Official Title of the study:
Serum Uric Acid Levels and Initial Presentation of Cardiovascular Diseases: a CALIBER Study

NCT number:
Not yet assigned

Date of the document:
23 January 2018
INTRODUCTION
Serum uric acid (SUA) is the final metabolic product of purine, which is generated in the liver and intestine by xanthine oxidase and is excreted from the urine [1]. SUA level has been observed to be associated with the risk of developing cardiovascular diseases (CVDs) in some studies [2-4], while other studies showed controversial results [5-7]. Although currently remains unclear, several pathologic mechanisms have been promoted to explain the effect of SUA on CVDs, including deleterious effects on endothelial function, oxidative metabolism, platelet adhesiveness, hemorheology, and aggregation [8-14].

An umbrella review published in June 2017 on BMJ [7] summarized evidence from clinical studies investigating the relationship between SUA levels and CVDs. None of them estimated the lifetime incidence associated with SUA attributable to the initial presentation of specific CVDs. Although one previous study investigated the association between SUA and the initial presentations of coronary heart disease, the outcome was composite [5]. Besides, previous studies have not estimated the association between SUA and eight endpoints, namely stable angina, unstable angina, subarachnoid haemorrhage, intracerebral haemorrhage, peripheral arterial disease, abdominal aortic aneurysm, and cardiac arrest/sudden cardiac death, and total cardiovascular disease.

We will address these gaps in one large prospective British cohort. Our objectives are to estimate the association between SUA and future risk of 13 of the most common initial cardiovascular presentations.

METHODS
Study overview
We will first conduct an observational study in the CALIBER (Clinical disease research using LInked Bespoke studies and Electronic health Records) programme cohort [15] to estimate the association between SUA and the incidences of the initial presentation of 13 CVDs.

Study population
We will select anonymised patients from the CALIBER cohort [15], which links four sources of electronic health data in England: the Clinical Practice Research Datalink (CPRD) with primary healthcare records, the Hospital Episode Statistics (HES) with hospital discharge information, and the Office for National Statistics (ONS) mortality and social deprivation data.

A patient will be included if he or she: a) was registered between January 1, 1998, and September 2015; b) aged 30 years or older at study entry; c) had no record of
previous diagnosis of cardiovascular disease; d) had been followed up for at least one year before the study entry; e) had at least one uric acid measurement.

**Exposure**
The main exposure will be the baseline serum uric acid (SUA) level measured within one year of the study entry as recorded in primary care. If a patient had more than one measurement, the mean of the values will be taken. We will perform sensitivity analyses to use the SUA level recorded closest to the study entry.

Since gout is an important complication of hyperuricaemia and is associated with CVDs, we will also classify the patient state at baseline as with gout or with no gout, regardless of SUA levels, based on the diagnosis within one year of the study entry. Participants who developed new-onset gout during follow-up will be analysed according to their baseline status of no gout.

**Covariates**
The following baseline covariates will be considered:

a) Sociodemographic: sex, age and index of multiple deprivation;

b) Cardiovascular risk factors: body mass index, systolic blood pressure, high-density lipoprotein (HDL) cholesterol, low-density lipoprotein (LDL) cholesterol, triglyceride, and estimated glomerular filtration rate, smoking status, alcohol consumption;

c) Medication use: any prescription in the year before index date of blood pressure lowering drugs (including diuretics, beta blockers, ACEIs, ARBs), statins, non-steroidal anti-inflammatory drugs (NSAIDs), aspirin, immunosuppressant drug, and in women only, oestrogen oral contraceptives and hormone replacement therapy;

d) Comorbidities (defined by Read codes, a coded thesaurus of clinical terms used in the UK, in primary care data and by ICD-10 codes in hospital care data): Cancer, diabetes, asthma or atopy, COPD, connective tissue disease, inflammatory bowel diseases.

For continuous variables (BMI, systolic blood pressure, and cholesterol), we will use the most recent value in the year before the study entry. For alcohol consumption, we will use the most recent record in the five years before entry to classify participants’ drinking behaviour into five categories: non-drinkers (Read codes as “teetotaller” or “non-drinker”), former drinkers (Read codes as “stopped drinking alcohol” or “ex-drinker”), occasional drinkers (Read codes as “drinks rarely” or “drinks occasionally”), current moderate drinkers (Read codes as “alcohol intake within recommended sensible limits” or “light drinker”), and heavy drinkers (Read codes as “alcohol intake above recommended sensible drinking limits” or “hazardous alcohol use”).
We will impute missing covariate values using multiple imputation, as implemented in the “mice” algorithm in the statistical package R. We will extract the following values of the continuous variables to use as auxiliary variables for multiple imputation: the first measurement after the time window and the most recent measurement before the time window, along with the timing of these measurements relative to the study entry. We will check whether the imputations are plausible by comparing plots of the distribution of recorded and imputed values of all variables.

**Follow up**
The observation period for all participants will begin on the date when all the study inclusion criteria are fulfilled. Participants will be followed up until the occurrence of an initial presentation of cardiovascular disease, death, or de-registration from the practice, whichever occurred first.

**Endpoints**
Endpoints will be the initial presentation of cardiovascular disease as any of the following 13 CVDs diagnosed in primary care (defined by Read codes), secondary care (defined by ICD-10 codes), or at death (defined by ICD-10 codes): Stable angina, unstable angina, myocardial infarction, unheralded coronary heart disease death, heart failure, cardiac arrest/sudden cardiac death, transient ischaemic attack, ischaemic stroke, subarachnoid haemorrhage, intracerebral haemorrhage, peripheral arterial disease, abdominal aortic aneurysm, atrial fibrillation, and a composite total cardiovascular disease (all 13 cardiovascular diseases combined). An example is given by the definition of atrial fibrillation [16]

**Statistical analysis**
We will analyse the SUA level as a categorical variable in order to avoid presuming a particular shape for the association with CVDs. Since the normal range for SUA and the diagnosis cut-off for hyperuricemia diagnosis varies [17-19], we will use the following cut-offs:

1) 3 mg/dL (180 umol/L, the lower limit of target SUA level for long-term management recommended by the 2016 updated EULAR guideline [20]),
2) 4 mg/dL (240 umol/L),
3) 5 mg/dL (300 umol/L, the target SUA for severe cases or patients with tophi recommended by various guidelines [20-23]),
4) 6 mg/dL (360 umol/L, the general target SUA for long-term management recommended by various guidelines [20-23]),
5) 7 mg/dL (420 umol/L).

If the shape of the association between SUA and CVDs is found to be linear, we will then perform analyses with SUA level as a continuous variable.

Quantitative variables will be summarized in mean and standard deviation and qualitative variables in number and percentages. Between group differences of
parametric values will be compared using analysis of variance (ANOVA) and that of categorical variables using chi-square tests.

We will plot crude cumulative incidence curves for each specific cardiovascular endpoint for people with SUA levels between 240-300 umol vs. 360-420 umol/L, using age as the timescale, under a competing risks framework (i.e., participants could experience only one initial presentation).

We will use multivariable Cox regression to calculate cause-specific hazards for associations between categories of SUA levels and initial presentations of cardiovascular disease. We will plot Schoenfeld residuals to verify the proportional hazards assumption.

In the primary analysis, we will adjust for age (linear and quadratic), sex, index of multiple deprivation, diabetes, smoking status, BMI, systolic blood pressure, HDL cholesterol, LDL cholesterol, and prescription of statins or antihypertensive medication in the year before study entry. The baseline hazard function of each model will be stratified by general practice and sex, and we will use multiple imputation to account for missing covariate data.

Assuming independence in effects, heterogeneity in associations across CVD endpoints for the fourth vs. the first quintile will be assessed with tau-square statistic, which is the between-CVD endpoint variance, and $I^2$ that can be interpreted as the proportion of the total variation in estimates that is due to heterogeneity.

We will conduct a series of sensitivity analyses:

a) Covariate adjustment: We will conduct analyses adjusted for age and sex only, and analyses adjusted for age, sex, and cardiovascular risk factors. We will assess interactions with age and sex.

b) Gout: We will categories patients into with gout and with no gout, regardless of SUA.

c) Imputed data: We will check whether the imputations are plausible by comparing plots of the distribution of recorded and imputed values of all variables.

d) Data source: We will conduct analysis ignoring endpoints recorded only in primary care (CPRD).

e) Endpoints: We will conduct analysis restricting endpoints to fatal endpoints.

f) Time: We will conduct analysis restricting participants to individuals who entered the study after 2004 (when recording of covariates would be expected to be more complete after introduction of the Quality and Outcomes Framework to incentivise good performance in the care of specified diseases).

We will conduct all analyses using R 3.4.1.
Role of funding
The funders had no role in the study design, data collection, data analysis, data interpretation, or writing of the report.
References


