

High Coffee Consumption, Brain Volume and Risk of Dementia and Stroke

Kitty Pham (BMedRadSc)^a, Anwar Mulugeta (PhD)^{a,b,c}, Ang Zhou (PhD)^{a,b}, John T O'Brien (DM FRCPsych)^d, David J Llewellyn (PhD)^{e,f} and Elina Hyppönen (PhD)^{a,b*}

^a *Australian Centre for Precision Health, University of South Australia, Clinical & Health Sciences, Adelaide, Australia.*

^b *South Australian Health and Medical Research Institute, Adelaide, Australia.*

^c *Department of Pharmacology and Clinical Pharmacy, College of Health Science, Addis Ababa University, Addis, Ababa, Ethiopia.*

^d *Department of Psychiatry, University of Cambridge, Cambridge, England.*

^e *College of Medicine and Health, University of Exeter, Devon, England.*

^f *Alan Turing Institute, London, England.*

This is an Accepted Manuscript of an article published by Taylor & Francis in NUTRITIONAL NEUROSCIENCE on 24 Jun 2021, available online:

<http://www.tandfonline.com/10.1080/1028415X.2021.1945858>.

*Corresponding Author:

Professor Elina Hyppönen, Australian Centre for Precision Health, University of South Australia Cancer Research Institute, South Australian Health & Medical Research Institute (SAHMRI) Level 8, GPO Box 2471, Adelaide SA 5001, Australia.

Contact number: +61 8 8302 2518

Email: elina.hypponen@unisa.edu.au

Biographical notes for the authors:

Kitty Pham is a current PhD candidate with the Australian Centre for Precision Health, Clinical & Health Sciences, University of South Australia. Her background is in medical radiation. She is currently registered as a Medical Radiation Practitioner (Radiation Therapist). Her PhD research project investigates the impacts of serum lipids and dietary fats on brain health outcomes. She is interested in MRI neuroimaging indicators for dementia disease risk and the application of large population cohort data to clinical settings.

Dr Anwar Mulugeta is a Research Associate at Nutritional and Genetic Epidemiology Research Group and Steering committee member at the Australian Centre for Precision Health. He has training in Genetic Epidemiology, Medical Pharmacology and Pharmacy. He is currently affiliated with Department of Pharmacology and Clinical Pharmacy, Addis Ababa University (AAU), Ethiopia. He previously worked as Clinical Trial Coordinator for AAU and Massachusetts General Hospital collaborative project. He focuses on understanding the contributions of genetics, environment and the interplay between them on disease risk including cognitive impairment, dementia, depression, and other health outcomes. He applies various genetic and bioinformatic approaches including mendelian randomizations, phenome-wide associations, genome-wide associations, gene-environment interactions, and machine learning.

Dr Ang Zhou has training in medical sciences and biostatistics. He is working as a research associate in the nutritional and genetic epidemiology group at the Australian Centre for Precision Health. Ang applies methodologies in genetic epidemiology to understand causal associations between modifiable lifestyle factors and health outcomes related to cognitive and cardiometabolic functions. His research also involves understanding gene-environment interplay on health outcomes.

Professor John T O'Brien is a Professor at the Department of Psychiatry at the University of Cambridge. His research interests include the role of biomarkers, especially MRI, SPECT and PET imaging, in the differential and early diagnosis of dementia, including identifying those 'at risk' of future cognitive decline and developing markers of onset and progression of disease. He is interested in late-life depression, especially the role of vascular and inflammatory factors in precipitating and perpetuating depression in late-life, and the ability of vascular interventions to improve or prevent depressive and cognitive symptoms. Professor O'Brien is also the NIHR Clinical Research Network National Specialty Lead for Dementia and a member of several clinical guideline groups which undertake a number of multimodal MR, ligand PET and MEG studies, longitudinal clinical and e-record studies, and a number of clinical studies in dementia, including trials of pharmacological and non-pharmacological management.

Professor David J Llewellyn is a Professor at the University of Exeter Medical School and a Fellow at the Alan Turing Institute. He also holds an honorary contract with Devon Partnership NHS Trust. He has advanced training in epidemiology and data science. His research aims to enhance the timely detection of dementia, with a focus on developing strategies for primary and secondary prevention. He uses a combination of evidence synthesis, data science and machine learning to develop new translational insights to identify more effective interventions and enhance the diagnostic pathway for dementia. He is an expert on the evaluation of cognitive function and dementia and is a member of the scientific advisory boards of the English Longitudinal Study of Ageing and the UCL Centre for Longitudinal Studies. He sits on Alzheimer's Research UK's Grant Review Board and their Clinical Policy Advisory Panel. David is also the Exeter Institute for Data Science and Artificial Intelligence Clinical Theme Lead and the Turing Exeter University Clinical Lead. He sits on the Steering Committee and leads the Clinical Advisory Group of the Early Detection of Neurodegeneration Initiative and is Director of the DEMON Network.

Professor Elina Hyppönen is the Director of the Australian Centre for Precision Health, and a Professor in Nutritional and Genetic Epidemiology at the University of South Australia. She is also Senior Principal Research Fellow at the South Australian Health and Medical Research Institute, Honorary Professor in Epidemiology at University College London, and Adjunct Professor at Tampere University. Professor Hyppönen has an interdisciplinary academic background, with academic qualifications in epidemiology, medical statistics, nutrition, and public health. She leads the Nutritional and Genetic Epidemiology group which has a focus on using genetic tools to inform on dietary and lifestyle guidelines for optimal health. She has a long-term research interest in life-course and intergenerational epidemiology, and an extensive track record in gene and risk factor discovery. Her current interests are related to implementing genome wide analyses and systems epidemiology approaches to establish effective strategies for disease prediction and prevention.

High Coffee Consumption, Brain Volume and Risk of Dementia and Stroke

Background: Coffee is a highly popular beverage worldwide, containing caffeine which is a central nervous system stimulant.

Objectives: We examined whether habitual coffee consumption is associated with differences in brain volumes or the odds of dementia or stroke.

Methods: We conducted prospective analyses of habitual coffee consumption on 398,646 UK Biobank participants (age 37-73 years), including 17,702 participants with MRI information. We examined the associations with brain volume using covariate adjusted linear regression, and with odds of dementia (4,333 incident cases) and stroke (6,181 incident cases) using logistic regression.

Results: There were inverse linear associations between habitual coffee consumption and total brain (fully adjusted β per cup -1.42, 95% CI -1.89, -0.94), grey matter (β -0.91, 95% CI -1.20, -0.62), white matter (β -0.51, 95% CI -0.83, -0.19) and hippocampal volumes (β -0.01, 95% CI -0.02, -0.003), but no evidence to support an association with white matter hyperintensity (WMH) volume (β -0.01, 95% CI -0.07, 0.05). The association between coffee consumption and dementia was non-linear ($P_{\text{non-linearity}}=0.0001$), with evidence for higher odds for non-coffee and decaffeinated coffee drinkers and those drinking >6 cups/day, compared to light coffee drinkers. After full covariate adjustment, consumption of >6 cups/day was associated with 53% higher odds of dementia compared to consumption of 1-2 cups/day (fully adjusted OR 1.53, 95% CI 1.28, 1.83), with less evidence for an association with stroke (OR 1.17, 95% CI 1.00, 1.37, $p=0.055$).

Conclusion: High coffee consumption was associated with smaller total brain volumes and increased odds of dementia.

Keywords: coffee; brain volume; dementia; stroke; volumetric MRI; UK Biobank; prospective cohort

Subject classification codes: N/A

Introduction

Coffee is among the most popular non-alcoholic beverage after water, with the global consumption exceeding nine billion kilograms each year (1). Coffee contains caffeine, which is a central nervous system stimulant (2). Due to high lipid solubility and structural similarity to adenosine, caffeine readily crosses the blood brain barrier to competitively bind with adenosine receptors. It stimulates the adrenal gland secretion of catecholamines, an excitatory neurotransmitter mediating the effects on cognition, emotion, and motor function (3). Hence, caffeine is known to have broad physiological influences, affecting sleep patterns, mood, motor activity, heart rate, body core temperature and oxygen consumption (4, 5).

There are mechanisms which could mediate a potential role of coffee on brain morphology and the risk of dementia and stroke. For example, brain volume changes may be caused by the antagonist binding of caffeine to adenosine receptors, inducing morphological change in mossy-fibres or pyramidal cells and a reduction of information transmission (6). However, there are only a handful of studies looking into the associations between coffee intake and brain volume (7-9). Most studies are longitudinal, varying in size between 45 and 2,914 participants, and report inconsistent findings which suggest either a positive (7), U-shaped (8) or negative (9) association between coffee consumption and regional or total brain volumes. The most recent study is a randomised control trial, which reports an association between higher caffeine intake and larger reductions in GMV (10). Another aspect of interest within the brain are the white matter hyperintensities (WMH), which are small lesions frequently associated with local small vessel damage, implicated in increased stroke and dementia risk (11). In a recent study, higher coffee consumption was associated with a lower WMH volume in older women while no association was seen in men (12). They suggested that the blockage of adenosine receptors by higher coffee intakes may decrease amyloid accumulation in the brain to reduce production of WMH.

A Mendelian randomisation study has provided evidence for a potential adverse effect of higher coffee consumption on the risk of Alzheimer's disease, with each cup of coffee increasing the risk by 26% (13). However, observational evidence is inconsistent, and of the 11 most recent longitudinal cohort studies on habitual coffee intake and dementia risk, five supported a possible protective effect (14-18), whereas six studies found no association (19-24). Also, the majority (four of seven) of the recent studies on stroke have reported an inverse relationship with coffee consumption (25-28), while three found no association (29-31). There is also some evidence for a U-shaped association with respect to both outcomes, where moderate drinkers were reported to have the lowest stroke risk, with some elevation both for non-drinkers and high coffee consumers (32, 33). In the current study, we used information from up to 398,646 participants in the UK Biobank (17,702 participants with information on brain neuroimaging biomarkers) to investigate the association between habitual patterns of coffee consumption and brain neuroimaging biomarkers and the odds of dementia and stroke.

Materials and Methods

Study Population

The UK Biobank is a long-term prospective epidemiological study of 502,504 participants with deep genotypic and phenotypic data (34). Individuals aged 37-73 years were recruited in 22 assessment centres across the United Kingdom, between March 13, 2006 and October 1, 2010. During baseline assessment, information was collected using touchscreen questionnaires, verbal interview, and physical examination, including a collection of blood, urine, and saliva samples. An imaging sub-study was incorporated into the UK Biobank in 2014 with the target of conducting brain, heart and body MRI imaging on 100,000 participants (35). The analyses on disease outcomes were conducted in up to 398,646 participants who were of white British ancestry and who had information on habitual coffee

consumption, while the brain volume analyses conducted in up to 17,702 participants with valid data on brain imaging (Figure 1). We allow for a maximum of two genetically related individuals (36), with further exclusions based on discrepancies between self-reported and genetically determined sex, and history of stroke or dementia at baseline.

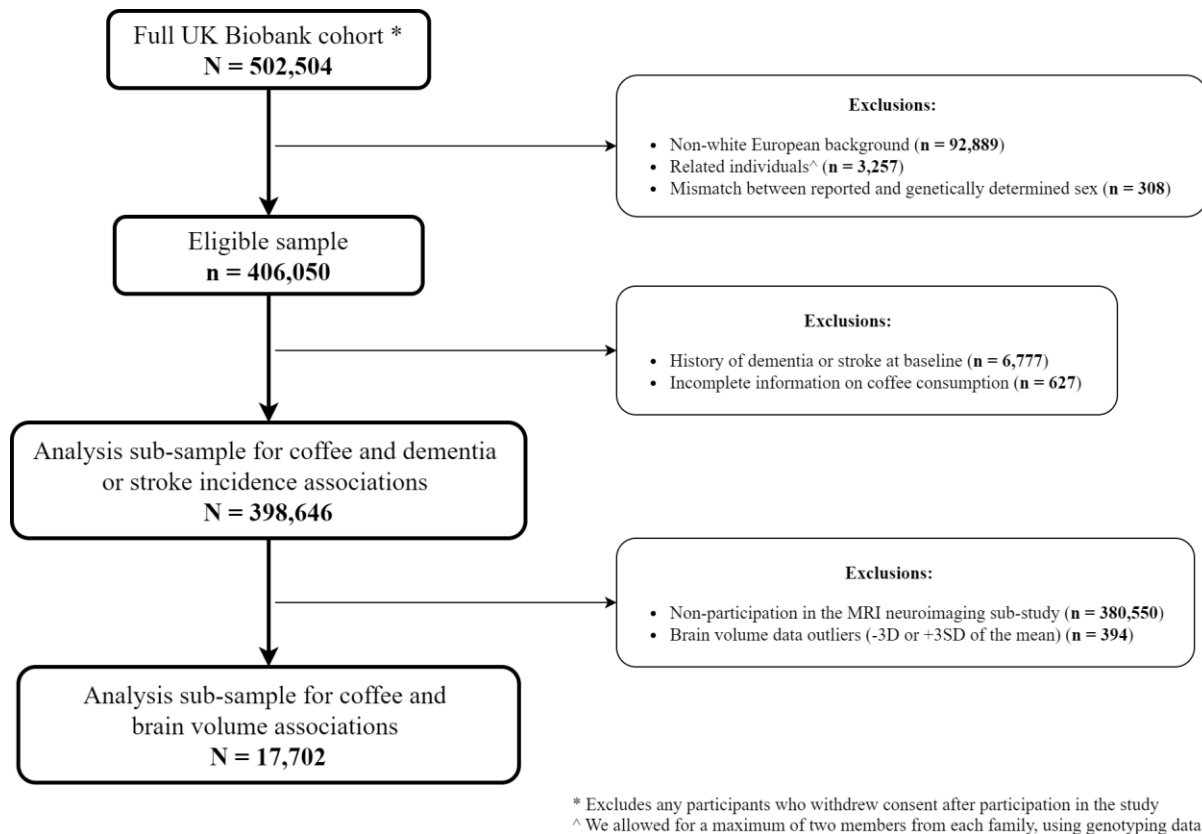


Figure 1. Flowchart showing the sample restriction from the UK Biobank cohort to the final analysis samples.

Ethics

As a secondary analysis of data, the current study relies upon the consents of subjects at their participation in the UK Biobank data collection studies and poses no additional risks to the participants. The UK Biobank, governed by the UK Biobank Ethics and Governance Council, holds Human Tissue Authority licencing and Research Tissue Bank approval, and oversees all access and use of the database (37). Ethical approval for the UK Biobank was granted by the National Information Governance Board for Health and Social Care and North West

Multicentre Research Ethics Committee (11/NW/0382). Participants provided electronic consent to use their anonymized data and sample for health-related research and for the UK Biobank to access their health-related records (34).

Coffee Consumption

Habitual coffee intake was self-reported as cups per day, as part of the touchscreen questionnaire at baseline with the question “How many cups of coffee do you drink a day?” (38). Participants were asked to include decaffeinated (decaf) coffee intake in their response. Decaffeinated consumers were identified from response to the question “What type of coffee do you usually drink?”. Coffee consumption was grouped in seven categories, including non-drinkers, decaffeinated coffee drinkers, and caffeinated coffee drinkers consuming <1 cup/day, 1-2 cups/day, 3-4 cups/day, 5-6 cups/day and >6 cups/day. “Coffee drinkers” in the text refer to participants who report intake of caffeinated coffee. For sensitivity analyses we constructed separate indicators restricting the data to decaffeinated coffee drinkers and according to tea consumption, using intake categories as above.

Total and regional brain volumes

Total and regional brain volumes were acquired in the MRI imaging sub-study, commencing 4 – 6 years after baseline assessment. Brain MRI scans were conducted in accordance with the UK Biobank Brain Imaging Protocol (39). A Siemens Skyra 3T scanner (running VD13A SP4 software) was used, with a 32-channel RF-receive head coil and a large 256mm superior-inferior field of view, extending from the top of the cranium to the neck/mouth region. The T1-weighted structural MRI had a 5-minute imaging sequence with TR 2000.0ms, TE 2.01ms and a 1x1x1mm spatial resolution (40). The T2-FLAIR structural imaging was undertaken in 6 minutes with TR 5000.0ms, TE 395.0ms and spatial resolution of 1.05x1x1mm.

After acquisition, images underwent an automated pipeline developed by the UK Biobank, based on the FMRIB software library (FSL) of MRI brain imaging data (41). The general processing pipeline involved reconstruction of data from DICOM to NIFTI format, organisation of data files, anonymization of images (digital masking of face and ears) and full 3D gradient distortion correction (40). Total brain, white matter, grey matter and hippocampal volumes were generated from processed T1 images, while combined analyses of T1 and T2-FLAIR data quantified WMH volume. Further details of each processing pipeline have been comprehensively documented in the UK Biobank Brain Imaging Protocol (39). We excluded extreme outliers with brain volumes below 3 SD or above 3 SD of the mean (394 participants).

Dementia and stroke incidence

Information on incident dementia and stroke were collated by the UK Biobank using primary care data, hospital admission electronic health records (EHR), national death registers and self-reported medical conditions in the UK Biobank touchscreen questionnaire (42). We used information on centrally adjudicated diagnoses on dementia and stroke, including sensitivity analyses according to dementia subtypes (vascular dementia, Alzheimer's disease, and frontotemporal dementia) (43, 44). Incident dementia and stroke cases were defined as any case identified after the baseline assessment and before the end of follow-up in February 1, 2018.

Covariates

Covariates were all collected during the UK Biobank baseline assessment (34), with full description provided in Supplementary Table 1. Townsend deprivation index reflecting area deprivation was determined based on the participants postcode. Most socioeconomic (highest level of education, income and employment), health related (history of common mental

disorders (CMD), sleep duration, insomnia, stressful events within the past 2 years, and long-standing illness) and other covariates (water intake, smoking, alcohol consumption, intensity of physical activity, tea consumption, stressful life-events, processed meat, and fruit consumption) were self-reported by participants in the touchscreen questionnaire. Weight (kg) and height (m) were measured and used to calculate body mass index (BMI, weight (kg)/height (m)²). BMI was grouped according to the World Health Organisation nutritional status classifications as underweight (<18.5 kg/m²), normal (18.5 – 25 kg/m²), overweight (25 – 30 kg/m²) and obese (≥30 kg/m²) (45). Blood pressure was averaged from two automated readings, taken a few moments apart on an Omron device, and adjusted for any blood pressure lowering medications (46). Whole body water mass was estimated from impedance measures and serum Urea Creatine ratio was measured from blood samples taken at baseline using Beckman Coulter AU5800 analysers (47, 48).

Statistical analyses

We explored the association between coffee consumption and total brain volume using linear regression and associations with incident dementia and stroke with logistic regression. We checked for non-linear relationships between coffee and our outcomes using log likelihood ratio tests comparing models with coffee as a categorical (non-linear) variable and coffee as a linear variable. We tested for variation in the association between coffee and brain volumes or disease odds by age and sex, using a multiplicative term (e.g. coffee x age), and in the absence of interaction present results for age groups and both sexes combined. In order to account for relatedness between participants, models were weighted (1 - kinship coefficient) (36). We used light coffee drinkers (1-2 cups/day), instead of non-drinkers, as the reference group to avoid bias from individuals with poor health, who may avoid coffee due to their health status (49, 50).

As sensitivity analyses, we repeated all analyses using information on decaffeinated coffee consumption and tea consumption as exposures. Since older age and smoking are known to be risk factors for dementia and stroke, we ran the analyses in a subset of participants aged over 60 years and for a subset of non-smokers only. For dementia, we conducted analyses stratified by dementia subtypes. Finally, to assess the role of reverse causation in the dementia findings, we conducted further sensitivity analyses allowing for a 4-year lag time between baseline assessment and dementia diagnosis.

Covariate adjustments were conducted in stages, starting with a basic model (sex, age, age squared, assessment centre), and then adjusted further for extended covariates including socioeconomic (education, Townsend deprivation index, income and employment), health (history of CMD, systolic blood pressure, sleep duration, insomnia, and long standing illness), anthropometric (height, BMI, and BMI squared), hydration indicators (water intake, water mass and Urea : Creatine ratio) or lifestyle factors (smoking, alcohol consumption, type of physical activity, tea consumption, stressful life-events, processed meat, and fruit consumption). Fully adjusted models included all the above covariates.

To account for missing covariate information, we used multiple imputation chained equation (MICE), creating fifteen imputed datasets with the aim of minimising sampling variability within the imputation process (51). Imputed variables included education, Townsend deprivation index, income, employment status, systolic blood pressure, sleep duration, insomnia, history of long-standing illnesses, height, BMI, water intake, total water body mass, UCR, smoking, alcohol, physical activity, tea consumption, stressful events, processed meat intake and fresh fruit intake. Each variable was missing in only <4.74% of the participants, except systolic blood pressure (8.76%) and income (13.84%). Analyses were repeated restricting participants to those with complete data on all covariates (n=280,330) and

as findings were not affected only data from the full sample is presented. All analyses were performed using the STATA SE version 14.1 software (52).

Results

Baseline characteristics of the UK Biobank

Within the full study population, 54.3% were female, and 66.6% reported no long-standing illness (Table 1). Coffee consumption was associated with the range of health and lifestyle covariates (Table 1, Supplementary Table 2). There were some differences in brain volumes by sex, age, BMI, education, and the reported history of depression and long-standing illness, with the same factors which tended to be associated with smaller brain volume also associated with higher incidence of dementia and stroke, except for sex and BMI. Differences in regional brain volume showed in part similar patterns to total brain volume indicators. Patterns of brain volume and dementia/stroke incidence by the full range of covariates can be found in Supplementary Table 2.

Coffee consumption and brain volume

Coffee consumption had a linear inverse association with total brain volume, with consistent patterns for grey matter, white matter, and hippocampal volumes (for all, $p_{\text{non-linearity}} > 0.42$) (Table 2). Compared to the basic model adjusted for age, sex and assessment centre, further inclusion of anthropometric measures, dehydration indicators, and lifestyle factors, each attenuated the associations, but evidence for an inverse trend remained even after full adjustment (Supplementary Table 3). These inverse trends between coffee consumption and brain volume indicators did not differ by sex or age (for all tests, $P_{\text{interaction}} > 0.30$). Coffee consumption was not associated with WMH volume before or after covariate adjustment. There was little evidence for an association between decaffeinated coffee intake and brain

volume measures (Supplementary Table 4).

Coffee consumption and odds of dementia and stroke

There was non-linear association between coffee consumption and the odds of dementia, with slightly higher odds seen with non-coffee or decaf drinkers, and more notable increases for participants in the highest coffee consumption categories ($p_{\text{non-linearity}}=0.0001$) (Figure 2). For stroke, the association between higher coffee consumption, but not coffee abstinence, was explained by covariate adjustment. Association between high coffee consumption and dementia was less affected by adjustment, and odds remained elevated for participants drinking >6 cups/day even after accounting for socioeconomic, anthropometric, lifestyle, dehydration, sleep, and disease related covariates. No differences in estimates were found among men and women or among different age groups (for all tests, $P_{\text{interaction}}>0.22$).

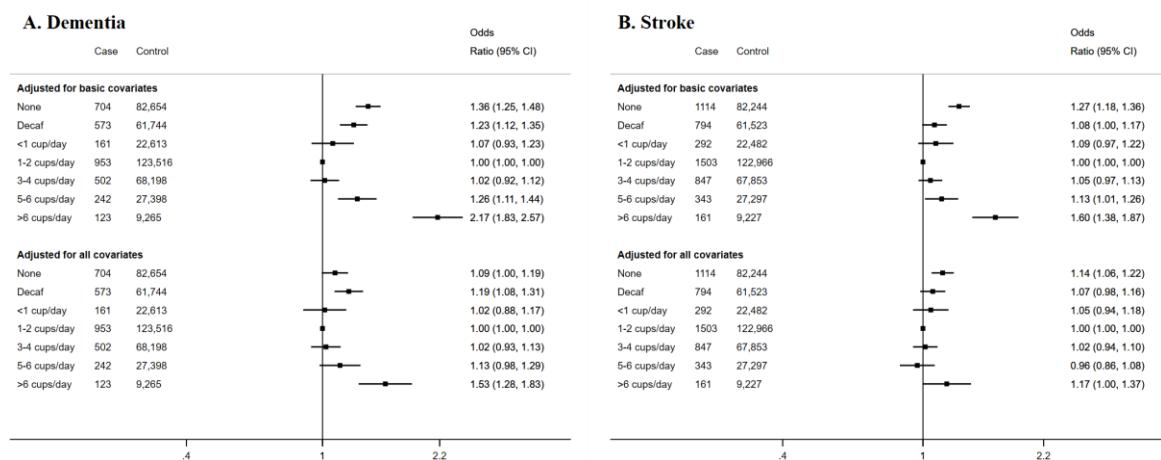


Figure 2. Association of coffee consumption with incidence of dementia (panel A) and stroke (panel B), using logistic regression models.

Basic covariates: age, age square, sex, and assessment centre.

Socioeconomic status: basic covariates + education, Townsend deprivation index, income, and employment status.

Diseases status: basic covariates + depression, systolic blood pressure, sleep duration, insomnia, and long-standing illnesses.

Anthropometric measures: basic covariates + height, body mass index, and body mass index square.

Dehydration indicators: basic covariates + water intake, whole body water mass, and urine-creatinine ratio.

Lifestyle factors: basic covariates + smoking, alcohol, intensity of physical activity, tea consumption, stress-related lifestyle factors, processed meat intake, and fresh fruit intake.

All covariates: adjusted for all covariates listed above.

Analyses in decaffeinated coffee drinkers found the same pattern of association as the main analyses on caffeinated coffee (Supplementary Figure 1). Dementia and stroke analyses in participants aged over 60 years (Supplementary Figure 2) and participants who were non-smokers (Supplementary Figure 3) both showed similar results to the main findings. Sensitivity analyses separating dementia diagnoses by sub-type confirmed the association between coffee intake and Alzheimer's disease, with inconclusive findings for vascular and frontotemporal dementia (Supplementary Figure 4). Analyses allowing for a 4-year lag time between coffee consumption and dementia incidence showed the same pattern of association as the main analyses (Supplementary Figure 5).

Sensitivity analyses on tea consumption did not provide any evidence for association with brain volumes (Supplementary Table 5) or the odds of dementia or stroke (Supplementary Figure 6), before or after covariate adjustment.

Discussion

Coffee is one of the most consumed beverages in the world after water, with potentially large implications for public health. While the immediate effects of caffeinated coffee on alertness and cognition appear typically beneficial (53), our results and an earlier genetic study suggest increased odds of dementia by high coffee intakes (13). Furthermore, in line with a recent placebo controlled trial demonstrating a dose-dependent GMV lowering effect for caffeine administration, we also found evidence for linear inverse associations between higher habitual coffee consumption and total, grey matter, white matter, and hippocampal volumes, providing further support for effects of coffee on the brain (10). While our study cannot confirm underlying causality of association, these results warrant carefully controlled studies to clarify the beneficial and adverse effects of coffee on the brain.

Our finding for a possible effect of higher coffee consumption and GMV is supported by a recent double-blind cross-over RCT (10). The study was conducted in 20 healthy males aged 18 – 35 years who identified as habitual caffeine consumers. Study was carefully controlled and the participants underwent 9 days with fixed sleep schedules and self-administered caffeine or placebo capsules (3 doses of 150 mg per day), followed by sleep electroencephalography (EEG) and T1-weighted MRI imaging in a controlled laboratory environment on the 10th day. Individual caffeine metabolism and compliance with the self-administered caffeine schedule was assessed by analysing fingertip sweat for caffeine and caffeine metabolites. The results were independent of sleep, as shown by sleep EEG testing and sleep diary monitoring. Despite the small sample size, the study accounts for individual differences in caffeine metabolism by conducting within-subject comparisons. These findings support the association seen in our observational study and suggest a possibly causal explanation. Additionally, a visual working memory test (N-back test) was conducted 4 times throughout the laboratory session and compared between caffeine and placebo conditions. The study found decreases in working memory test performance with caffeine vs. placebo, in conjunction with the observed GMV decreases, tentatively proposing a detrimental effect on brain functioning with smaller brain volumes, which would support our findings with high coffee consumption and an increased odds of dementia.

To our knowledge, the largest of the earlier studies looking at the effects of habitual coffee consumption on the brain was conducted in the Rotterdam cohort (n=2,914) which is similar to UK Biobank with respect to age (mean age 59 years) (9). Compatible with our findings, they reported an inverse linear association between habitual coffee consumption and total hippocampal volume, and directionally consistent but non-significant associations for total and grey matter volumes. Other studies are notably smaller, and for example, in a group of young women (n=45, 13-30 years), coffee consumption was not associated with total brain

volumes but was suggested to have a U-shaped association with hippocampal volume, with higher values seen both for non-drinkers and higher consumption groups (8). While our middle to older aged population did not suggest age related interactions with respect to the associations of coffee and brain volume indicators, a study on elderly diabetics (n=185, 64-86 years) found no evidence for an association with white matter volume but a possible age dependent association with grey matter volume (54). Again, we did not observe an association between coffee intake and WMH, while an earlier study observed some evidence for an inverse relation in women but not for men (n=317 and 324, respectively) (12).

A recent meta-analysis did not find convincing evidence for a linear association between coffee intake and dementia while in line with our findings there was some suggestion for a U-shaped association and an elevated risk for those consuming at least five cups per day (p-curvature=0.08) (22). Evidence for a potential need for caution with respect to dementia risk, have been provided by large-scale Mendelian randomization analyses (17,008 cases and 37,154 controls), where higher genetically instrumented coffee intake was seen to lead to an increased risk of Alzheimer's disease (13). In contrast, genetic evidence for the role of coffee intake on stroke risk remains inconclusive (32), and while we observed some evidence for a U-shaped association, evidence for an association with heavy coffee consumption was not strong after full covariate adjustment. In studies of coffee consumption, non-drinkers often have a higher disease risk compared to coffee drinkers, which may reflect reverse causality arising from abstinence from coffee consumption due to poor reaction (raised blood pressure, increased heart rate) or existing health concerns (5, 18, 25, 27, 28, 32, 50).

Competitive binding of caffeine to adenosine receptors in the brain is a well-known physiological mechanism and has been implicated in coffee consumption studies as the cause

of changes in brain morphology (7). Changes in brain volume are theorised to be caused by structural change in mossy-fibres and pyramidal cells, initiated by the antagonist binding of caffeine (55). Other effects could be mediated through the cardiovascular system, with relevance for both vascular dementia and AD (56). For example, the consumption of unfiltered coffee (and other types of coffee, to a lesser extent) increases serum total cholesterol, LDL cholesterol and triglyceride levels (57). These lipid abnormalities are a leading risk factor for atherosclerosis, which has been associated with AD, independently of other brain pathologies (58). However, we did not find strong evidence for caffeine driven mechanisms, as the patterns seen with decaffeinated coffee were similar to patterns with caffeinated coffee, and tea consumption (containing varying amounts of caffeine) was not associated with any of the outcomes.

With respect to volumetric brain neuroimaging data, the sample size is an order of magnitude greater than the largest of earlier studies (9). Another major strength with our study was the availability of a range of socioeconomic, disease status, anthropometric, dehydration related, and lifestyle covariates, allowing for extensive confounder adjustments. However, we cannot fully exclude the possibility of residual confounding, as full information on diet was only available for a sub-sample, and factors such as dehydration were measured at baseline rather than at the same time as brain MRI. The use of self-reported data and the possibility of lifestyle changes between baseline and MRI or covariate measurement is a further limitation. Smoking is an important confounder, but sensitivity analyses restricting data to non-smokers provided identical results. Long follow-up, although advantageous for the study, does not exclude reverse causation, which could arise as olfactory and taste dysfunction is an early symptom of dementia, and may cause sufferers to consume more coffee (59, 60). However, we conducted sensitivity analyses excluding information from the first four years after baseline, obtaining similar results. We also acknowledge that

information on baseline mild cognitive impairment was not available within the UK Biobank. Disease diagnoses were determined by ICD codes based on information from hospital inpatient records and primary health care data, and while that represents a good pragmatic solution for a study of this scale, we cannot exclude potential for related misclassification.

The UK biobank is subject to a healthy volunteer bias, and the sample is not representative of the general population (61). Associations in other ethnic populations should be further investigated in more diverse samples as the current study is restricted to White British individuals. The convenience and chain sampling methods may introduce bias, though this has been minimised by variation of socioeconomic status, and population density between recruitment areas (34). Importantly for the current study, there also remains ambiguity in the individual interpretation of a cup serving of coffee, however, this is a broader issue as the standardisation of a unit of coffee is yet to be established in literature, particularly for observational studies (62). Furthermore, we only had information about the type of coffee *usually* consumed by participants (preventing us from identifying individuals drinking both caffeinated and decaffeinated coffee), and there also were no details for different types of coffee and tea or other potential sources of caffeine, limiting our ability to examine caffeine related mechanisms.

In conclusion, higher coffee consumption was linearly associated with smaller total brain, grey matter, white matter, and hippocampal volumes, with our study also supporting the proposition that heavy coffee consumption may be associated with an increased odds of dementia.

Data availability statement:

All data will be available to approved users of the UK Biobank upon application.

Acknowledgements:

Conflicts of interest statement: The authors do not have any conflicts of interest to declare.

Funding details: This work was supported by the National Health and Medical Research Council under Grants GNT1157281 and GNT1123603.

Author contributions: The author's responsibilities were as follows – KP analysed data and prepared the first draft; AM, AZ advised on analyses, analysed the data, and drafted the manuscript; EH conceptualised, funded, and supervised the study and drafted the manuscript; JTO and DL along with all other authors interpreted the results, revised the paper and approved the final manuscript

References:

1. International Coffee Organization [ICO]. Annual Review 2017-18. London: International Coffee Organization 2019.
2. Gonzalez de Mejia E, Ramirez-Mares MV. Impact of caffeine and coffee on our health. *Trends Endocrinol Metab* 2014;25(10):489-92. doi: 10.1016/j.tem.2014.07.003.
3. Nehlig A, Armspach JP, Namer IJ. SPECT assessment of brain activation induced by caffeine: no effect on areas involved in dependence. *Dialogues Clin Neurosci* 2010;12(2):255-63.
4. Spaeth AM, Goel N, Dinges DF. The Cumulative Neurobehavioral and Physiological Effects of Chronic Caffeine Intake: Individual Differences and Implications for the Use of Caffeinated Energy Products. *Nutr Rev* 2014; 72(1): 34-47. doi: 10.1111/nure.12151.
5. Geethavani G, Rameswarudu M, Reddy, RR. Effect of Caffeine on Heart Rate and Blood Pressure. *Int J Sci Res* 2014; 4(2). Retrieved from <http://www.ijsrp.org/>.
6. Rebola N, Lujan R, Cunha RA, Mulle C. Adenosine A2A receptors are essential for long-term potentiation of NMDA-EPSCs at hippocampal mossy fiber synapses. *Neuron* 2008;57(1):121-34. doi: 10.1016/j.neuron.2007.11.023.
7. Haller S, Montandon ML, Rodriguez C, Herrmann FR, Giannakopoulos P. Impact of Coffee, Wine, and Chocolate Consumption on Cognitive Outcome and MRI Parameters in Old Age. *Nutrients* 2018;10(10). doi: 10.3390/nu10101391.
8. Perlaki G, Orsi G, Kovacs N, Schwarcz A, Pap Z, Kalmar Z, Plozer E, Csatho A, Gabriel R, Komoly S, et al. Coffee consumption may influence hippocampal volume in young women. *Brain Imaging Behav* 2011;5(4):274-84. doi: 10.1007/s11682-011-9131-6.
9. Araújo LF, Mirza SS, Bos D, Niessen WJ, Barreto SM, van der Lugt A, Vernooij MW, Hofman A, Tiemeier H, Ikram MA. Association of Coffee Consumption with MRI Markers and Cognitive Function: A Population-Based Study. *J Alzheimers Dis* 2016;53(2):451-61. doi: 10.3233/jad-160116.

10. Lin Y, Weibel J, Landolt H, Santini F, Meyer M, Brunmair J, Meier-Menches, Gerner C, Borgwardt S, Cajochen C, Reichert C. Daily Caffeine Intake Induces Concentration-Dependent Medial Temporal Plasticity in Humans: A Multimodal Double-Blind Randomized Controlled Trial. *Cereb Cortex* 2021; 0(0): 1-11. doi: doi: 10.1093/cercor/bhab005.
11. Debette S, Markus HS. The clinical importance of white matter hyperintensities on brain magnetic resonance imaging: systematic review and meta-analysis. *Bmj* 2010;341:c3666. doi: 10.1136/bmj.c3666.
12. Ritchie K, Artero S, Portet F, Brickman A, Muraskin J, Beanino E, Ancelin ML, Carrière I. Caffeine, cognitive functioning, and white matter lesions in the elderly: establishing causality from epidemiological evidence. *J Alzheimers Dis* 2010;20 Suppl 1:S161-6. doi: 10.3233/jad-2010-1387.
13. Larsson SC, Traylor M, Malik R, Dichgans M, Burgess S, Markus HS. Modifiable pathways in Alzheimer's disease: Mendelian randomisation analysis. *Bmj* 2017;359:j5375. doi: 10.1136/bmj.j5375.
14. Arab L, Biggs ML, O'Meara ES, Longstreth WT, Crane PK, Fitzpatrick AL. Gender differences in tea, coffee, and cognitive decline in the elderly: the Cardiovascular Health Study. *J Alzheimers Dis* 2011;27(3):553-66. doi: 10.3233/jad-2011-110431.
15. Driscoll I, Shumaker SA, Snively BM, Margolis KL, Manson JE, Vitolins MZ, Rossom RC, Espeland MA. Relationships Between Caffeine Intake and Risk for Probable Dementia or Global Cognitive Impairment: The Women's Health Initiative Memory Study. *J Gerontol A Biol Sci Med Sci* 2016;71(12):1596-602. doi: 10.1093/gerona/glw078.
16. Eskelinen MH, Kivipelto M. Caffeine as a protective factor in dementia and Alzheimer's disease. *J Alzheimers Dis* 2010;20 Suppl 1:S167-74. doi: 10.3233/jad-2010-1404.
17. Santos C, Lunet N, Azevedo A, de Mendonça A, Ritchie K, Barros H. Caffeine intake is associated with a lower risk of cognitive decline: a cohort study from Portugal. *J Alzheimers Dis* 2010;20 Suppl 1:S175-85. doi: 10.3233/jad-2010-091303.
18. Sugiyama K, Tomata Y, Kaiho Y, Honkura K, Sugawara Y, Tsuji I. Association between Coffee Consumption and Incident Risk of Disabling Dementia in

- Elderly Japanese: The Ohsaki Cohort 2006 Study. *J Alzheimers Dis* 2016;50(2):491-500. doi: 10.3233/jad-150693.
19. Fischer K, Melo van Lent D, Wolfsgruber S, Weinhold L, Kleineidam L, Bickel H, Scherer M, Eisele M, van den Bussche H, Wiese B, et al. Prospective Associations between Single Foods, Alzheimer's Dementia and Memory Decline in the Elderly. *Nutrients* 2018;10(7). doi: 10.3390/nu10070852.
 20. Gelber RP, Petrovitch H, Masaki KH, Ross GW, White LR. Coffee intake in midlife and risk of dementia and its neuropathologic correlates. *J Alzheimers Dis* 2011;23(4):607-15. doi: 10.3233/jad-2010-101428.
 21. Laitala VS, Kaprio J, Koskenvuo M, Rähä I, Rinne JO, Silventoinen K. Coffee drinking in middle age is not associated with cognitive performance in old age. *Am J Clin Nutr* 2009;90(3):640-6. doi: 10.3945/ajcn.2009.27660.
 22. 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). doi: 10.3390/nu10101501.
 23. Mirza SS, Tiemeier H, de Bruijn RF, Hofman A, Franco OH, Kiefte-de Jong J, Koudstaal PJ, Ikram MA. Coffee consumption and incident dementia. *Eur J Epidemiol* 2014;29(10):735-41. doi: 10.1007/s10654-014-9943-y.
 24. Liu QP, Wu YF, Cheng HY, Xia T, Ding H, Wang H, Wang ZM, Xu Y. 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. doi: 10.1016/j.nut.2015.11.015.
 25. Larsson SC, Virtamo J, Wolk A. Coffee consumption and risk of stroke in women. *Stroke* 2011;42(4):908-12. doi: 10.1161/strokeaha.110.603787.
 26. Lee J, Lee J-E, Kim Y. Relationship between coffee consumption and stroke risk in Korean population: the Health Examinees (HEXA) Study. *Nutrition Journal* 2017;16(1):7. doi: 10.1186/s12937-017-0232-y.
 27. Liebeskind DS, Sanossian N, Fu KA, Wang HJ, Arab L. The coffee paradox in stroke: Increased consumption linked with fewer strokes. *Nutr Neurosci* 2016;19(9):406-13. doi: 10.1179/1476830515y.0000000035.
 28. Lopez-Garcia E, Rodriguez-Artalejo F, Rexrode KM, Logroscino G, Hu FB, van Dam RM. Coffee consumption and risk of stroke in women. *Circulation* 2009;119(8):1116-23. doi: 10.1161/circulationaha.108.826164.

29. de Koning Gans JM, Uiterwaal Cuno SPM, van der Schouw Yvonne T, Boer Jolanda MA, Grobbee Diederick E, Verschuren WMM, Beulens Joline WJ. Tea and Coffee Consumption and Cardiovascular Morbidity and Mortality. *Arteriosclerosis, Thrombosis, and Vascular Biology* 2010;30(8):1665-71. doi: 10.1161/ATVBAHA.109.201939.
30. Floegel A, Pischon T, Bergmann MM, Teucher B, Kaaks R, Boeing H. Coffee consumption and risk of chronic disease in the European Prospective Investigation into Cancer and Nutrition (EPIC)-Germany study. *Am J Clin Nutr* 2012;95(4):901-8. doi: 10.3945/ajcn.111.023648.
31. Kokubo Y, Iso H, Saito I, Yamagishi K, Yatsuya H, Ishihara J, Inoue M, Tsugane S. The impact of green tea and coffee consumption on the reduced risk of stroke incidence in Japanese population: the Japan public health center-based study cohort. *Stroke* 2013;44(5):1369-74. doi: 10.1161/strokeaha.111.677500.
32. Nordestgaard AT, Nordestgaard BG. Coffee intake, cardiovascular disease and all-cause mortality: observational and Mendelian randomization analyses in 95 000-223 000 individuals. *Int J Epidemiol* 2016;45(6):1938-52. doi: 10.1093/ije/dyw325.
33. Deng C, Lu Q, Gong B, Li L, Chang L, Fu L, Zhao Y. Stroke and food groups: an overview of systematic reviews and meta-analyses. *Public Health Nutr* 2018;21(4):766-76. doi: 10.1017/s1368980017003093.
34. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, 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. doi: 10.1371/journal.pmed.1001779.
35. Miller KL, Alfaro-Almagro F, Bangerter NK, Thomas DL, Yacoub E, Xu J, Bartsch AJ, Jbabdi S, Sotiropoulos SN, Andersson JL, et al. Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nat Neurosci* 2016;19(11):1523-36. doi: 10.1038/nn.4393.
36. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust relationship inference in genome-wide association studies. *Bioinformatics* 2010;26(22):2867-73. doi: 10.1093/bioinformatics/btq559.
37. UK Biobank. UK Biobank Ethics and Governance Framework. Stockport: UK Biobank, 2007.

38. UK Biobank. Internet: <http://biobank.ndph.ox.ac.uk/showcase/> (accessed 26 August 2019).
39. UK Biobank. UK Biobank Brain Imaging - Acquisition Protocol Stockport: UK Biobank, 2014.
40. Smith SM, Alfaro-Almagro F, Miller KL. UK Biobank Brain Imaging Documentation Stockport Welcome Centre for Integrative Neuroimaging (WIN-FMRIB), 2020.
41. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. FSL. *Neuroimage* 2012;62(2):782-90. doi: 10.1016/j.neuroimage.2011.09.015.
42. UK Biobank. First Occurrence of Health Outcomes Defined by 3-character ICD10 code. Stockport UK Biobank, 2019.
43. Bush K, Wilkinson T, Schnier C, Nolan J, Sudlow C. Definitions of Dementia and the Major Diagnostic Pathologies, UK Biobank Phase 1 Outcomes Adjudication. Stockport: UK Biobank, 2018.
44. Schnier C, Bush K, Nolan J, Sudlow C. Definitions of Stroke for UK Biobank Phase 1 Outcomes Adjudication 1.1 ed. Stockport: UK Biobank, 2017.
45. World Health Organization [WHO]. Internet: <http://www.euro.who.int/en/health-topics/disease-prevention/nutrition/a-healthy-lifestyle/body-mass-index-bmi> ((accessed 3 July 2019).
46. UK Biobank. Blood Pressure, Version 1.0. Stockport: UK Biobank, 2011.
47. UK Biobank. Body Composition Measurement, Version 1.0. Stockport: UK Biobank, 2011.
48. UK Biobank. Biomarker assay quality procedures: approaches used to minimise systematic and random errors (and the wider epidemiological implications), Version 1.2. Stockport: UK Biobank, 2019.
49. 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.
50. Hyppönen E, Zhou A. Cardiovascular symptoms affect the patterns of habitual coffee consumption. *Am J Clin Nutr*: 2021; 00:1-6.
<https://doi.org/10.1093/ajcn/nqab014>.
51. Sterne JA, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ*. 2009;338:b2393. Published 2009 Jun 29. doi:10.1136/bmj.b239347.

52. StataCorp LLC. STATA IC 14.1. Texas: StataCorp, 2015.
53. McLellan TM, Caldwell JA, Lieberman HR. A review of caffeine's effects on cognitive, physical and occupational performance. *Neurosci Biobehav Rev.* 2016;71:294-312. doi:10.1016/j.neubiorev.2016.09.001
54. West RK, Ravona-Springer R, Livny A, Heymann A, Shahar D, Leroith D, Preiss R, Zukran R, Silverman JM, Schnaider-Beeri M. Age Modulates the Association of Caffeine Intake With Cognition and With Gray Matter in Elderly Diabetics. *J Gerontol A Biol Sci Med Sci* 2019;74(5):683-8. doi: 10.1093/gerona/gly090.
55. Rivera-Oliver M, Díaz-Ríos M. Using caffeine and other adenosine receptor antagonists and agonists as therapeutic tools against neurodegenerative diseases: a review. *Life Sci.* 2014;101(1-2):1-9. doi:10.1016/j.lfs.2014.01.083
56. You D, Kim Y, Ha A, Lee Y, Kim S, Kim C, Lee S, Choi D, Lee J. Possible Health Effects of Caffeinated Coffee Consumption on Alzheimer's Disease and Cardiovascular Disease. *Toxicol Res;* 27(1): 7-10. doi: 10.5487/TR.2011.27.1.007.
57. Cai L, Ma D, Zhang Y, Liu Z, Wang P. The effect of coffee consumption on serum lipids: a meta-analysis of randomized controlled trials. *Eur J Clin Nutr.* 2012;66(8):872-877. doi:10.1038/ejcn.2012.68
58. Iadecola, C. Revisiting atherosclerosis and dementia. *Nat Neurosci* 23, 691–692 (2020). <https://doi.org/10.1038/s41593-020-0626-6>
59. Alves, J., Petrosyan, A., & Magalhães, R. (2014). Olfactory dysfunction in dementia. *World journal of clinical cases*, 2(11), 661–667. <https://doi.org/10.12998/wjcc.v2.i11.661>
60. Sakai, M., Kazui, H., Shigenobu, K., Komori, K., Ikeda, M., & Nishikawa, T. (2017). Gustatory Dysfunction as an Early Symptom of Semantic Dementia. *Dementia and geriatric cognitive disorders extra*, 7(3), 395–405. <https://doi.org/10.1159/000481854>
61. Fry A, Littlejohns TJ, Sudlow C, Doherty N, Adamska L, Sprosen T, Collins R, Allen NE. 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. doi: 10.1093/aje/kwx246.

62. Poole R, Kennedy OJ, Roderick P, Fallowfield JA, Hayes PC, Parkes J. Coffee consumption and health: umbrella review of meta-analyses of multiple health outcomes. *Bmj* 2017;359:j5024. doi: 10.1136/bmj.j5024.

Table 1. Summary of coffee consumption, total brain volume and incidence of dementia and stroke, by baseline characteristics of participants in the UK Biobank.

	n (%)	Coffee drinker (%)	>6 cups/day coffee drinkers (%)	TBV [^] Median (IQR)	Incident dementia Rate* (n)	Incident stroke Rate* (n)	
Sex							
Male	182,370 (45.7)	68.4	2.9	1488.8 (1441.0, 1536.0)	8.8 (2,447)	15.0 (3,595)	
Female	216,276 (54.3)	59.3	1.8	1514.8 (1464.4, 1564.7)	5.7 (1,886)	8.9 (2,586)	
	P		<1 x 10 ⁻³⁰⁰		3.5 x 10 ⁻¹³⁰	3.0 x 10 ⁻³²	1.4 x 10 ⁻⁷⁴
Age							
39 – 49 years	87,573 (22.0)	61.9	2.9	1555.1 (1512.7, 1595.6)	1.1 (158)	3.7 (493)	
50 – 59 years	131,944 (33.1)	63.3	2.5	1505.2 (1463.6, 1548.8)	3.3 (652)	7.7 (1,459)	
60 – 73 years	179,129 (44.9)	64.4	1.9	1457.1 (1416.3, 1500.4)	13.8 (3,523)	17.7 (4,229)	
	P		1.2 x 10 ⁻³⁰		<1 x 10 ⁻³⁰⁰	1.31 x 10 ⁻²⁷¹	1.2 x 10 ⁻²⁵⁷
BMI							
Underweight, <18.5 kg/m ²	1,987 (0.5)	58.8	2.7	1525.8 (1468.2, 1556.2)	9.5 (31)	13.6 (37)	
Normal, [18.5, 25) kg/m ²	129,989 (32.6)	63.0	1.9	1511.0 (1458.4, 1561.6)	6.6 (1,211)	9.5 (1,644)	
Overweight, [25, 30) kg/m ²	169,951 (42.6)	64.4	2.3	1496.3 (1448.3, 1547.0)	7.3 (1,766)	11.7 (2,609)	
Obese, ≥30 kg/m ²	95,534 (24.0)	62.5	2.8	1497.6 (1448.5, 1543.0)	8.8 (1,232)	14.4 (1,826)	
Missing	1,185 (0.3)	58.0	2.4	1491.2 (1429.8, 1554.1)	28.3 (91)	29.1 (65)	
	P		9.0 x 10 ⁻⁹		5.2 x 10 ⁻¹¹	0.010	8.4 x 10 ⁻¹¹
Education							
None	69,128 (17.4)	58.1	2.9	1482.1 (1434.7, 1526.3)	15.0 (1,539)	18.4 (1,672)	
Intermediate (NVQ/CSE/A-levels)	142,597 (35.8)	62.2	2.5	1508.3 (1458.1, 1558.0)	6.4 (1,278)	10.8 (2,026)	
High (degree/professional)	183,641 (46.0)	66.6	2.0	1501.3 (1451.3, 1551.2)	5.3 (1,395)	9.8 (2,723)	
Missing	3,280 (0.8)	57.5	4.4	1485.5 (1442.1, 1541.5)	19.0 (121)	13.3 (102)	
	P		<1 x 10 ⁻³⁰⁰		0.004	5.8 x 10 ⁻⁴⁵	5.8 x 10 ⁻¹⁹
Stressful Events (last 2 years)							
None	224,381 (56.3)	64.2	2.0	1498.1 (1448.1, 1548.3)	7.3 (2,258)	11.4 (3,361)	
Serious illness or injury	35,024 (8.8)	61.0	2.9	1501.2 (1447.0, 1553.9)	13.8 (752)	18.0 (865)	
Family death or illness	106,567 (26.7)	62.9	2.2	1506.3 (1456.6, 1556.2)	5.9 (910)	10.2 (1,458)	
Marital separation or divorce	7,434 (1.9)	63.6	3.4	1520.6 (1474.7, 1564.9)	5.7 (76)	8.6 (94)	

Financial difficulties	23,313 (5.8)	63.1	3.9	1518.9 (1468.7, 1564.8)	6.6 (250)	12.1 (398)	
Missing	1,927 (0.5)	61.5	2.0	1495.8 (1458.5, 1546.5)	17.6 (87)	15.4 (57)	
	P		6.7×10^{-21}		0.10	1.3×10^{-10}	1.3×10^{-8}
History of depression							
No	376,132 (94.3)	63.7	2.2	1502.0 (1451.6, 1551.8)	7.2 (3,899)	11.6 (5,800)	
Yes	22,514 (5.7)	59.3	3.4	1504.1 (1457.5, 1556.2)	12.8 (434)	12.8 (381)	
	P		1.1×10^{-18}		0.005	3.0×10^{-37}	1.9×10^{-5}
Long standing illness							
No	265,333 (66.6)	64.9	2.1	1505.2 (1454.7, 1555.1)	4.5 (1,619)	8.9 (3,088)	
Yes	124,426 (31.2)	60.5	2.7	1491.4 (1443.8, 1542.4)	13.4 (2,518)	17.5 (2,902)	
Missing	8,887 (2.2)	62.6	2.6	1500.8 (1445.8, 1557.4)	13.7 (196)	14.8 (191)	
	P		7.3×10^{-204}		1.6×10^{-6}	3.3×10^{-131}	4.5×10^{-71}

P values were from log likelihood ratio tests using logistic regression (for coffee drinking), linear regression (for brain volume) or logistic regression (for disease risk), adjusting for age, sex, and assessment centre, and weighted for 1-kinship coefficient to account for relatedness

^ Total brain volume, information shown for the imaging subsample (n = 17,702)

* rate per 10,000 person- years.

Table 2. Differences in brain volume indicators (in cm³) by habitual caffeinated coffee consumption (cups per day), using linear regression models.

		Beta (95% CI)	P
Total brain volumes (n = 17,702)	Adjusted for basic covariates	-1.865 (-2.310, -1.421)	2.2 x 10 ⁻¹⁶
	Adjusted for all covariates	-1.416 (-1.893, -0.938)	7.4 x 10 ⁻⁹
Grey matter volumes (n = 17,702)	Adjusted for basic covariates	-1.294 (-1.565, -1.024)	7.4 x 10 ⁻²¹
	Adjusted for all covariates	-0.910 (-1.196, -0.624)	7.4 x 10 ⁻¹⁰
White matter volumes (n = 17,702)	Adjusted for basic covariates	-0.571 (-0.867, -0.275)	1.5 x 10 ⁻⁴
	Adjusted for all covariates	-0.506 (-0.825, -0.186)	0.0020
Hippocampal volumes (n = 17,689)	Adjusted for basic covariates	-0.015 (-0.022, -0.007)	8.2 x 10 ⁻⁵
	Adjusted for all covariates	-0.011 (-0.018, -0.003)	0.0092
Left hippocampal volumes (n = 17,689)	Adjusted for basic covariates	-0.007 (-0.011, -0.003)	0.0016
	Adjusted for all covariates	-0.005 (-0.009, -0.0004)	0.031
Right hippocampal volumes (n = 17,689)	Adjusted for basic covariates	-0.008 (-0.012, -0.004)	1.2 x 10 ⁻⁴
	Adjusted for all covariates	-0.006 (-0.010, -0.001)	0.015
White matter hyperintensity volumes (n = 16,730)	Adjusted for basic covariates	0.040 (-0.014, 0.093)	0.15
	Adjusted for all covariates	-0.011 (-0.068, 0.046)	0.71

Basic covariates: age, age square, sex, and assessment centre

Socioeconomic status: basic covariates + education, Townsend deprivation index, income, and employment status

Diseases status: basic covariates + depression, systolic blood pressure, sleep duration and insomnia, and longstanding illness

Anthropometric measures: basic covariates + height, body mass index, and body mass index square

Dehydration indicator: basic covariates + water intake, whole body water mass, and urine-creatinine ratio

Lifestyle factors: basic covariates + smoking, alcohol, type of physical activity, tea consumption, and stress-related lifestyle factors, processed meat, fresh fruit

All covariates: adjusted for all covariates listed above.