Official title: Coffee Intake Biomarker Validation and Associations with Respiratory Diseases

Approved Date: August 22, 2022

Study Protocol and Statistical Analysis Plan

Recruitment of study participants

Study participants (n=100) will be recruited via flyers (e.g., shopping malls, MRT stations, neighbourhood shops, community centres, food centres, tertiary institutions such as NUS (Saw Swee Hock School of Public Health, National University of Singapore), and door-to-door), advertisements at on-campus notice boards, bus-stops and newspapers. Recruitment email within NUS and NTU (Lee Kong Chian School of Medicine, Food Science & Technology and Office of Faculty Housing, Nanyang Technological University) will be sent by the respective administrative officers on behalf of the project team members. Recruitment will also be conducted through online platforms such as Telegram, Facebook and Instagram, as well as snowballing and convenience sampling through referrals. Participants will be screened for their eligibility before they are enrolled into the study. Exclusion criteria would include pregnancy, medication use, presence or history of diseases (e.g., cancer, cardiometabolic, endocrine), major surgical operations, depression, anxiety, caffeine/lactose intolerance, fear of needles and/or blood disorders. A short questionnaire will be administered to each participant to obtain information on their basic demographics, lifestyle and medical history. Inclusion criteria would include regular coffee drinkers aged 21-64 years old and of Chinese ethnicity. All participants will be compensated for their time (S$300). Written informed consent will be obtained from each participant prior to the commencement of the study. Approvals from the Institutional Review Boards of NUS and NTU will also be obtained.

Briefing session and informed consent (two weeks prior to intervention study)

A briefing session, along with informed consent, will take place two weeks before commencing the intervention study at NTU Experimental Medicine Building level 1 (interview room) or through an arranged Zoom call. Participants who opted for an arranged zoom call will be sent a scheduled meeting link, where he/she will be briefed by the recruited research assistant and/or PI/ Co-I. Researchers will explain the project aims and intervention study, and answer questions that the participants may have. If

the participants agree, we will record their informed consent. Over zoom, informed consent will be conducted via a video zoom call, in presence of a third-party witness (either an individual added in the call or physically present with the participant at the time of the call) and the participant will be sign the consent with an e-signature. This will also be performed by the research assistant and should be completed within half an hour. Participants will be briefed on how to collect their 24-hour urine (during Day 0- the day before the intervention or otherwise known as Day 1) and provided with urine collection containers and esky bags. They will also be given or emailed handouts on foods/drinks not allowed to be consumed one week before the intervention study and during the study.

Multi-dose intervention study design

The accurate measurement of all coffee-derived compounds and metabolites in urine and blood will be carried out in a pilot intervention study, where coffee (Coffee A) will be fed to participants (n=100) in an increasing dose over three weeks (one dose per week, two dose per week and three dose per week).

Their next coffee intervention will be Coffee B. Coffee B will be the highest dosage (three dose per week) of a different blend of coffee.

Pre-study Commencement:

Before commencing the study (Day 0), participants will be requested to collect their total volume urine (24h urine) in a ‘urine care package’ consisting of urine collection containers, zip lock bags, ice packs, an esky and an instruction sheet that would have been delivered to them prior. The participants will also be requested to fast overnight (8-12h) without caloric intake.

Study Visits and Blood Collection:

On Day 1 of the intervention study, participants will be split into small groups, consisting of at most 50 participants each visit over at least two days. This will take

place at the NTU Experimental Medicine Building. Upon arrival, participants will be required to go to the seminar room to provide us with their filled urine containers and complete questionnaires. They will then proceed to level 1 (team-based learning room) for blood drawing. This is a designated, safe, clean, quiet/comfortable room. Qualified personnel/registered nurses from Caring Hearts Nursing Pte Ltd have been engaged to provide mobile phlebotomy service through the duration of the intervention study and will perform the blood-taking procedure.

A total of 11 blood collections will be performed. Subjects will be divided into two groups (Group A and Group B) for logistics reason. For Group A, blood will be collected into 6 mL vacutainers on Days 1, 8, 15,22, 29, 36, 43, 50, 57, 60, 63 of the intervention study (over 10 consecutive weeks). For Group B, blood will be collected into 6 mL vacutainers on Days 2, 9, 16, 23, 30, 37, 44, 51, 58, 61, 64 of the intervention study (over 10 consecutive weeks).

Coffee Intervention:

Subjects will then proceed to a designated area outside the seminar room for coffee drinking and to receive coffee grounds for the remaining days of the week. They will receive pre-packaged coffee grounds to prepare the low coffee dose at home for the remainder of the week. Ground coffee will be purchased from a local coffee provider and packed into individual coffee sachets. The coffee grounds will be packed by an external food packing service, in a food-grade certified premise. The letters/ documentation indicating Hazard Analysis Critical Control Point (HACCP) certification of the vendor's premise has been provided.

The starting dose of coffee A per dose is ≤10 g coffee powder for low dose, ≤20 g coffee powder for medium dose and ≤40 g coffee powder for high dose. On the other hand, coffee B will be of high dosage (≤40g). Participants will be asked to brew the coffee for a maximum of 10 minutes in a cup of hot water, minimum 150 mL. The concept of this intervention study design is to stimulate habitual drinking of coffee. Hence, participants are required to consume the coffee doses daily but are not restricted in the time of coffee consumption, they may take the coffee in the morning or afternoon.

At the end of the low coffee dose intervention (Day 7), participants will be requested to collect their 24h urine and return on Day 8, where fasting serum samples will be collected from them. This is followed by a week without any coffee consumption to investigate the post consumption kinetics, i.e., how long the candidate biomarkers remain present in the biofluids after stopping coffee consumption. At the end of the 'coffee abstinence' week (Day 14), again, participants will be requested to collect their 24h urine and fast overnight for the serum sampling on Day 15. On the same day (Day 15), participants will be provided with the medium dose coffee and the next set of pre-packaged coffee grounds to prepare the medium coffee dose at home for the remainder of the week. This is repeated for the high dose coffee A and high dose coffee B interventions. Each study visit is estimated to be completed within a maximum of two hours.

Tea Intervention:

A single dose will be provided to participants, in the form of one tea-bag from a generic brand. Pre-packaged commercial tea bags will be purchased from a supermarket. Participants will be asked to brew the tea-bag for a maximum of 5 minutes in a cup of hot water (200 mL). From two days before commencing the study until the end of the intervention study, participants will be requested to avoid certain foods that may confound the study design including other coffee and tea that are not provided by us, coffee or tea-flavoured foods, caffeine-containing beverages (i.e., Coca-Cola, energy drinks), supplements containing caffeine or of herbal/botanical origin, chocolate and chocolate-containing foods. We will organise a briefing with the participants before commencing the study, and provide a pamphlet containing pictures and a detailed list of "allowed" and "non-allowed" foods/beverages. We will also ask the participants to keep an electronic daily dietary record and take photos of their meals to ensure that they maintain compliance with the study design. Participants who consume coffee products outside of our study design will be flagged and excluded in the sensitivity analysis. With a sample size of 100 participants, we will stagger their visit for the serum collection/handing out of coffee grounds into morning sessions for 3 days, i.e., 33-34 participants per day. The session (08.00-11.00) will consist of serum sampling, briefing, consent taking, etc from 08.00-11.00 and processing of serum/urine samples in the

laboratory from 09.00-13.00.

Sample processing and extraction

Urine will be collected (weekly, 20mL) in urine specimen containers, transferred to 50mL tubes, centrifuged (1500g, 10 min, 4°C) and aliquoted into 1.5mL tubes. Blood will be collected (weekly, 6mL) in Vacutainer® serum tubes and allowed to clot/stand at room temperature. After 30-40 min, serum will be separated from entire blood by centrifugation (2000g, 10 min, 24°C) and aliquoted into 1.5mL tubes. All samples will be processed on the same day that they are collected. Coffee and tea samples provided to the participants will also be collected (weekly, 50mL each) with a portion aliquoted into 1.5mL tubes for a detailed chemical analysis as the composition of local coffee brews has not been studied. Samples will be stored at -80°C until all samples have been collected at the end of the intervention study and processed in a single batch. Briefly, thawed samples will be added with deuterated internal standards (for measure of control for extraction). After, urine samples will be centrifuged (4000g, 10 min, 4°C) and transferred into glass vials. Serum samples will be de-proteinised with acidified methanol, centrifuged (4000g, 10 min, 4°C) and transferred into glass vials. Coffee and tea aliquots will be centrifuged (4000g, 10 min, 4°C) to separate the particulate pellet from aqueous phase and fat layer on the surface, if present. The fat layer will be removed with a micro-spatula, diluted 50-fold and transferred into glass vials. All samples will be stored at -80°C until ready for analysis.

Targeted and untargeted metabolomic profiling

Urine and serum samples will undergo a combination of targeted and untargeted metabolomic profiling at the Singapore Phenome Centre located at NTU, which has state-of-the-art mass spectrometers and strong technical support. Firstly, targeted detection and quantification of a list of previously proposed coffee biomarkers will be conducted on the intervention urine and serum samples using a Waters Acquity Ultra-High Performance Liquid Chromatography-Triple Quadrupole Mass Spectrometry (UHPLC-TQ-MS), and commercially available standards and deuterated standards. For atractyligenin, its lack of standard previously limited its detection but we have

successfully synthesised this compound from its atractyloside precursor, which can be used for level I identification and accurate quantification in this project. Briefly, thawed urine, serum, coffee and tea samples will be analysed directly on the UHPLC-TQ-MS. Secondly, to explore potentially novel biomarkers and/or metabolites related to the consumption of the two types of coffee relevant to Singapore (black coffee and coffee with additives), untargeted metabolomics will be conducted on the urine and serum samples using a Waters Acquity UHPLC-QTOF-MS, operating in positive and negative electrospray ionisation modes with a mass scan range from 50 to 1000 m/z. Samples will be processed as described above. For each matrix, a quality control extract of pooled urine/serum from all study participants will be injected at the beginning and end of each run, and in-between every 10 samples.

Data Processing and Compound Identification

Quantification of the biomarker compounds from the targeted method will be processed using MassLynx (Waters-specific software). Processing of the untargeted metabolomics data will be carried out using an open access platform, Galaxy WorkFlow4Metabolomics (http://workflow4metabolomics.org/). Only de-identified datasets without subject identifiers will be uploaded to the open access platform. In untargeted metabolomics, the identity of detected ions is a priori unknown and the multi-step identification process will follow as previously described using publicly available databases, e.g., Metlin, PhytoHub, HMDB, m/zCloud, etc. As identification of unknown ions is still very much a bottleneck in metabolomics studies, an in-house database of all known coffee precursors and metabolites will also be created to facilitate this identification process, and will be built with monoisotopic mass, chemical formula, m/z, and fragmentation pattern and spectra obtained from commercially available and deuterated standards.

Statistical analysis

The concentrations of compounds identified and quantified from the targeted method will be used to build calibration curves relative to their dose-response in Microsoft Excel 2016 and R version 3.4. Assuming a drop-out rate of 33% (one-third), we will have a

total of 67 participants in for this pilot study. Cross-validated principal component analysis and multivariate linear regression models will be used to assess the compounds that best discriminate the multi-dose coffee consumption in the untargeted approach (e.g., having a variable importance in projection (VIP) >2) using SIMCA version 15, and prioritised for identification.