geq$52,000£, 25.27% compared with 21.36%, respectively), but were less likely to have high-level tea intake ($\geq$4 cups per day, 13.31% compared with 46.27%, respectively) or enough fruit and vegetable consumption (28.38% compared with 31.45%, respectively). A lower level of total estradiol was observed among individuals with high-level coffee intake, compared to individuals with low-level coffee intake; this was, however, only noted among women, regardless of menopause status or HRT use (**Supplemental Table 3**). Similar results were also observed when comparing individuals with moderate-level coffee consumption to those with low-level coffee consumption, with the exception of BMI and smoking status.

PheWAS

Among the 174 medical conditions with a prevalence $\geq$ 0.5%, 31 showed statistically significant associations with high-level coffee consumption, using low-level coffee consumption as a reference (**Figure 2; Supplemental Table 4**). The majority

TABLE 1 Characteristics of the study participants[1]

Characteristics	Low-level coffee consumption (*n* = 113,995)	Moderate-level coffee consumption (*n* = 204,140)	High-level coffee consumption (*n* = 77,404)
Age at recruitment, y	56.0 (48.7–62.3)	58.5 (50.7–63.7)	57.0 (49.7–62.8)
Follow-up time, y	11.8 (11.1–12.5)	11.8 (11.1–12.5)	11.8 (11.1–12.5)
Sex			
Female	66,094 (57.98%)	112,445 (55.08%)	37,486 (48.43%)
Male	47,901 (42.02%)	91,695 (44.92%)	39,918 (51.57%)
Townsend deprivation index			
<−3.64	26,217 (23.00%)	55,060 (26.97%)	20,012 (25.85%)
−3.64 to −2.14	27,304 (23.95%)	52,885 (25.91%)	19,980 (25.81%)
−2.14 to 0.55	28,990 (25.43%)	51,060 (25.01%)	19,332 (24.98%)
≥0.55	31,341 (27.49%)	44,874 (21.98%)	17,993 (23.25%)
Unknown	143 (0.13%)	261 (0.13%)	87 (0.11%)
Household income, £			
<18,000	21,872 (19.19%)	33,752 (16.53%)	12,838 (16.59%)
18,000 to 52,000	49,673 (43.57%)	90,967 (44.56%)	34,747 (44.89%)
≥52,000	24,353 (21.36%)	50,631 (24.80%)	19,557 (25.27%)
Unknown	18,097 (15.88%)	28,790 (14.10%)	10,262 (13.26%)
BMI, kg/m2			
<24.1	30,842 (27.06%)	55,227 (27.05%)	16,081 (20.78%)
24.1 to 29.9	55,818 (48.97%)	104,665 (51.27%)	39,884 (51.53%)
≥29.9	26,648 (23.38%)	43,453 (21.29%)	21,146 (27.32%)
Unknown	687 (0.60%)	795 (0.39%)	293 (0.38%)
Smoking status			
<5 cigarettes/d	3434 (3.01%)	7585 (3.72%)	3590 (4.64%)
5 to 14 cigarettes/d	2657 (2.33%)	4358 (2.13%)	3190 (4.12%)
≥15 cigarettes/d	4006 (3.51%)	5055 (2.48%)	6399 (8.27%)
Former smoker	34,361 (30.14%)	69,657 (34.12%)	27,133 (35.05%)
Never smoker	69,052 (60.57%)	116,760 (57.20%)	36,738 (47.46%)
Unknown	485 (0.43%)	725 (0.36%)	354 (0.46%)
Alcohol drinking status[2]			
Low-risk drinking	63,249 (55.48%)	110,126 (53.95%)	38,816 (50.15%)
Hazardous drinking	29,113 (25.54%)	68,932 (33.77%)	27,521 (35.56%)
Harmful drinking	6421 (5.63%)	11,344 (5.56%)	5212 (6.73%)
Former drinker	4908 (4.31%)	4409 (2.16%)	2570 (3.32%)
Never drinker	8296 (7.28%)	6036 (2.96%)	2005 (2.59%)
Unknown	2008 (1.76%)	3293 (1.61%)	1280 (1.65%)
Coffee type[3]			
Regular	22,400 (19.65%)	164,586 (80.62%)	61,784 (79.82%)
Decaffeinated	5412 (4.75%)	37,531 (18.38%)	14,722 (19.02%)
Unknown	875 (0.77%)	2023 (0.99%)	898 (1.16%)
Not available	85,308 (74.83%)	0 (0.00%)	0 (0.00%)
Total physical activity[4]			
Low	22,655 (19.87%)	37,982 (18.61%)	16,308 (21.07%)
Moderate	44,380 (38.93%)	87,827 (43.02%)	30,535 (39.45%)
High	23,637 (20.74%)	41,639 (20.40%)	15,729 (20.32%)
Unknown	23,323 (20.46%)	36,692 (17.97%)	14,832 (19.16%)
Fruit and vegetable consumption[5]			
Inadequate	77,963 (68.39%)	135,825 (66.54%)	55,339 (71.49%)
Adequate	35,855 (31.45%)	68,153 (33.39%)	21,969 (28.38%)
Unknown	177 (0.16%)	162 (0.08%)	96 (0.12%)
Tea intake[6]			
Low	11,713 (10.28%)	26,613 (13.04%)	31,774 (41.05%)
Moderate	49,267 (43.22%)	123,843 (60.67%)	35,226 (45.51%)
High	52,751 (46.27%)	53,515 (26.21%)	10,301 (13.31%)
Unknown	264 (0.23%)	169 (0.08%)	103 (0.13%)

[1] The values are reported as the median (lower quantile–upper quantile) for continuous variables and *n* (%) for categorical variables. Low-, moderate-, and high-level coffee consumption were defined as drinking <1, 1–3, or ≥4 cups of coffee per day. MET, metabolic equivalent task.

[2] For a current drinker, the alcohol consumption level was calculated by converting the reported number of glasses to the UK standard unit for each type of alcohol and summing up different types of alcohol intakes. Low-risk drinking was defined as an alcohol consumption level ≤14 units/week, hazardous drinking was defined as 14–35 units/week (for women) or 14–50 units/week (for men), and harmful drinking was defined as ≥35 units/week (for women) or ≥50 units/week (for men).

[3] Coffee drinkers were asked to select the type of coffee they usually consumed from the following mutually exclusive responses: decaffeinated coffee (any type), instant coffee, ground coffee, other type of coffee, or "do not know or prefer not to answer." The categories of instant coffee, ground coffee, and other type of coffee are combined as regular coffee.

[4] The total physical activity amount was calculated by summing the MET weighted time spent in vigorous, moderate, and walking activities. Low, moderate, and high physical activities were defined as total physical activity amounts <798, 798–3,552, or ≥3,552 MET min/wk.

[5] Inadequate fruit or vegetable consumption was defined as eating less than 5 portions of fruit and vegetables per day.

[6] Low-, moderate-, and high-level tea intake were defined as tea consumption <1, 1–4, or ≥4 cups per day.

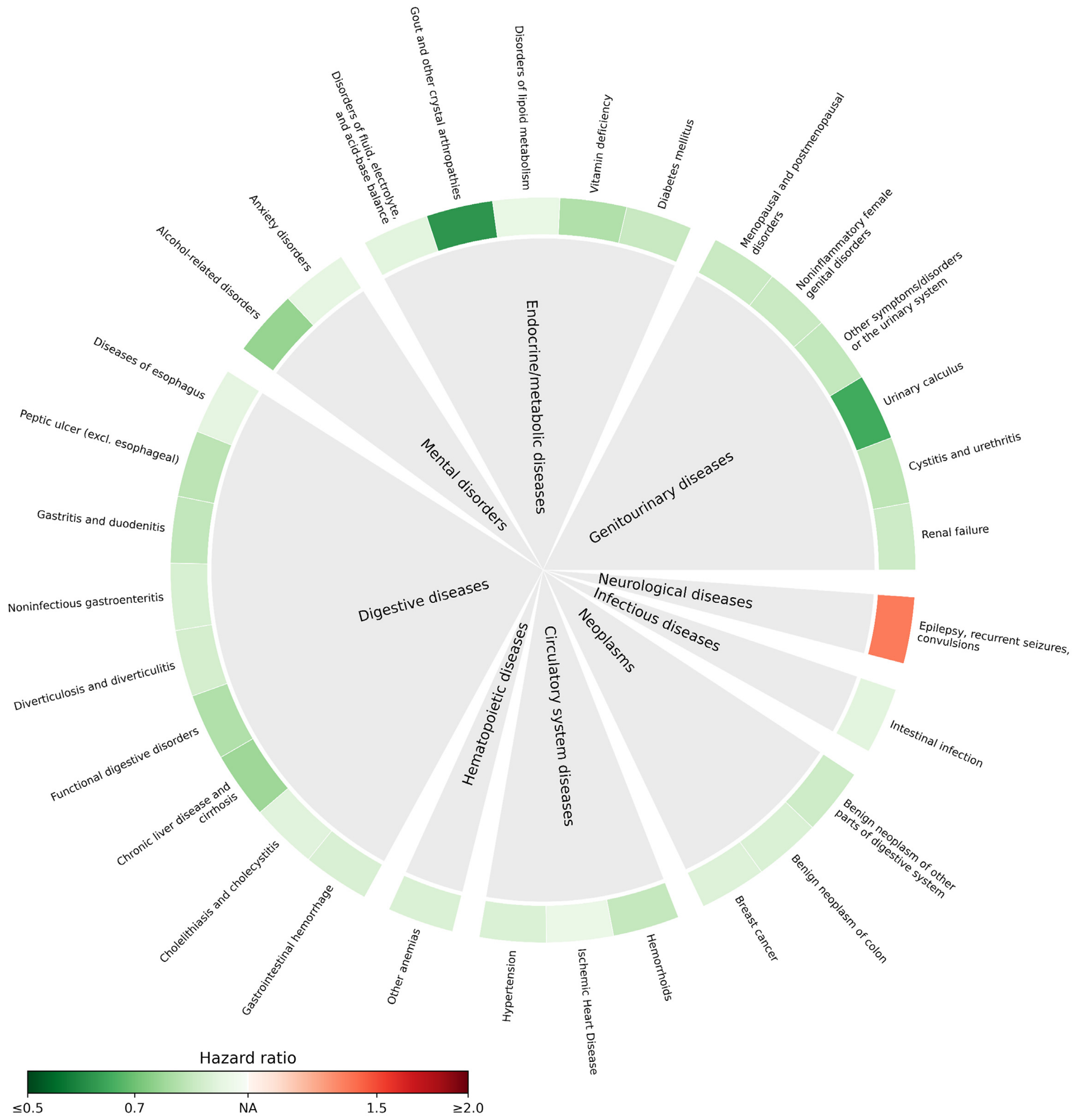

FIGURE 2 RRs of subsequent medical conditions for high-level coffee consumption compared to low-level coffee consumption ($n = 191{,}399$). The outer ring shows the point estimates of HRs of identified medical conditions that were statistically significantly associated with high-level coffee consumption after correction for multiple testing (i.e., false discovery rate–adjusted P value < 0.05). HRs were derived from a Cox model, adjusted for age, sex, Townsend deprivation index, household income, BMI, tea intake, smoking status, alcohol drinking status, physical activity, and fruit and vegetable consumption. The red color indicates a higher risk (i.e., HR > 1) and the green color shows a lower risk (i.e., HR < 1). The degree of color represents the magnitude of the corresponding association. Detailed results are shown in Supplemental Table 4.

of these conditions (96.77%; 30 of 31) showed an inverse association with high-level coffee consumption, with the smallest HRs observed for gout and other crystal arthropathies (HR, 0.61; 95% CI, 0.56–0.67), urinary calculus (HR, 0.65; 95% CI, 0.59–0.72), and alcohol-related disorders (HR, 0.75; 95% CI: 0.69–0.82). A positive association was noted between high-level coffee intake and epilepsy (HR, 1.37; 95% CI, 1.21–1.54).

Diseases networks in relation to high-level coffee consumption: disease-trajectory and comorbidity network analyses

Among a total of 435 possible disease pairs constructed among the 30 medical conditions with an inverse association with high-level coffee intake, 305 were retained after the selection based on prevalence and comorbidity strength measures (**Supplemental**

Figure 1). We identified 68 D1→D2 pairs with clear temporal orders in the disease-trajectory analysis and 297 D1↔D2 pairs with confirmed comorbidity associations in the comorbidity network analysis.

Figure 3 shows an overview of the disease trajectories of medical conditions with an inverse association with high-level coffee consumption. Briefly, according to medical conditions placed in the first layer of the network, 4 major clusters were observed based on the similarities in the affected systems or etiologies. Cluster 1 consisted of medical conditions affecting the cardiometabolic system, where the disease tree thrived after the diagnoses of diabetes mellitus, hypertension, gout, and ischemic heart disease. Cluster 2 started with a group of gastrointestinal system–related diseases, including gastrointestinal hemorrhage, hemorrhoids, noninfectious gastroenteritis, and benign neoplasm of colon. Cluster 3 originated from alcohol-related disorders, which then linked to several downstream medical conditions, such as chronic liver disease and cirrhosis and renal failure. The last cluster consisted of a group of estrogen-related conditions that were specific for women, including noninflammatory female genital disorders, breast cancer, and menopausal and postmenopausal disorders.

Figure 4 shows the comorbidity network of medical conditions with an inverse association with high-level coffee consumption. The main modules identified from the network were largely comparable to the clusters observed in the disease-trajectory analysis, including modules predominated by conditions related to the upper (the module in the right upper corner, with mainly brown nodes) and lower (the module in the right lower corner, with mainly brown nodes) gastrointestinal tract and cardiometabolic diseases (orange and blue nodes), as well as modules centered around estrogen-related conditions (light brown nodes).

Subgroup and sensitivity analyses

The subgroup analyses by sex identified 19 and 29 medical conditions with statistically significant inverse associations with high-level coffee consumption in men and women, respectively (**Supplemental Figure 2; Supplemental Table 5**). In general, the lower risk in relation to high-level coffee intake was more pronounced for women, compared with men, with the largest difference noted for functional digestive disorders (1.00 for men compared with 0.73 for women; *P*-interaction = 0.01), followed by diabetes mellitus (0.92 compared with 0.74, respectively; *P*-interaction < 0.001) and gout (0.59 compared with 0.71, respectively; *P*-interaction = 0.03). Except for estrogen-related conditions, all 3 clusters observed in the main trajectory analysis were observed among men (**Supplemental Figure 3**). Among women, the cluster originating from alcohol-related disorders was not present, while an additional cluster starting from other urinary system symptoms or disorders was observed. In the comorbidity network analysis, we obtained largely similar comorbidity patterns for women as in the main analysis, but a simpler network, with fewer connections, was noted for men (**Supplemental Figure 4**).

We found a linear, dose-response association (monotonically decreasing) between the level of coffee consumption (cups per day) and 11 medical conditions, whereas a U-shaped, nonlinear association was noted for the other 19 conditions (**Supplemental Figure 5**). Further, in the analyses by type of coffee, we obtained largely attenuated estimates for high-level intake of decaffeinated coffee, with 5 medical conditions showing inverse associations (**Supplemental Figure 6**). For high-level intake of regular coffee, however, inverse associations were noted for 43 conditions, including all 30 conditions found to be statistically significant in the main analysis.

Largely similar result patterns were observed in the sensitivity analysis, where individuals with prevalent cardiovascular or gastrointestinal problems were excluded (**Supplemental Figure 7**). Likewise, we obtained identical results after adding 2 years of lag time in the analysis (**Supplemental Figure 8**). We observed attenuated estimates in the PheWAS analysis after excluding non–coffee drinkers from the group with low-level coffee intake—that is, comparing coffee drinkers with >4 cups per day to those with <1 but >0 cups per day—although a majority of the risk patterns remained (**Supplemental Figure 9**).

Discussion

In this large, prospective study based on the UK Biobank, we found that individuals with high-level coffee consumption (i.e., ≥4 cups per day) were at lower risk of 30 medical conditions, compared with individuals with low-level coffee consumption (i.e., <1 cup per day). Further disease-trajectory and comorbidity network analyses revealed 2 distinct clusters of medical conditions, affecting mainly the cardiometabolic and gastrointestinal systems. A subgroup analysis by sex suggested stronger results among women than men. We also found a cluster of estrogen-related medical conditions among women, as well as a cluster of alcohol-related disorders among men, in relation to high-level coffee consumption. Taken together, these results further suggest that habitual coffee consumption is likely beneficial for health for both men and women, possibly due to its associations with cardiometabolic and gastrointestinal diseases and diseases related to alcohol use and estrogen regulation.

Although few studies have explored disease networks subsequent to high-level coffee consumption, our finding of a lower risk for medical conditions, including diabetes mellitus, gout, hypertension, liver cirrhosis, and cholelithiasis, corroborates the results of earlier studies (3, 4, 32–34). Evidence on the role of coffee consumption on nonneoplastic gastrointestinal outcomes was, however, scarce (35). Further, as the disease-trajectory and comorbidity network analyses required a large study sample size with a large number of affected cases, we failed to confirm the previously suggested inverse associations of coffee consumption with some relatively rare cancer types (7). We also failed to confirm the previously reported inverse associations between coffee consumption and neurodegenerative diseases (e.g., dementia and Parkinson's disease) (5, 6). Nevertheless, our null result regarding the latter is in line with 2 recent studies based on data from the UK Biobank, where little evidence was found that habitual coffee consumption could improve cognitive function (36, 37).

Our findings support the notion that the potential health benefits of habitual coffee consumption may be mainly attributable to coffee's impact on the cardiometabolic and gastrointestinal systems. Given that multiple coffee components (e.g., caffeine) have antioxidant, anti-inflammatory, and antiproliferative effects,

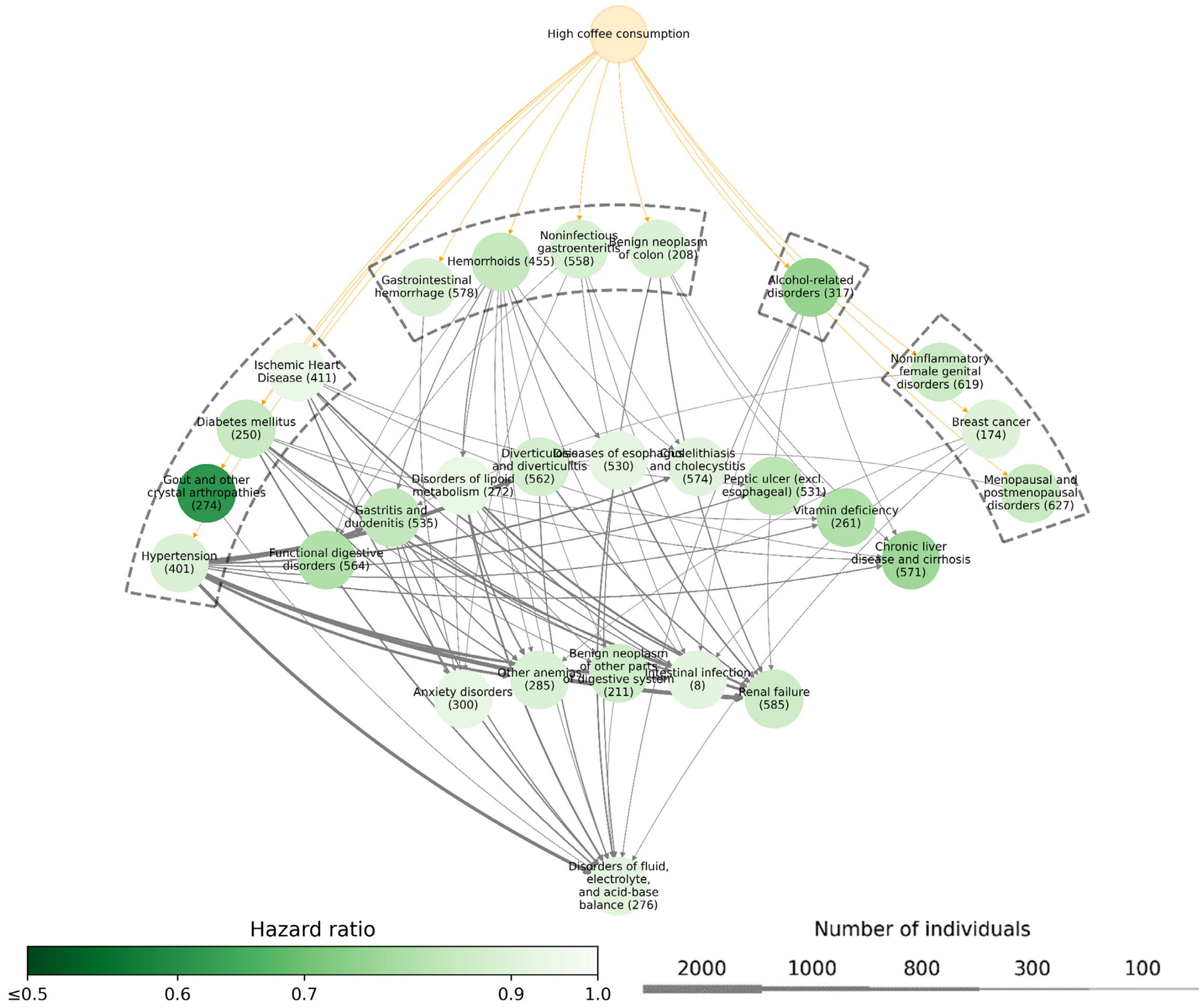
High coffee consumption
Noninfectious gastroenteritis (558)
Benign neoplasm of colon (208)
Hemorrhoids (455)
Gastrointestinal hemorrhage (578)
Alcohol-related disorders (317)
Ischemic Heart Disease (411)
Noninflammatory female genital disorders (619)
Diabetes mellitus (250)
Breast cancer (174)
Diverticulosis and diverticulitis (562)
Diseases of esophagus (530)
Cholelithiasis and cholecystitis (574)
Gout and other crystal arthropathies (274)
Disorders of lipoid metabolism (272)
Peptic ulcer (excl. esophageal) (531)
Menopausal and postmenopausal disorders (627)
Gastritis and duodenitis (535)
Vitamin deficiency (261)
Hypertension (401)
Functional digestive disorders (564)
Chronic liver disease and cirrhosis (571)
Benign neoplasm of other parts of digestive system (211)
Other anemias (285)
Intestinal infection (8)
Anxiety disorders (300)
Renal failure (585)
Disorders of fluid, electrolyte, and acid-base balance (276)
Hazard ratio
≤0.5
0.6
0.7
0.9
1.0
Number of individuals
2000
1000
800
300
100

FIGURE 3 Disease-trajectory network of medical conditions with a lower risk in relation to high-level coffee consumption (n = 191,399). Each node represents a medical condition, with the color of the node representing the HR of the corresponding medical condition when comparing individuals with high-level coffee consumption to those with low-level coffee consumption, according to an adjusted Cox model. The width of the lines connecting 2 nodes represents the number of individuals with the corresponding disease trajectory.

it is biologically plausible that coffee intake might modulate the risks of diseases in these systems. Previous studies have shown, for instance, the antioxidant effects of coffee components—namely, flavonoids, melanoidines, and chlorogenic acids (38)—although coffee is not the main dietary source for flavonoids. Similarly, evidence has suggested that the associations between coffee and metabolic diseases are likely due to the role of caffeine in increasing metabolic rates (39) and the antagonistic effects of phenolic compounds in coffee (40). Our finding of weaker associations for high-level intake of decaffeinated coffee, compared with high-level intake of regular coffee, also suggests the importance of caffeine.

The sex-specific disease-trajectory analysis showed a stronger result pattern among women, compared with men. This might be partially explained by the higher activity of CYP1A2, which is a main enzyme involved in caffeine degradation, in men than women (41). Although it remains unknown whether such a sex difference in catabolism is appliable to other components of coffee, the finding of an inverse association between high-level coffee intake and estrogen-related conditions is supported by the fact that women with high-level coffee intake had lower levels of circulating estradiol, consistent with previous reports (42). The association between high-level coffee intake and lower risks of chronic liver diseases is also supported by the previous literature (33, 43, 44). Although the underlying mechanisms for such association are currently unknown, our network analyses suggest that this association might be partly attributable to the lower risk of alcohol addiction among regular coffee drinkers, and primarily among men. However, as a positive correlation has also been suggested between alcohol and caffeine use (45), further investigations are warranted to better understand the correlations between coffee and alcohol intakes, as well as how coffee and alcohol intakes could jointly affect downstream diseases (e.g., chronic liver diseases).

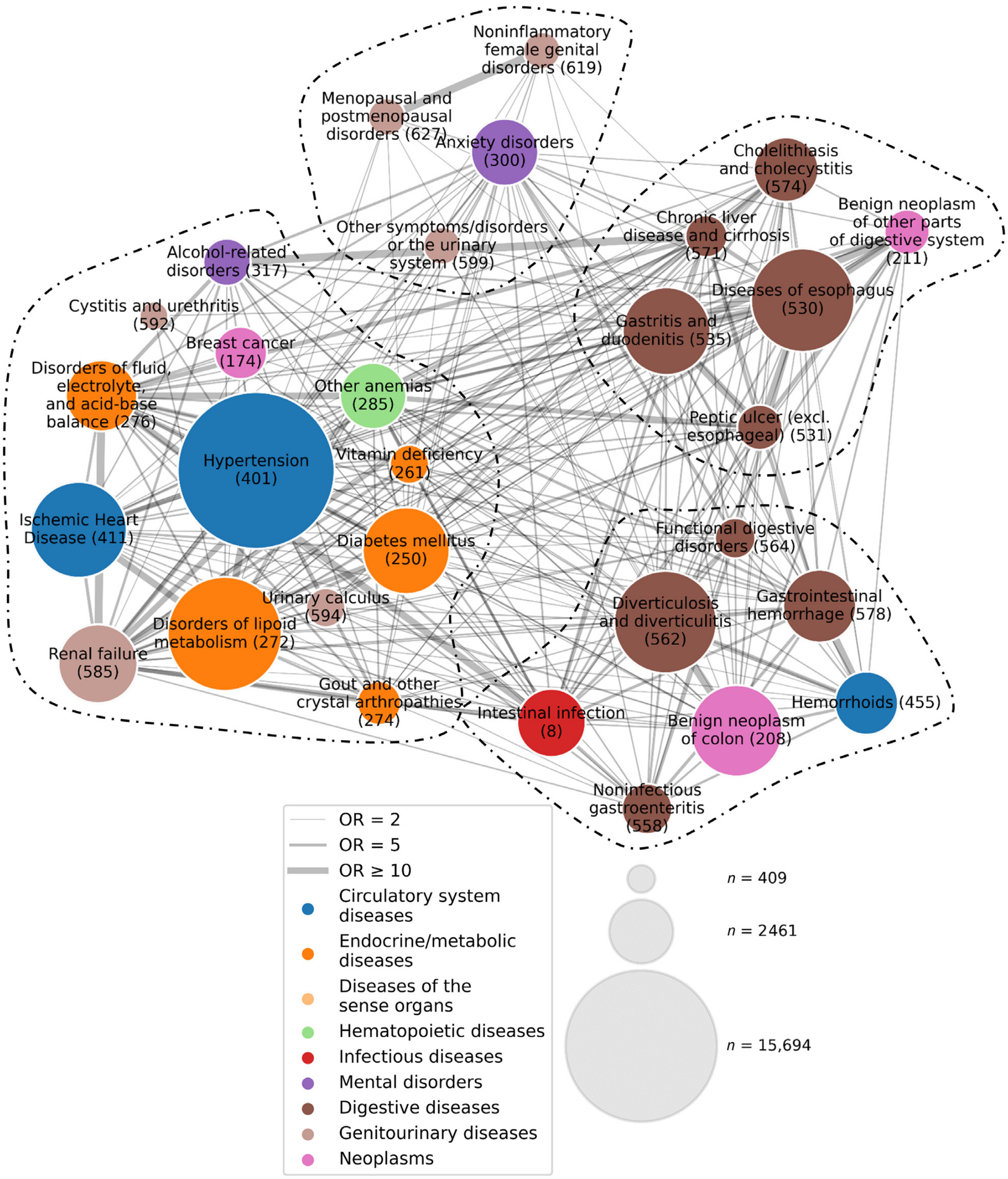
Noninflammatory female genital disorders (619)
Menopausal and postmenopausal disorders (627)
Anxiety disorders (300)
Cholelithiasis and cholecystitis (574)
Benign neoplasm of other parts of digestive system (211)
Chronic liver disease and cirrhosis (571)
Other symptoms/disorders or the urinary system (599)
Alcohol-related disorders (317)
Diseases of esophagus (530)
Cystitis and urethritis (592)
Gastritis and duodenitis (535)
Breast cancer (174)
Disorders of fluid, electrolyte, and acid-base balance (276)
Other anemias (285)
Peptic ulcer (excl. esophageal) (531)
Hypertension (401)
Vitamin deficiency (261)
Functional digestive disorders (564)
Ischemic Heart Disease (411)
Diabetes mellitus (250)
Urinary calculus (594)
Diverticulosis and diverticulitis (562)
Gastrointestinal hemorrhage (578)
Disorders of lipoid metabolism (272)
Renal failure (585)
Gout and other crystal arthropathies (274)
Intestinal infection (8)
Benign neoplasm of colon (208)
Hemorrhoids (455)
Noninfectious gastroenteritis (558)
OR = 2
OR = 5
OR ≥ 10
● Circulatory system diseases
● Endocrine/metabolic diseases
● Diseases of the sense organs
● Hematopoietic diseases
● Infectious diseases
● Mental disorders
● Digestive diseases
● Genitourinary diseases
● Neoplasms
n = 409
n = 2461
n = 15,694

FIGURE 4 Comorbidity network of medical conditions with a lower risk in relation to high-level coffee consumption ($n = 191,399$). Each node represents a medical condition and is labeled with its name and "phecode." The size and color of each node indicate the prevalence and the category of the corresponding medical condition, respectively (see legend). The width of the link represents the strength of the comorbidity association, measured by ORs obtained from an adjusted logistic regression. The network is partitioned into 4 modules using a Louvain algorithm, and nodes belonging to the same module are grouped together and separated from other nodes using dashed lines.

The major strengths of the study include the large sample size and long follow-up based on the UK Biobank. Additionally, the detailed information on sociodemographic and lifestyle factors collected at baseline for all participants enabled us to adjust for many important confounders in the analyses. Finally, by applying PheWAS, along with disease-trajectory and comorbidity network analyses, we for the first time visualized comprehensively the entire picture of subsequent health outcomes of habitual coffee consumption.

Our study also has limitations. First, habitual coffee intake might be affected by health conditions. For instance, as caffeine intake might increase blood pressure (46), individuals with preexisting cardiovascular (24, 47, 48) or gastrointestinal (25) conditions might avoid high-level coffee consumption. We can therefore not rule out the possibility that reverse causality has contributed to the observed associations to some extent. Also, as symptom onset is often earlier than the clinical diagnosis for many chronic diseases (49, 50), the temporal order from exposure to coffee intake to the studied medical conditions is not always clear. However, as similar results were obtained in the sensitivity analyses after excluding individuals with any diagnosis or symptoms of cardiovascular or gastrointestinal conditions at baseline, when restricting the analysis to coffee drinkers (i.e., comparisons amongst individuals with some level of coffee intake), and when starting the follow-up from 2 years after baseline (i.e., 2-year lag time), it is unlikely that our findings can be fully explained by reverse causation. Second, although multiple confounders, such as sociodemographic factors, smoking, alcohol drinking, and BMI, have been considered in the analysis, residual confounding may still be present to some extent due to unmeasured factors, such as overall diet quality and total energy intake. Regardless, due to the observational nature of the present study, it is likely premature to claim the noted associations as causal. Existing Mendelian randomization analyses have failed to demonstrate causal links between coffee intake and different health outcomes (51, 52). These studies have, however, relatively limited the statistical power and suboptimal instrumental variables due to the potential pleiotropy and heterogeneity of the genetic variants used.

We did not study other sources of caffeine intake, such as tea, in the analysis. Future studies with direct measurements of caffeine levels—for instance, through biospecimens or detailed dietary recalls—is therefore warranted to test the hypothesis that the noted associations of high-level coffee intake with different health outcomes are attributed to caffeine. The lack of full primary care data of the UK Biobank population prevented us from analyzing less severe medical conditions not treated by specialist care, and might have led to a delay in the ascertainment of medical conditions first treated in primary care. Finally, the present study is limited to participants of the UK Biobank, who are not directly representative of the entire UK population (53), making it difficult to generalize our findings to the whole United Kingdom or other populations.

In conclusion, this large, prospective study based on the UK Biobank revealed that habitual coffee consumption was associated with lower risks of many medical conditions, especially those in the cardiometabolic and gastrointestinal systems and those related to alcohol use and estrogen regulation.

We thank all the sponsors and team members involved in West China Biomedical Big Data Center and Med-X Center for Informatics, Sichuan University. This research was done using the UK Biobank Resource under Application 54803.

The authors' responsibilities were as follows – FF, HS: were responsible for the study's concept and design; YZ, HY, ZY, YH, YS, YQ: did the data and project management; CH, YZ, WC, XH: did the data cleaning and analysis; CH, FF, HS: interpreted the data; CH, YZ, WC, XH, FF, HS: drafted the manuscript; and all authors: agree to be accountable for all aspects of the work and read and approved the final manuscript. The authors report no conflicts of interest.

Data Availability

Data from UK Biobank are available to all researchers upon application. https://www.ukbiobank.ac.uk/.

References

1. Heckman MA, Weil J, Gonzalez de Mejia E. Caffeine (1, 3, 7-trimethylxanthine) in foods: A comprehensive review on consumption, functionality, safety, and regulatory matters. J Food Sci 2010;75(3):R77–87.
2. Kim Y, Je Y, Giovannucci E. Coffee consumption and all-cause and cause-specific mortality: A meta-analysis by potential modifiers. Eur J Epidemiol 2019;34(8):731–52.
3. Carlstrom M, Larsson SC. Coffee consumption and reduced risk of developing type 2 diabetes: A systematic review with meta-analysis. Nutr Rev 2018;76(6):395–417.
4. Ding M, Bhupathiraju SN, Satija A, van Dam RM, Hu FB. Long-term coffee consumption and risk of cardiovascular disease: A systematic review and a dose-response meta-analysis of prospective cohort studies. Circulation 2014;129(6):643–59.
5. Liu QP, Wu YF, Cheng HY, Xia T, Ding H, Wang H, et al. Habitual coffee consumption and risk of cognitive decline/dementia: A systematic review and meta-analysis of prospective cohort studies. Nutrition 2016;32(6):628–36.
6. Hong CT, Chan L, Bai CH. The effect of caffeine on the risk and progression of Parkinson's disease: A meta-analysis. Nutrients 2020;12(6):1860.
7. Wang A, Wang S, Zhu C, Huang H, Wu L, Wan X, et al. Coffee and cancer risk: A meta-analysis of prospective observational studies. Sci Rep 2016;6:33711.
8. Jensen AB, Moseley PL, Oprea TI, Ellesoe SG, Eriksson R, Schmock H, et al. Temporal disease trajectories condensed from population-wide registry data covering 6.2 million patients. Nat Commun 2014;5: 4022.
9. Han X, Hou C, Yang H, Chen W, Ying Z, Hu Y, et al. Disease trajectories and mortality among individuals diagnosed with depression: A community-based cohort study in UK Biobank. Mol Psychiatry 2021;26(11):6736–46.
10. Hidalgo CA, Blumm N, Barabasi AL, Christakis NA. A dynamic network approach for the study of human phenotypes. PLoS Comput Biol 2009;5(4):e1000353.
11. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK Biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med 2015;12(3):e1001779.
12. UK Biobank, Primary Care Linked Data Version 1.0. 2019. https://biobank.ndph.ox.ac.uk/showcase/showcase/docs/primary_care_data.pdf.
13. Thiru K, Hassey A, Sullivan F. Systematic review of scope and quality of electronic patient record data in primary care. BMJ 2003;326(7398):1070.
14. Burns EM, Rigby E, Mamidanna R, Bottle A, Aylin P, Ziprin P, et al. Systematic review of discharge coding accuracy. J Public Health 2012;34(1):138–48.
15. Vomero ND, Colpo E. Nutritional care in peptic ulcer. Arq Bras Cir Dig. 2014;27(4):298–302.
16. Katz PO, Gerson LB, Vela MF. Guidelines for the diagnosis and management of gastroesophageal reflux disease. Am J Gastroenterol 2013;108(3):308–28; quiz 29.
17. Hecht I, Achiron A, Man V, Burgansky-Eliash Z. Modifiable factors in the management of glaucoma: A systematic review of current evidence. Graefes Arch Clin Exp Ophthalmol. 2017;255(4):789–96.
18. Brown AC, Rampertab SD, Mullin GE. Existing dietary guidelines for Crohn's disease and ulcerative colitis. Exp Rev Gastroenterol Hepatol 2011;5(3):411–25.
19. Cozma-Petrut A, Loghin F, Miere D, Dumitrascu DL. Diet in irritable bowel syndrome: What to recommend, not what to forbid to patients! World J Gastroenterol 2017;23(21):3771–83.
20. UK Biobank, Reason lost to follow-up. 2022. Available from: https://biobank.ndph.ox.ac.uk/showcase/field.cgi?id=190.
21. Wu P, Gifford A, Meng X, Li X, Campbell H, Varley T, et al. Mapping ICD-10 and ICD-10-CM codes to phecodes: Workflow development and initial evaluation. JMIR Med Inform 2019;7(4):e14325.
22. Townsend P, Phillimore P, Beattie A. Health and deprivation: Inequality and the North. Kent, UK: Routledge; 1988.
23. Cassidy S, Chau JY, Catt M, Bauman A, Trenell MI. Cross-sectional study of diet, physical activity, television viewing and sleep duration in 233,110 adults from the UK Biobank; The behavioural phenotype of cardiovascular disease and type 2 diabetes. BMJ Open 2016;6(3):e010038.
24. Hypponen E, Zhou A. Cardiovascular symptoms affect the patterns of habitual coffee consumption. Am J Clin Nutr 2021;114(1): 214–9.

25. Soroko S, Chang J, Barrett-Connor E. Reasons for changing caffeinated coffee consumption: The Rancho Bernardo study. J Am Coll Nutr 1996;15(1):97–101.
26. Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. J R Stat Soc B Methodol 1995;57(1):289–300.
27. Grambsch PM, Therneau T. Proportional hazards tests and diagnostics based on weighted residuals. Biometrika 1994;81(3):515–26.
28. De Meo P, Ferrara E, Fiumara G, Provetti A. Generalized Louvain method for community detection in large networks. In: Sebastián Ventura, Ajith Abraham, Krzysztof Cios, Cristóbal Romero, Francesco Marcelloni, José Manuel Benítez, Eva Gibaja. Proceedings of the 11th International Conference on Intelligent Systems Design and Applications. IEEE; Cordoba, Spain. 2011: 88–93.
29. Pauli Virtanen, Ralf Gommers, SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods 2020;17(3):261–72.
30. Seabold Skipper, Josef Perktold. Statsmodels: Econometric and statistical modeling with python. 2010; Proceedings of the 9th Python in Science Conference.
31. Davidson Pilon. Lifelines: survival analysis in Python. Journal of Open Source Software 2019;4(40):1317.
32. Zhang Y, Yang T, Zeng C, Wei J, Li H, Xiong YL, et al. Is coffee consumption associated with a lower risk of hyperuricaemia or gout? A systematic review and meta-analysis. BMJ Open 2016;6(7): e009809.
33. Kennedy OJ, Roderick P, Buchanan R, Fallowfield JA, Hayes PC, Parkes J. Systematic review with meta-analysis: Coffee consumption and the risk of cirrhosis. Aliment Pharmacol Ther 2016;43(5):562–74.
34. Zhang YP, Li WQ, Sun YL, Zhu RT, Wang WJ. Systematic review with meta-analysis: Coffee consumption and the risk of gallstone disease. Aliment Pharmacol Ther 2015;42(6):637–48.
35. Nehlig A. Effects of coffee on the gastro-intestinal tract: A narrative review and literature update. Nutrients 2022;14(2):399.
36. Cornelis MC, Weintraub S, Morris MC. Caffeinated coffee and tea consumption, genetic variation and cognitive function in the UK Biobank. J Nutr 2020;150(8):2164–74.
37. Larsson SC, Orsini N. Coffee consumption and risk of dementia and Alzheimer's disease: A dose-response meta-analysis of prospective studies. Nutrients 2018;10(10):1501.
38. Godos J, Pluchinotta FR, Marventano S, Buscemi S, Li Volti G, Galvano F, et al. Coffee components and cardiovascular risk: Beneficial and detrimental effects. Int J Food Sci Nutr 2014;65(8):925–36.
39. Astrup A, Toubro S, Cannon S, Hein P, Breum L, Madsen J. Caffeine: A double-blind, placebo-controlled study of its thermogenic, metabolic, and cardiovascular effects in healthy volunteers. Am J Clin Nutr 1990;51(5):759–67.
40. van Dijk AE, Olthof MR, Meeuse JC, Seebus E, Heine RJ, van Dam RM. Acute effects of decaffeinated coffee and the major coffee components chlorogenic acid and trigonelline on glucose tolerance. Diabetes Care 2009;32(6):1023–5.
41. Nehlig A. Interindividual differences in caffeine metabolism and factors driving caffeine consumption. Pharmacol Rev 2018;70(2): 384–411.
42. Schliep KC, Schisterman EF, Mumford SL, Pollack AZ, Zhang C, Ye A, et al. Caffeinated beverage intake and reproductive hormones among premenopausal women in the Biocycle study. Am J Clin Nutr 2012;95(2):488–97.
43. Klatsky AL, Morton C, Udaltsova N, Friedman GD. Coffee, cirrhosis, and transaminase enzymes. Arch Intern Med 2006;166(11): 1190–5.
44. Kennedy OJ, Fallowfield JA, Poole R, Hayes PC, Parkes J, Roderick PJ. All coffee types decrease the risk of adverse clinical outcomes in chronic liver disease: A UK Biobank study. BMC Public Health 2021;21(1):970.
45. Kendler KS, Schmitt E, Aggen SH, Prescott CA. Genetic and environmental influences on alcohol, caffeine, cannabis, and nicotine use from early adolescence to middle adulthood. Arch Gen Psychiatry 2008;65(6):674–82.
46. Noordzij M, Uiterwaal CS, Arends LR, Kok FJ, Grobbee DE, Geleijnse JM. Blood pressure response to chronic intake of coffee and caffeine: A meta-analysis of randomized controlled trials. J Hypertens 2005;23(5):921–8.
47. De Giuseppe R, Di Napoli I, Granata F, Mottolese A, Cena H. Caffeine and blood pressure: A critical review perspective. Nutr Res Rev 2019;32(2):169–75.
48. Zhou A, Hypponen E. Long-term coffee consumption, caffeine metabolism genetics, and risk of cardiovascular disease: A prospective analysis of up to 347,077 individuals and 8368 cases. Am J Clin Nutr 2019;109(3):509–16.
49. Samuels TA, Cohen D, Brancati FL, Coresh J, Kao WH. Delayed diagnosis of incident type 2 diabetes mellitus in the ARIC study. Am J Manag Care 2006;12(12):717–24.
50. Breen DP, Evans JR, Farrell K, Brayne C, Barker RA. Determinants of delayed diagnosis in Parkinson's disease. J Neurol 2013;260(8):1978–81.
51. Cornelis MC, Munafo MR. Mendelian randomization studies of coffee and caffeine consumption. Nutrients 2018;10(10):1343.
52. Nicolopoulos K, Mulugeta A, Zhou A, Hypponen E. Association between habitual coffee consumption and multiple disease outcomes: A Mendelian randomisation phenome-wide association study in the UK Biobank. Clin Nutr 2020;39(11): 3467–76.
53. Fry A, Littlejohns TJ, Sudlow C, Doherty N, Adamska L, Sprosen T, et al. Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. Am J Epidemiol 2017;186(9):1026–34.