THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Genetically determined serum urate levels and cardiovascular and other diseases in UK Biobank cohort

OPEN ACCESS

Genetically determined serum urate levels and cardiovascular and other diseases in UK Biobank cohort: A phenome-wide mendelian randomization study

Xue Li[1], Xiangrui Meng[1], Yazhou He[1], Athina Spiliopoulou[2], Maria Timofeeva[3], Wei-Qi Wei[4], Aliya Gifford[4], Tian Yang[1], Tim Varley[5], Ioanna Tzoulaki[6,7], Peter Joshi[1], Joshua C. Denny[4], Paul Mckeigue[2], Harry Campbell[1], Evropi Theodoratou[1,8]*

1 Centre for Global Health, Usher Institute, University of Edinburgh, Edinburgh, United Kingdom, 2 Centre for Population Health Sciences, Usher Institute, University of Edinburgh, Edinburgh, United Kingdom, 3 Colon Cancer Genetics Group, Medical Research Council Human Genetics Unit, Medical Research Council, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, United Kingdom, 4 Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, Tennessee, United States of America, 5 Public Health and Intelligence, NHS National Services Scotland, Edinburgh, United Kingdom, 6 Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, United Kingdom, 7 Department of Hygiene and Epidemiology, University of Ioannina Medical School, Ioannina, Greece, 8 Edinburgh Cancer Research Centre, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, United Kingdom

* e.theodoratou@ed.ac.uk

Abstract

Background

The role of urate in cardiovascular diseases (CVDs) has been extensively investigated in observational studies; however, the extent of any causal effect remains unclear, making it difficult to evaluate its clinical relevance.

Methods and findings

A phenome-wide association study (PheWAS) together with a Bayesian analysis of tree-structured phenotypic model (TreeWAS) was performed to examine disease outcomes related to genetically determined serum urate levels in 339,256 unrelated White British individuals (54% female) in the UK Biobank who were aged 40–69 years (mean age, 56.87; SD, 7.99) when recruited from 2006 to 2010. Mendelian randomization (MR) analyses were performed to replicate significant findings using various genome-wide association study (GWAS) consortia data. Sensitivity analyses were conducted to examine possible pleiotropic effects on metabolic traits of the genetic variants used as instruments for urate. PheWAS analysis, examining the association with 1,431 disease outcomes, identified 13 distinct phecodes representing 4 disease groups (inflammatory polyarthropathies, hypertensive disease, circulatory disease, and metabolic disorders) and 9 disease outcomes (gout, gouty arthropathy, pyogenic arthritis, essential hypertension, coronary atherosclerosis, ischemic heart disease, chronic ischemic heart disease, myocardial infarction, and hypercholesterolemia) that

were associated with genetically determined serum urate levels after multiple testing correction ($p < 3.35 \times 10^{-4}$). TreeWAS analysis, examining 10,750 ICD-10 diagnostic terms, identified more sub-phenotypes of cardiovascular and cerebrovascular diseases (e.g., angina pectoris, heart failure, cerebral infarction). MR analysis successfully replicated the association with gout, hypertension, heart diseases, and blood lipid levels but indicated the existence of genetic pleiotropy. Sensitivity analyses support an inference that pleiotropic effects of genetic variants on urate and metabolic traits contribute to the observational associations with CVDs. The main limitations of this study relate to possible bias from pleiotropic effects of the considered genetic variants and possible misclassification of cases for mild disease that did not require hospitalization.

Conclusion

In this study, high serum urate levels were found to be associated with increased risk of different types of cardiac events. The finding of genetic pleiotropy indicates the existence of common upstream pathological elements influencing both urate and metabolic traits, and this may suggest new opportunities and challenges for developing drugs targeting a common mediator that would be beneficial for both the treatment of gout and the prevention of cardiovascular comorbidities.

Author summary

Why was this study done?

- Serum urate level has been extensively studied in epidemiological studies in relation to various diseases, but the extent of any causal effect is still unclear, making it difficult to evaluate the clinical importance of urate.

- Mendelian randomization (MR) uses naturally occurring genetic variants as instruments to infer the causal role of a risk factor in a disease or outcome of interest. Previous MR studies were typically hypothesis driven, and few studies have comprehensively investigated how serum urate level might influence overall health.

What did the researchers do and find?

- We implemented a phenome-wide association study (PheWAS) followed by a Bayesian analysis of tree-structured phenotypic model (TreeWAS) and MR analyses to explore the association between urate and a broad range of disease outcomes in the UK Biobank.

- We identified gout, hypertension, hypercholesterolemia, and a multitude of cardiovascular and cerebrovascular diseases (e.g., coronary atherosclerosis, myocardial infarction, angina pectoris, ischemic heart disease, heart failure, and cerebral infarction) that were associated with serum urate levels.

- MR analysis using various GWAS consortia data successfully replicated these associations but indicated the existence of genetic pleiotropy. Sensitivity analyses examining

the pleiotropic effects of urate genetic risk loci on a set of metabolic traits support an inference that genetic pleiotropy contributes to observational association between urate and different types of cardiac events.

What do these findings mean?

- There was little evidence of causality; instead, the pleiotropy of genetic variants on urate and metabolic traits is indicated for the observed associations with cardiovascular/metabolic diseases.

- The linked biological pathways between urate and metabolic traits indicated that the frequent coexistence of gout with hypertension, cardiovascular diseases (CVDs), and hyperlipidemia is a range of interrelated disease outcomes due to linked pathogenic components, rather than isolated events.

- These findings support the European League against Rheumatism (EULAR) recommendation of systematic screening and assessment of cardiovascular/metabolic comorbidities in gout patients.

- These findings may suggest new opportunities and challenges for developing drugs targeting a more distal mediator that would be beneficial for both the treatment of gout and the prevention of cardiovascular/metabolic comorbidities.

Introduction

The role of urate has been explored in a large number of observational studies in relation to a multitude of health outcomes [1]. Apart from gout, compelling evidence exists for the associations between high serum urate level and an increased risk of non–crystal deposition disorders, including hypertension, cardiovascular diseases (CVDs), and metabolic syndrome [2,3]. Although considerable research efforts have been made in trying to understand the pathological role of urate in such disorders, its causal role has not been clearly established. Therefore, it has been argued that either these associations are confounded by other risk factors, such as obesity, or they represent reverse causality [4].

As is typical in complex traits, genetic determinants are implicated in the regulation of serum urate levels. Genetic studies among twins and families have reported a substantial heritable component of serum urate level with an estimated heritability of 40%–60% [5,6]. The genetic determinants of serum urate level have been explored in several genome-wide association studies (GWASs) [7–10] and the wealth of resultant data allows for the identification and application of genetic variants as instruments to help separate causal from noncausal associations, given that genotypes are generally independent of environmental exposures and the transmission of genetic information is usually unidirectional. Investigating the associations between genetic variants related to serum urate and disease outcomes might help provide causal evidence in support of the hypotheses that link urate to multiple clinical disorders. Previous mendelian randomization (MR) studies using the genetic variants as instruments of serum urate levels reported inconsistent findings [11–14]. While some supported a causal effect on health outcomes beyond gout (e.g., diabetic macrovascular disease, CVD mortality,

and sudden cardiac death), the majority reported no causal relationships [1]. Specifically, most of the negative results of previous MR studies are perhaps due to the selection of genes selectively involved in renal handling of urate, while a substantial portion of cardiovascular risk is probably due to pleiotropic genes controlling for xanthine oxidase activity and urate production.

Our recently published MR–phenome-wide association study (PheWAS) analysis on the interim release data of UK Biobank (n = 120,091) provided an overview of the disease outcomes that were associated with the urate genetic risk loci [15]. Our study demonstrated that serum urate level shared the same genetic risk loci with multiple disease outcomes, particularly those related to cardiovascular/metabolic diseases and autoimmune disorders [15]. These findings provide a rationale for further investigating whether these cross-phenotype associations are causal. Although we have applied multiple methodologies to distinguish the PheWAS associations that were causal from those due to pleiotropy or genetic linkage, the use of the interim release data of UK Biobank set power limitations to our investigation and did not allow us to investigate less prevalent phenotypes. The release of the full UK Biobank GWAS genotype dataset provides a unique opportunity to further explore the previous MR-PheWAS findings, repeat analysis with the larger available cohort, and include phenotypes that were not investigated in the previous study due to an insufficient number of cases.

In this study, we performed an updated phenome-wide mendelian randomization study (PWMR) by using data from the full UK Biobank cohort. A weighted polygenic risk score (GRS) incorporating effect estimates of multiple genetic risk loci taken from the most recent and largest GWAS of serum urate was employed as a proxy of serum urate level [8]. The framework of phenome was defined by using both the PheCODE schema (also used in the previous MR-PheWAS) [15] and a novel Bayesian analysis framework, termed TreeWAS (tree-structured phenotypic model) [16]. Any replication of previous findings and/or novel findings was further explored in this study.

Methods

This study is reported as per the STROBE guideline (S1 STROBE Checklist). UK Biobank has ethics approval from the North West Multi-Centre Research Ethics Committee (11/NW/0382). Appropriate informed consent was obtained from participants and ethical approval was covered by the UK Biobank, from which data for this work were obtained (under approved data request application ID 10775). Any additional ethical approval was adjudged unnecessary for the present study. Although there is no formal or documented protocol for this study, the main analyses of PheWAS, TreeWAS, and the replication study were prespecified; the sensitivity analysis for the GRS of genetic polymorphisms involved in renal handling of urate was supplemented according to the reviewer's comments to better interpret the findings.

UK Biobank

UK Biobank is a large-scale, population-based prospective cohort study, which recruited over 500,000 participants aged between 40 and 69 years in 2006–2010 and combined extensive measurement of baseline data and genotype data with linked national medical records (e.g., inpatient hospital episode records, cancer registry, and death registry) for longitudinal follow-up. This study was constrained to a subset of unrelated White British individuals with high-quality genotype data in order to minimize the influence of diverse population structure within UK Biobank. Details about genotype data and phenotype data and the procedures of quality control are described in S1 Text.

Weighted genetic risk score

To generate a genetic proxy for serum urate, genetic variants associated with urate were searched across the GWAS catalogue and literature. Thirty-one genetic variants associated with urate among European populations were identified from previous GWASs [7,8] and were selected as components of the genetic proxy for serum urate level. The overall proportion of variance (adjusted R^2) of urate explained by the 31 genetic variants was around 7% [8]. The SNP effect on urate (effect size and standard error [SE]) was taken from the largest meta-analysis of GWASs performed by the Global Urate Genetics Consortium (GUGC) [8]. A weighted GRS was constructed by incorporating effect estimates of the 31 urate variants for UK Biobank participants. Specifically, the GRS was created by adding up the number of urate-increasing alleles for each SNP weighted for the SNP effect size on serum urate level (regression beta coefficients) and then adding this weighted score for all 31 SNPs.

Phenome framework

We analyzed three phenotypic datasets (i.e., inpatient hospital records, cancer registry data, and death registry data) available in the UK Biobank database. As we were interested in disease phenotypes, the ontology of the phenome was defined based on the ICD codes in the electronic medical records. We pooled the hospital episode data, cancer registry data, and death registry data together and included both the primary and secondary ICD codes. Individual ICD codes could not be directly used to define the phenome, as they represent specific sub-phenotypes of a similar set of diseases, instead of independent phenotypes. To account for the correlations between ICD codes, we applied two strategies: (i) the PheCODE schema that has been recently updated and successfully adopted in our previous MR-PheWAS [15]; and (ii) a novel Bayesian analysis of a TreeWAS that was developed by researchers from the Wellcome Trust Centre for Human Genetics [16].

PheCODE schema. The PheCODE system was developed to combine one or more related ICD codes into distinct disease groups [17]. To develop a phenotyping method applicable to the ICD-10 coding system in UK Biobank, we created a map to match ICD-9/10 codes to phecodes [15]. The latest version of the PheCODE system includes 1,866 hierarchical phenotype codes that could be directly matched to the ICD-9/10 codes and provides a scheme to automatically exclude patients that have similar or potentially overlapping disease states from the corresponding control group (e.g., excluding type 1 diabetes from being in the control group when analyzing the phenotype of type 2 diabetes). The PheCODE map is made publicly accessible via the link https://phewascatalog.org/phecodes_icd10.

TreeWAS. A novel Bayesian analysis on a TreeWAS has recently been developed to interrogate the increasingly specific sub-phenotypes defined by the ICD-10 coding system. It has been suggested that this model has higher statistical power for detecting genotype-phenotype associations [16]. In principle, this phenotyping method models the genetic coefficients across all phenotypes as a set of random variables. To model the correlations of the hierarchical tree-like structure of ICD-10 codes (termed as TreeWAS), a Markov process is applied to allow the genetic coefficients to evolve down the tree trunk and branches. The tree structure is determined based on the classification hierarchy of the ICD-10 coding system, in which each node in the tree represents a clinical term in the classification. More details about the tree-structured phenotyping process are described elsewhere [16].

Statistical analysis

To take advantage of both phenotyping models, we explored the association between the weighted GRS of urate and the phenome framework defined by both the PheCODE schema

(described as PheWAS analysis) and the tree-structured phenotypic model (described as Tree-WAS analysis). The correlation with weighted GRS was examined for a number of potential confounding factors including sex, age, body mass index (BMI), assessment center, and the first 5 genetic principal components (PCs). In the PheWAS analysis, associations between weighted GRS and phecodes (with no fewer than 20 cases) were examined by logistic regression. Given that phenotypes investigated are not totally independent in the PheCODE system, because multiple levels of phenotypic granularity were used for the definition of the case-control groups, we applied the false discovery rate (FDR) method (corresponding to the FDR of $q < 0.05$) to account for multiple comparisons instead of the more stringent Bonferroni correction [18]. In the TreeWAS analysis, associations between the weighted GRS and the phenome variables were tested by the Bayesian network analysis at both terminal and internal nodes of the tree structure. The marginal posterior probability (PP) for each node in the tree (where its genetic coefficient was nonzero) and the corresponding maximum a posteriori (MAP) effect estimate with 95% credible interval were determined by using the MAP estimator. Any association with any node of the tree at the PP ≥ 0.95 was reported for further investigation. Details about the TreeWAS analysis have been described previously [16]. All the statistical analyses were implemented by R 3.3.2.

Replication in MR-base database

To validate findings, PheWAS associations were further examined in the MR-base database for replication in different populations [19,20]. We used this platform to make causal inference by performing two-sample MR analysis using available GWAS consortia data. We applied the simplest inverse variance weighted mendelian randomization (MR IVW) approach as crude analysis; if there was horizontal pleiotropy that violated the assumptions of the MR IVW, we applied a mixture-of-experts machine learning framework of mendelian randomization (MR-MoE) to predict the performance of three main classes of MR analytical approaches (mean-based, median-based, and mode-based methods) in the context of different models of pleiotropy and then selected the most likely unbiased causal estimate for each specific circumstance [20]. Full details of these MR approaches, including their different assumptions, are provided in S1 Text and S1 Table. The schematic presentation of the overall study design is shown in Fig 1.

Sensitivity analyses

We created a separate GRS of seven genetic polymorphisms, which are involved in renal handling of urate (six urate transporter-coding genes: *SLC22A12*, *SLC2A9*, *ABCG2*, *SLC22A11*, *SLC17A1*, *SLC16A9*, and one urate transporter-related scaffolding gene, *PDZK1*), as a sensitivity analysis. We then performed additional sensitivity analyses to further explore any pleiotropic associations. To identify genetic variants showing pleiotropy, we examined their association with a set of metabolic traits (i.e., BMI, waist-to-hip ratio [WHR], total cholesterol [TC], low-density lipoprotein cholesterol [LDL-c], high-density lipoprotein cholesterol [HDL-c], fasting glucose, 2-hour glucose, glycoproteins, systolic blood pressure [SBP], and diastolic blood pressure [DBP]) through publicly available resources from various GWAS consortia (a summary of these GWASs is provided in S1 Text). An association was declared as pleiotropic when these GWAS summary data reported any association between the serum urate risk loci and these metabolic traits at $p < 1.61 \times 10^{-3}$ (the threshold was determined based on the Bonferroni correction with a significance level of $\alpha = 0.05$ divided by the number of 31 serum urate risk loci analyzed in this study). These 31 urate genetic risk loci were then divided into five categories, accordingly: (i) urate-specific loci, including 14 SNPs with no pleiotropic effect
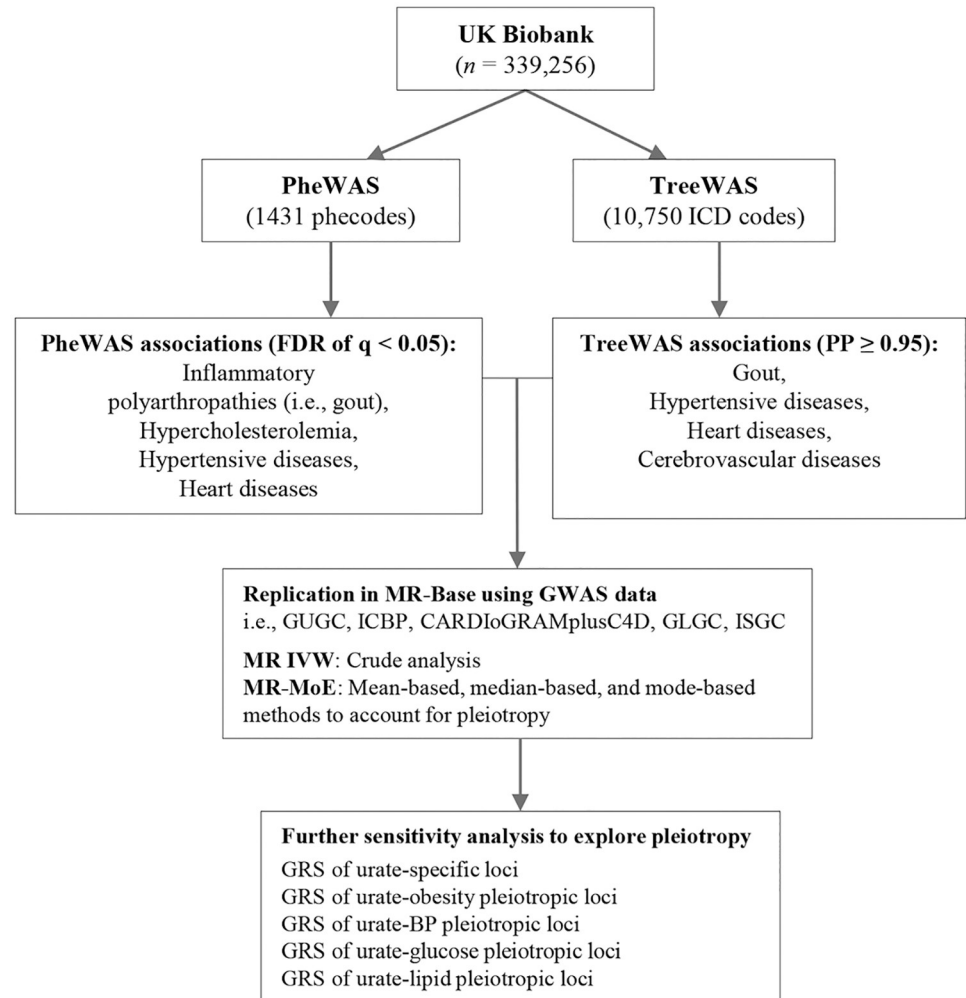
Fig 1. Schematic representation of the study design. BP, blood pressure; CARDIoGRAMplusC4D, Coronary ARtery DIsease Genome wide Replication and Meta-analysis (CARDIoGRAM) plus The Coronary Artery Disease (C4D) Genetics consortium; FDR, false discovery rate; GLGC, Global Lipids Genetic Consortium; GRS, polygenic risk score; GUGC, Global Urate Genetic Consortium; GWAS, genome-wide association study; ICBP, International Consortium for Blood Pressure; ISGC, Ischaemic stroke Genetic Consortium; MR IVW, inverse variance weighted mendelian randomization; MR-MoE, a mixture-of-experts machine learning framework of mendelian randomization; PheWAS, phenome-wide association study; PP, posterior probability; TreeWAS, tree-structured phenotypic model.

https://doi.org/10.1371/journal.pmed.1002937.g001

on the examined metabolic traits (S2 Table); (ii) urate-obesity pleiotropic loci, including 10 SNPs with pleiotropic effects on BMI or WHR (S3 Table); (iii) urate-BP pleiotropic loci, including 10 SNPs with pleiotropic effects on blood pressures (BPs) (i.e., DBP and SBP) (S4 Table); (iv) urate-lipid pleiotropic loci, including 6 SNPs with pleiotropic effects on lipids (i.e., TC, LDL-c, and HDL-c) (S5 Table); and (v) urate-glucose pleiotropic loci, including 3 SNPs with pleiotropic effects on blood glucose (fasting glucose, 2-hour glucose, glycoproteins) (S6 Table). A set of GRSs were created accordingly to recalculate the effect estimates in PheWAS analysis.

Results

We included 339,256 unrelated White British individuals from the full UK Biobank cohort, consisting of 157,146 men and 182,110 women. The mean age of the study population was

56.87 (SD: 7.99) and the mean BMI was 27.40 (SD: 4.76) kg/m^2 at the time of recruitment. Other sociodemographic characteristics of the study population are summarized in S7 Table. The mean value of weighted GRS among the study population was 0.44 (SD: 0.31), which is equivalent to 0.44 mg/dL of serum urate level. The correlations between the weighted GRS and potential confounding factors (i.e., age, sex, BMI, assessment center, and the PCs) are provided in S7 Table. Of these, two variables (i.e., assessment center and the PCs) were statistically significantly correlated with the weighted GRS and therefore were adjusted as covariates.

PheWAS and TreeWAS associations

Within the study population, we identified 10,750 unique ICD-10 codes and 3,113 ICD-9 codes in total. After mapping the diagnostic ICD-10/9 codes in UK Biobank to phecodes, the phenome defined by PheCODE schema consisted of 1,853 distinct phecodes among the study population. After filtering out the phecodes with less than 20 cases, PheWAS analysis was performed for 1,431 phecodes (median number of cases: 345 [range, 20–107,298]) that could be classified into 17 broadly related disease categories (Table 1). Associations with the weighted GRS of urate were examined for 1,431 case-control groups, leading to an adjusted significance threshold of $p < 3.35 \times 10^{-4}$ (corresponding to the FDR of q < 0.05) to account for multiple testing. The derived summary PheWAS data are provided in S1 Data. Of these, 13 phecodes were identified to be associated with genetically determined high serum urate level at $p < 3.35 \times 10^{-4}$ (Table 2). These phecodes represent 4 disease groups: inflammatory polyarthropathies ($p = 4.97 \times 10^{-19}$), hypertensive disease ($p = 6.02 \times 10^{-7}$), circulatory disease ($p = 3.29 \times 10^{-4}$), and metabolic disorders (p = 3.33×10^{-4}); and 9 disease outcomes: gout ($p = 4.27 \times 10^{-123}$), gouty arthropathy ($p = 1.39 \times 10^{-5}$), pyogenic arthritis (p = 2.87×10^{-4}), essential hypertension ($p = 6.26 \times 10^{-7}$), coronary atherosclerosis ($p = 1.17 \times 10^{-5}$), ischemic heart disease ($p = 1.73 \times 10^{-5}$), chronic ischaemic heart disease ($p = 1.52 \times 10^{-5}$), myocardial infarction ($p = 5.23 \times 10^{-5}$), and hypercholesterolemia ($p = 3.34 \times 10^{-4}$).

In the Bayesian analysis framework, containing 10,750 diagnostic terms, a total of 27 parent/child nodes of ICD-10 terms were identified with PP $\geq$ 0.95. They were clustered mainly in five branches of the hierarchical tree structure (Fig 2 and S8 Table): (i) block M10: gout (PP = 1.00) and its sub-phenotypes M10.0 (idiopathic gout) and M10.9 (gout, unspecified); (ii) block I10-I15: hypertensive disease (PP > 0.99) and its sub-phenotype I10 (essential hypertension); (iii) block I20-I25: ischemic heart diseases (PP > 0.99), and its sub-phenotypes: I20 (angina pectoris), I21 (acute myocardial infarction), I25 (chronic ischemic heart disease), I25.1 (atherosclerotic heart disease), and I25.2 (old myocardial infarction); (iv) block I30-I52: other forms of heart disease (PP > 0.99), and its sub-phenotypes I50 (heart failure) and I50.1 (left ventricular failure); and (v) block I60-I69: cerebrovascular diseases (PP > 0.99), and its sub-phenotype I10 (cerebral infarction).

Findings from PheWAS and TreeWAS were generally consistent in their associations with gout, hypertensive disease, and heart diseases, while a number of additional sub-phenotypes were identified by TreeWAS. Association with the disease group of inflammatory polyarthropathies was statistically significant in PheWAS (OR = 1.27, 95% CI: 1.21–1.34, $p = 4.97 \times 10^{-19}$) but had a moderate PP in TreeWAS (OR = 1.07, 95% CI: 1.06–1.08, PP = 0.76). We examined the specific diseases included in this disease group (M05-M06: rheumatoid arthritis [RA], M07: psoriatic and enteropathic arthropathies, M08-09: juvenile arthritis, M10: gout, and M11-14: arthropathies and other arthritis), and only gout had a significant association in PheWAS analysis. Association with cerebrovascular diseases had a high PP in TreeWAS (OR = 1.07, 95% CI: 1.06–1.08, PP > 0.99) but did not reach the significance threshold of PheWAS (OR = 1.08, 95% CI: 0.99–1.16, $p = 0.070$), although their estimates were of the

Table 1. The number of phenotypes and cases in each disease category.

Disease categories	Number of phenotypes	Number of cases		
		Minimum	Mean	Maximum
Circulatory diseases	140	434	3,581	107,298
Congenital anomalies	45	102	230	1,480
Dermatological diseases	74	283	2,544	89,976
Diseases in sense organs	104	253	1,228	31,845
Digestive diseases	143	551	3,123	62,862
Neoplasms	129	493	2,558	84,098
Infectious diseases	48	190	958	8,600
Endocrine and metabolic diseases	103	154	1,590	35,954
Hematopoietic diseases	40	228	1,200	10,095
Neurological diseases	69	224	1,180	32,194
Respiratory diseases	71	674	2,448	49,782
Mental disorders	64	260	1,493	23,226
Genitourinary diseases	140	655	2,536	82,964
Pregnancy complications	28	237	914	7,518
Musculoskeletal diseases	109	347	2,847	59,852
Clinical symptoms	27	711	3,741	33,553
Injuries and poisonings	97	388	1,079	13,303

https://doi.org/10.1371/journal.pmed.1002937.t001

same direction. We recalculated the PheWAS estimates by adding up self-reported stroke cases to increase statistical power, but the corresponding estimates were still not statistically significant (OR = 1.05, 95% CI: 0.99–1.13, p = 0.130).

Replication in MR-base database

To validate the findings, we performed two-sample MR analyses on associated diseases (i.e., gout, RA, coronary heart disease [CHD], myocardial infarction, ischemic stroke) or on their corresponding intermediate traits or surrogate outcomes (i.e., SBP, DBP, TC, LDL-c, HDL-c) (Table 3). Results from MR IVW suggested that a genetically determined high serum urate

Table 2. Disease outcomes associated with the weighted GRS of urate in PheWAS analysis.

Phecode	Disease outcomes	n_cases	n_controls	beta	SE	OR (95% CI)	p-value
274.1	Gout	2,532	335,108	1.682	0.071	5.37 (4.67–6.18)	4.27×10^{-123}
714	Inflammatory polyarthropathies	15,408	320,862	0.244	0.027	1.27 (1.21–1.34)	4.97×10^{-19}
401	Hypertension	63,694	274,477	0.076	0.015	1.07 (1.05–1.11)	6.02×10^{-7}
401.1	Essential hypertension	63,442	274,477	0.077	0.015	1.08 (1.05–1.11)	6.26×10^{-7}
411.4	Coronary atherosclerosis	25,795	311,554	0.096	0.022	1.10 (1.05–1.14)	1.17×10^{-5}
274.11	Gouty arthropathy	88	335,108	1.631	0.375	5.10 (2.45–10.66)	1.39×10^{-5}
411.8	Chronic ischemic heart disease	25,567	311,554	0.095	0.022	1.09 (1.05–1.14)	1.52×10^{-5}
411	Ischemic heart disease	25,617	311,554	0.094	0.022	1.09 (1.05–1.14)	1.73×10^{-5}
411.2	Myocardial infarction	9,829	311,554	0.138	0.034	1.14 (1.07–1.22)	5.23×10^{-5}
711.1	Pyogenic arthritis	270	277,590	0.742	0.205	2.10 (1.41–3.13)	2.87×10^{-4}
459.9	Circulatory disease	107,298	230,622	0.046	0.013	1.04 (1.02–1.07)	3.29×10^{-4}
277	Metabolic disorders	35,954	302,209	0.067	0.019	1.07 (1.03–1.11)	3.33×10^{-4}
272.11	Hypercholesterolemia	27,040	308,948	0.077	0.021	1.08 (1.04–1.12)	3.34×10^{-4}

Abbreviations: GRS, polygenic risk score; PheWAS, phenome-wide association study

https://doi.org/10.1371/journal.pmed.1002937.t002
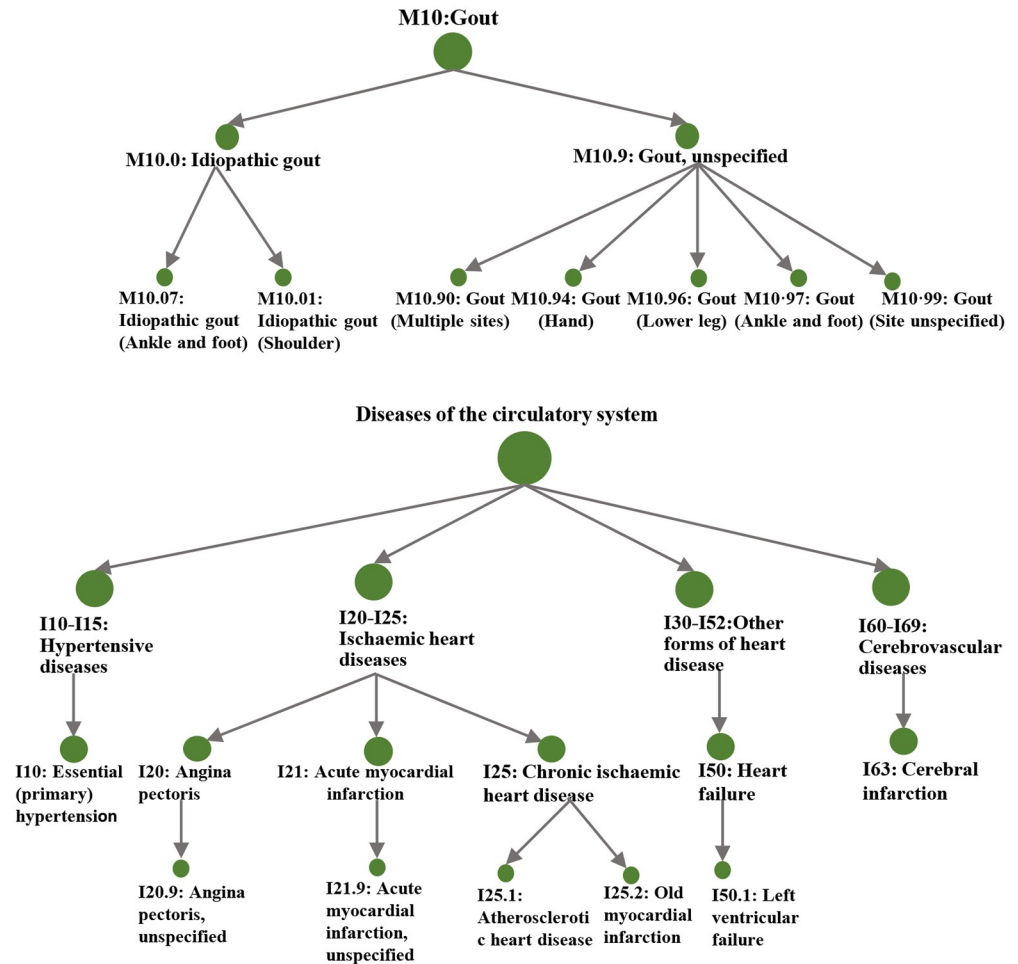
Fig 2. A hierarchical structure of disease outcomes associated with urate in TreeWAS analysis. TreeWAS, tree-structured phenotypic model.

https://doi.org/10.1371/journal.pmed.1002937.g002

level was associated with increased risk of gout (OR = 4.53, 95% CI: 3.64–5.64; p_{effect} = 9.66 × 10^{-42}), DBP (OR = 1.04, 95% CI: 1.02–1.08; p_{effect} = 0.044), SBP (OR = 1.03, 95% CI: 1.00–1.06; p_{effect} = 0.050), CHD (OR = 1.10, 95% CI: 1.02–1.19; p_{effect} = 0.014), myocardial infarction (OR = 1.11, 95% CI: 1.02–1.20; p_{effect} = 0.017), and decreased level of HDL-c (OR = 0.93, 95% CI: 0.88–0.98; p_{effect} = 0.007) but had no effect on ischemic stroke (OR = 1.03, 95% CI: 0.93–1.14; p_{effect} = 0.582). A test for directional horizontal pleiotropy indicated the existence of pleiotropy on the causal estimates of DBP ($p_{pleiotropy}$ = 0.014), SBP ($p_{pleiotropy}$ = 0.003), CHD ($p_{pleiotropy}$ = 0.008), myocardial infarction ($p_{pleiotropy}$ = 0.008), and HDL-c ($p_{pleiotropy}$ = 0.016), indicating the MR IVW estimates are likely biased. Using the MR-MoE analysis to select the most appropriate MR approach to deal with different models of pleiotropy, we derived a statistically significant causal estimate only for gout (OR = 4.50, 95% CI: 3.62–5.59, p_{effect} = 3.35 × 10^{-77}) (Table 3). Causal estimates from each of the MR analytical approaches are provided in S9–S17 Tables.

Table 3. Replication of MR effect estimates in MR-base database.

Outcome	beta	SE	OR (95% CI)	p_{effect}	$p_{pleiotropy}$[*]	n_cases	n_total	Data source
Replication of significant PheWAS associations								
Gout								
PheWAS	1.682	0.071	5.37 (4.67–6.18)	4.27×10^{-123}	- -	2,532	337,640	UKBB
MR IVW	1.511	0.112	4.53 (3.64–5.64)	9.66×10^{-42}	0.485	2,115	67,259	GUGC
MR-MoE	1.504	0.081	4.50 (3.62–5.59)	3.35×10^{-77}	- -			
Hypertension								
PheWAS	0.076	0.015	1.07 (1.05–1.11)	6.02×10^{-7}	- -	63,694	338,171	UKBB
DBP								
MR IVW	0.042	0.020	1.04 (1.02–1.08)	0.044	0.014	- -	69,395	ICBP
MR-MoE	0.034	0.019	1.03 (0.99–1.07)	0.084	- -	- -		
SBP								
MR IVW	0.031	0.015	1.03 (1.00–1.06)	0.050	0.003	- -	69 395	ICBP
MR-MoE	0.008	0.007	1.01 (0.99–1.02)	0.266	- -	- -		
CHD								
PheWAS	0.094	0.022	1.09 (1.05–1.14)	1.73×10^{-5}	- -	25,617	337,171	UKBB
MR IVW	0.098	0.038	1.10 (1.02–1.19)	0.014	0.008	60,801	123,504	CARDIoGRAMplusC4D
MR-MoE	0.047	0.028	1.05 (0.99–1.10)	0.086	- -			
Myocardial infarction								
PheWAS	0.138	0.034	1.14 (1.07–1.22)	5.23×10^{-5}	- -	9,829	321,383	UKBB
MR IVW	0.105	0.041	1.11 (1.02–1.20)	0.017	0.008	43,676	128,199	CARDIoGRAMplusC4D
MR-MoE	0.058	0.030	1.06 (0.99–1.12)	0.055	- -			
Hypercholesterolemia								
PheWAS	0.077	0.021	1.08 (1.04–1.12)	3.34×10^{-4}	- -	27,040	335,988	UKBB
TC								
MR IVW	0.028	0.036	1.03 (0.96–1.10)	0.440	0.602	- -	173,082	GLGC
MR-MoE	0.005	0.026	1.00 (0.95–1.06)	0.848	- -			
HDL-c								
MR IVW	-0.075	0.026	0.93 (0.88–0.98)	0.007	0.016	- -	187,167	GLGC
MR-MoE	-0.030	0.015	0.97 (0.94–1.01)	0.058	- -			
LDL-c								
MR IVW	0.011	0.023	1.05 (0.95–1.16)	0.627	0.175	- -	187,365	GLGC
MR-MoE	0.011	0.023	1.05 (0.95–1.16)	0.627	- -			
Replication of additional TreeWAS associations								
Ischemic stroke								
PheWAS	0.071	0.040	1.08 (0.99–1.16)	0.070	- -	9,528	338,172	UKBB
MR IVW	0.029	0.052	1.03 (0.93–1.14)	0.586	0.290	10,307	19,326	ISGC
MR-MoE	0.029	0.052	1.03 (0.93–1.14)	0.586	- -			

[*]Test for directional horizontal pleiotropy.

Abbreviations: CARDIoGRAMplusC4D, Coronary ARtery DIsease Genome wide Replication and Meta-analysis (CARDIoGRAM) plus The Coronary Artery Disease (C4D) Genetics consortium; CHD, coronary heart disease; DBP, diastolic blood pressure; GLGC, Global Lipids Genetics Consortium; GUGC, Global Urate Genetics Consortium; HDL-c, high-density lipoprotein cholesterol; ICBP, International Consortium of Blood Pressure; ISGC, Ischaemic stroke Genetic Consortium; LDL-c, low-density lipoprotein cholesterol; MR, mendelian randomization; MR IVW, inverse variance weighted mendelian randomization; MR-MoE, a mixture-of-experts machine learning framework of mendelian randomization; PheWAS, phenome-wide association study; SBP, systolic blood pressure; TC, total cholesterol; TreeWAS, tree-structured phenotypic model; UKBB, UK Biobank

https://doi.org/10.1371/journal.pmed.1002937.t003

Sensitivity analyses

PheWAS analysis using the GRS of 7 SNPs involved in renal handling of urate showed significant associations with gout ($p = 3.04 \times 10^{-91}$) and related diseases (e.g., inflammatory polyarthropathies, gouty arthropathy) after FDR correction. A GRS of the remaining 24 SNPs (excluding genetic polymorphisms involved in renal handling of urate) showed significant associations with gout, hypertension, hypercholesterolemia, and CVDs (e.g., coronary atherosclerosis, ischemic heart diseases, and myocardial infarction) (S18 Table). Given that most of the related outcomes were CVDs, we performed further sensitivity analyses to examine the potential of any pleiotropy effect of urate risk variants on metabolic traits. We recalculated the PheWAS estimates by using a number of GRSs created based on their association with a set of metabolic traits (Fig 3 and S19 Table), and the specific metabolic traits investigated were further determined by the availability of summary GWAS data. The GRS of urate-specific loci was only associated with gout and its upper disease group of inflammatory polyarthropathies, but not with any cardiovascular/metabolic diseases. In contrast, the GRSs of pleiotropic loci on obesity, BP, lipids, and glucose showed significant association with both gout and the CVDs. Specifically, the GRS of pleiotropic loci on lipids was significantly associated with all CVDs, including hypertensive diseases (i.e., essential hypertension), heart diseases (i.e., ischemic heart diseases), and metabolic disorders (i.e., hypercholesterolemia). Additionally, the GRS of pleiotropic loci on glucose was significantly associated with diabetes (i.e., type 2 diabetes). When removing any group of pleiotropic loci from the creation of GRS, their associations with hypertensive diseases, heart diseases, and metabolic disorders were not statistically significant (S20 Table). The effects of pleiotropic loci (mapped with genes) on serum urate level against their effects on four representative disease outcomes were plotted in S1 Fig, in which the two urate transporter genes (*SLC2A9* and *ABCG2*) are recognised as the leading loci driving the association with gout, the *GCKR* gene is the leading locus driving the association with hypercholesterolemia, and the *PTPN11/ATXN2* gene is the leading locus driving the association with hypertension and ischemic heart diseases.

Discussion

The present study demonstrated that genetically determined high serum urate level was consistently associated with increased risk of several disease groups, including inflammatory polyarthropathies (e.g., gout and gouty arthropathy), hypertensive diseases (e.g., essential hypertension), heart diseases (e.g., coronary atherosclerosis, myocardial infarction, angina pectoris, ischemic heart disease, and heart failure), and metabolic disorders (e.g., hypercholesterolemia). This study, using data from the full UK Biobank cohort ($n = 339,256$), verified the associations discovered in the previous MR-PheWAS study based on the interim release of UK Biobank genetic data ($n = 120,091$) [15] and identified a number of new sub-phenotypes of diseases (e.g., gouty arthropathy, angina pectoris, and heart failure). Some disease outcomes (e.g., disorders of iron metabolism, celiac disease) reported in the previous study were not identified in the present study, as these associations were derived from the genetic linkage disequilibrium between two single variants and therefore were diluted by the use of a weighted GRS of multiple genetic instruments. Association between urate and the risk of gout, hypertension, CHD, myocardial infarction, and a decreased level of HDL-c was successfully replicated in different European populations by analyzing various GWAS consortia data documented in the MR-base database [19], but a causal relationship was only supported for gout. Overall, findings from the current study support the observational associations between high serum urate level and increased risk of hypertensive diseases, heart diseases, and metabolic disorders and also

Fig 3. Network plot of the sensitivity analyses of PheWAS using different sets of weighted GRS. The red circles represent the disease outcomes associated with the weighted GRS of the 31 urate genetic rick loci; the blue circles represent disease outcomes associated with the weighted GRS of urate-specific risk loci; the green circles represent diseases outcomes associated with urate-obesity pleiotropic loci; the orange circles represent disease outcomes associated with urate-lipid pleiotropic loci; and the pink circles represent disease outcomes associated with urate-lipid pleiotropic loci. E70-E90, metabolic disorders; GRS, genetic risk score; I10-I15, hypertensive diseases; I20-I25, ischemic heart diseases; M05-M14, inflammatory polyarthropathies; PheWAS, phenome-wide association study.

https://doi.org/10.1371/journal.pmed.1002937.g003

indicated that these associations were more likely due to genetic pleiotropy instead of causality.

A recent umbrella review summarized the published MR studies and examined the causal relationship of serum urate level with a wide range of health outcomes, including gout, cardiovascular, metabolic, and neurocognitive disorders, and for the majority of investigated traits, causality was not verified [1]. There were nine disease outcomes (e.g., diabetic macrovascular disease, arterial stiffness [internal diameter of carotid artery], adverse renal events, Parkinson disease, lifetime anxiety disorders, memory performance, CVD mortality, sudden cardiac

death, and gout) reported to have a nominally statistically significant causal relationship with urate, but most of them presented with discordant results between MR studies or suffered from methodological limitations (e.g., inadequate study power, invalid genetic instruments), and only that for gout was verified based on convincing evidence.

Specifically, our finding that genetically predicted serum urate level is causally associated with increased risk of gout is not surprising, as it is well known that the causal factor of gout is represented by the monosodium urate crystals (MSUs), which leads to acute local inflammation in joints. Moreover, this study also detected an association between urate and the disease group of inflammatory polyarthropathies. To investigate if there were any other types of inflammatory polyarthropathies (beyond gout) associated with urate, we examined the association of urate with all specific diseases included in this group, but none of them were statistically significant. When excluding gout from this disease group, the association was not statistically significant any longer, indicating the observed association was actually driven by gout.

Numerous epidemiological studies have reported that elevated serum urate level is related to increased risk of hypertension, and their relationship has been consistent, showing a dose-response relationship of similar magnitude [21]. Findings from our current study support this association, but the magnitude of estimated effect size (OR = 1.07, 95% CI: 1.05–1.11) is smaller than that of traditional epidemiological studies [22]. In our PheWAS, TreeWAS, and MR IVW analysis, we consistently showed a moderate association between urate and different types of heart disease, including coronary atherosclerosis, angina pectoris, ischemic heart diseases, acute/old myocardial infarction, and heart failure; however, the MR-MoE analysis did not support the causal inference after accounting for the presence of pleiotropy.

Large epidemiological studies have established an association between high serum urate level and the increased risk of metabolic disorders [23]. The NHANES III survey study suggested that a high serum urate level was associated with increased levels of serum LDL-c, triglycerides, TC, and apolipoprotein-B and a decreased level of HDL-c [24]. Our study further strengthened this epidemiological evidence and highlighted an association between urate and hypercholesterolemia. Our MR IVW analysis replicated the corresponding association with its surrogate outcome (i.e., HDL-c) but suggested the presence of pleiotropy instead of causality. Additionally, epidemiological studies have also indicated that high serum urate level is associated with increased risk of diabetes [25]. However, this association was not detected in the main PheWAS or TreeWAS analysis, while sensitivity analysis using the GRS of urate-glucose pleiotropic loci (i.e., *GCKR*, *IGF1R*, and *SLC16A9*) identified significant association with type 2 diabetes.

To explore how genetic pleiotropy influences the association with cardiovascular/metabolic diseases, we analysed all 31 urate loci across a set of metabolic traits and identified 14 SNPs (urate-specific loci) that were exclusively associated with urate and 17 SNPs (pleiotropic loci) that were associated with metabolic traits. When examining the urate-specific loci, their GRSs were only associated with gout and its upper disease group of inflammatory polyarthropathies, but not with any cardiovascular or metabolic diseases. In contrast, when categorizing the pleiotropic loci into different groups (e.g., GRS of urate-obesity loci, GRS of urate-BP loci, GRS of urate-lipid loci, and GRS of urate-glucose loci), the GRSs of pleiotropic loci showed consistent associations with both gout and the cardiovascular/metabolic diseases. When removing any group of pleiotropic loci from the creation of GRS (e.g., GRS of urate without pleiotropic loci on BP, or GRS of urate without pleiotropic loci on lipids), their association with heart diseases and metabolic disorders was not statistically significant. Based on these findings, our study suggests that the association between urate and CVDs is probably due to the pleiotropic effects of genetic variants on urate and metabolic traits.

Examining the associations between individual urate genetic risk loci and the related disease outcomes highlighted two loci, *GCKR* and *PTPN11/ATXN2*, which drive their association with hypercholesterolemia, hypertension, and ischemic heart disease. Pathway network analysis of the leading pleotropic genes provides some clues on how genetic pleiotropy contributes to the association between urate and cardiovascular/metabolic disease. Genetic variation in *GCKR* is shown to be associated with concentrations of urate, triglyceride, and glucose [26]. The most plausible explanation for this observation is that *GCKR* affects both serum urate and triglyceride and glucose levels by a common unconfirmed mediator, which is proposed to be glucose-6-phosphate [27]. The *GCKR* controls the hepatic production of glucose-6-phosphate, which is catabolized for triglyceride synthesis via glycolysis, pyruvate, and acetyl coenzyme A, while glucose-6-phosphate is also a precursor of purine (uric acid) metabolism [27]. Additionally, gene functional annotation of *PTPN11/ATXN2* highlights another subnetwork around hemostasis pathways, including platelet activation, aggregation, and sensitization (activated by LDL-c) [28], and these may be relevant to the observed association with hypertension and heart diseases; but how this gene influences serum urate level has not yet been clearly demonstrated.

The detection of a multitude of cross-phenotype associations in this study adds to our understanding of the extent of shared genetic/biological components between urate and metabolic traits. Further characterizing the associations between urate and disease outcomes as causal or pleiotropic contributes to our knowledge of how the role of urate should be interpreted and used in clinical practice in the management of related disease conditions. Given that the observational associations between urate and cardiometabolic diseases are more likely due to pleiotropy rather than causality, our study supports the notion that urate could be a predictor but probably not a direct target for the development of compounds that could reduce cardiovascular/metabolic disease risk. The linked biological pathways between urate and metabolic traits indicated that the frequent coexistence of gout with hypertension, CVDs, and hyperlipidemia is a range of interrelated disease outcomes due to linked pathogenic components, rather than isolated events. This supports the European League against Rheumatism (EULAR) recommendation of systematic screening and assessment of cardiovascular/metabolic comorbidities in gout patients [29]. The classification of high serum urate levels due to renal handling dysfunction or high urate production would improve the identification of gout patients with higher risk of metabolic and CVD, and promote a more selective and effective use of urate-lowering drugs. The finding of genetic pleiotropy indicates the existence of common upstream pathological elements influencing both urate and metabolic traits, and this may suggest new opportunities and challenges for developing drugs targeting a common mediator that would be beneficial for both the treatment of gout and the prevention of cardiovascular/metabolic comorbidities. This study has focused on the detection of cross-phenotype associations and highlighted the importance of pleiotropy in the links of these complex diseases. We have made efforts to try to understand the cross-phenotype association in the context of a pleiotropy model, but functionally characterizing the underlying biological mechanisms remains a challenge in this field and is worthy of further investigation.

The strengths of this study include its potential to examine a broad spectrum of disease outcomes related to urate and to reflect the shared biological relevance among associated phenotypes, given that previous MR studies were typically hypothesis driven and few studies have comprehensively investigated how serum urate level might influence overall health. Compared with the previous MR-PheWAS [15], the present study extends the prior findings by combining genetic risk loci of urate into a weighted GRS, exploring genetic pleiotropy on a set of metabolic traits systematically, investigating more disease outcomes, assessing their associations with >3-fold more cases, examining consistency of findings across two different phenotyping models to reduce the probability of false positive/negative findings due to factors related to the

model, and replicating the findings by performing two-sample MR in different populations. Our study demonstrated the performance of two phenotyping models by accounting for the differences in the specificity and granularity of different phenome definitions and by characterizing the phenotypic correlations among different levels of ICD hierarchy. TreeWAS is shown to increase statistical power and can detect new associations missed by conventional PheWAS [16]. One of the major accomplishments of this study together with the previous MR-PheWAS has been the establishment of a framework or workflow for PheWAS [15]. We believe this study would be an excellent starting point for researchers who plan to use the UK Biobank resource to comprehensively interrogate the clinical significance of biomarkers. The updated version of PheCODE schema used in this study is made available for researchers who are interested in performing PheWAS in UK Biobank.

This study also has limitations. The causal inference in our study is limited by the common difficulty of pleiotropy caused by the use of multiple genetic instruments. Although we have performed sensitivity analyses by grouping the pleiotropic loci based on metabolic traits and exploring their association separately, there is still a probability of undetected pleiotropy or the possibility that the relatively weak causal effects of urate on diseases were concealed by the strong pleiotropic effects of the genetic variants on metabolic traits. Moreover, as most cases were identified from the inpatient hospital records, this may have impaired the coverage of case ascertainment, especially for the diseases that do not usually cause events for hospitalization. The incorporation of self-reported data would improve this limitation, but it is also likely to mistakenly include patients who do not have a true diagnosis and introduce information bias. As UK Biobank is currently performing disease adjudication and processing linkages to general practice records and outpatient data, a widely covered and accurately defined criteria of case ascertainment for PheWAS study would be possible in the future.

Conclusions

Overall, when taking together the findings from PheWAS/TreeWAS, MR replication, and sensitivity analyses, we conclude a robust association between urate and a group of diseases including gout, hypertensive diseases, heart diseases, and metabolic disorders of lipids, but the causal role of urate is only supported in gout. Our study indicates that the association between urate and CVDs is probably due to the pleiotropic effects of genetic variants on urate and metabolic traits. These findings support that urate could be a good predictor for the cardiovascular/metabolic disease risk. Further investigation on therapies targeting the shared biological pathways between urate and metabolic traits would be beneficial for the treatment of gout and the primary prevention of cardiovascular/metabolic comorbidities.

Supporting information

S1 STROBE Checklist.
(DOCX)

S1 Text. Supplementary methods.
(DOCX)

S1 Data. Data underlying this study.
(XLSX)

S1 Table. A summary of MR analytical approaches and their assumptions. MR, mendelian randomization.
(DOCX)

S2 Table. A summary of genetic risk variants identified in previous urate GWAS. GWAS, genome-wide association study.
(DOCX)

S3 Table. A summary of pleiotropic loci on urate and obesity traits.
(DOCX)

S4 Table. A summary of pleiotropic loci on urate and BPs. BP, blood pressure.
(DOCX)

S5 Table. A summary of pleiotropic loci on urate and lipids.
(DOCX)

S6 Table. A summary of pleiotropic loci on urate and glucose.
(DOCX)

S7 Table. Association between the GRS of urate and potential confounding factors. GRS, polygenic risk score.
(DOCX)

S8 Table. Phenotypes associated with the weighted GRS of urate in TreeWAS analysis. GRS, polygenic risk score; TreeWAS, tree-structured phenotypic model.
(DOCX)

S9 Table. Results from MR-MoE analysis for urate and gout. MR-MoE, a mixture-of-experts machine learning framework of mendelian randomization.
(DOCX)

S10 Table. Results from MR-MoE analysis for urate and DBP. DBP, diastolic blood pressure; MR-MoE, a mixture-of-experts machine learning framework of mendelian randomization.
(DOCX)

S11 Table. Results from MR-MoE analysis for urate and SBP. MR-MoE, a mixture-of-experts machine learning framework of mendelian randomization; SBP, systemic blood pressure.
(DOCX)

S12 Table. Results from MR-MoE analysis for urate and CHD. CHD, coronary heart disease; MR-MoE, a mixture-of-experts machine learning framework of mendelian randomization.
(DOCX)

S13 Table. Results from MR-MoE analysis for urate and myocardial infarction (MI). MR-MoE, a mixture-of-experts machine learning framework of mendelian randomization.
(DOCX)

S14 Table. Results from MR-MoE analysis for urate and TC. MR-MoE, a mixture-of-experts machine learning framework of mendelian randomization; TC, total cholesterol.
(DOCX)

S15 Table. Results from MR-MoE analysis for urate and HDL-c. HDL-c, high-density lipoprotein cholesterol; MR-MoE, a mixture-of-experts machine learning framework of mendelian randomization.
(DOCX)

S16 Table. Results from MR-MoE analysis for urate and LDL-c. LDL-c, low-density lipoprotein cholesterol; MR-MoE, a mixture-of-experts machine learning framework of mendelian randomization.
(DOCX)

S17 Table. Results from MR-MoE analysis for urate and ischemic stroke (IS). MR-MoE, a mixture-of-experts machine learning framework of mendelian randomization.
(DOCX)

S18 Table. Sensitivity analysis by using the GRS of genetic polymorphisms involved in renal handling of urate. GRS, polygenic risk score.
(DOCX)

S19 Table. Sensitivity analysis by including pleiotropic loci on metabolic traits.
(DOCX)

S20 Table. Sensitivity analysis by excluding the pleiotropic loci of metabolic traits.
(DOCX)

S1 Fig. Scatterplots of the effects of pleiotropic loci (mapped with genes) on serum urate levels against their effects on disease outcomes. The size of the points was scaled to be inversely proportional to the standard errors of the effect sizes. Two urate transporter genes (*SLC2A9* and *ABCG2*) are recognized as the leading loci driving the association with gout, the *GCKR* gene is the leading locus driving the association with hypercholesterolemia, and the *PTPN11/ATXN2* gene is the leading locus driving the association with hypertension and ischemic heart diseases.
(TIF)

Author Contributions

Conceptualization: Xue Li, Ioanna Tzoulaki, Joshua C. Denny, Paul Mckeigue, Harry Campbell, Evropi Theodoratou.

Data curation: Xue Li, Xiangrui Meng, Wei-Qi Wei, Aliya Gifford, Tian Yang, Tim Varley.

Formal analysis: Xue Li, Wei-Qi Wei.

Funding acquisition: Evropi Theodoratou.

Methodology: Xue Li, Xiangrui Meng, Yazhou He, Athina Spiliopoulou, Maria Timofeeva, Peter Joshi, Joshua C. Denny, Paul Mckeigue, Evropi Theodoratou.

Resources: Athina Spiliopoulou, Joshua C. Denny, Harry Campbell.

Software: Xue Li, Xiangrui Meng, Yazhou He, Peter Joshi.

Supervision: Ioanna Tzoulaki, Paul Mckeigue, Harry Campbell, Evropi Theodoratou.

Writing – original draft: Xue Li.

Writing – review & editing: Xiangrui Meng, Yazhou He, Athina Spiliopoulou, Maria Timofeeva, Wei-Qi Wei, Aliya Gifford, Tian Yang, Tim Varley, Ioanna Tzoulaki, Peter Joshi, Joshua C. Denny, Paul Mckeigue, Harry Campbell, Evropi Theodoratou.

References

1. Li X, Meng X, Timofeeva M, Tzoulaki I, Tsilidis KK, Ioannidis JP, et al. Serum uric acid levels and multiple health outcomes: umbrella review of evidence from observational studies, randomised controlled

trials, and Mendelian randomisation studies. BMJ. 2017; 357:j2376. https://doi.org/10.1136/bmj.j2376 PMID: 28592419

2. Feig DI, Kang DH, Johnson RJ. Uric Acid and Cardiovascular Risk. N Engl J Med. 2008; 359(17):1811–21. https://doi.org/10.1056/NEJMra0800885 PMID: 18946066

3. Borghi C, Rosei EA, Bardin T, Dawson J, Dominiczak A, Kielstein JT, et al. Serum uric acid and the risk of cardiovascular and renal disease. J Hypertens. 2015; 33(9):1729–41;. https://doi.org/10.1097/HJH.0000000000000701 PMID: 26136207

4. Mazzali M, Kanbay M, Segal MS, Shafiu M, Jalal D, Feig DI, et al. Uric acid and hypertension: cause or effect? Curr Rheumatol Rep. 2010; 12(2):108–17. https://doi.org/10.1007/s11926-010-0094-1 PMID: 20425019

5. Emmerson BT, Nagel SL, Duffy DL, Martin NG. Genetic control of the renal clearance of urate: a study of twins. Ann Rheum Dis. 1992; 51(3):375–7. https://doi.org/10.1136/ard.51.3.375 PMID: 1575585

6. Wilk JB, Djousse L, Borecki I, Atwood LD, Hunt SC, Rich SS, et al. Segregation analysis of serum uric acid in the NHLBI Family Heart Study. Hum Genet. 2000; 106(3):355–9. https://doi.org/10.1007/s004390051050 PMID: 10798367

7. Kolz M, Johnson T, Sanna S, Teumer A, Vitart V, Perola M, et al. Meta-analysis of 28,141 individuals identifies common variants within five new loci that influence uric acid concentrations. PLoS Genet. 2009; 5(6):e1000504. https://doi.org/10.1371/journal.pgen.1000504 PMID: 19503597

8. Kottgen A, Albrecht E, Teumer A, Vitart V, Krumsiek J, Hundertmark C, et al. Genome-wide association analyses identify 18 new loci associated with serum urate concentrations. Nat Genet. 2013; 45(2):145–54. https://doi.org/10.1038/ng.2500 PMID: 23263486

9. Dehghan A, Kottgen A, Yang Q, Hwang SJ, Kao WL, Rivadeneira F, et al. Association of three genetic loci with uric acid concentration and risk of gout: a genome-wide association study. Lancet. 2008; 372 (9654):1953–61. https://doi.org/10.1016/S0140-6736(08)61343-4 PMID: 18834626

10. Vitart V, Rudan I, Hayward C, Gray NK, Floyd J, Palmer CN, et al. SLC2A9 is a newly identified urate transporter influencing serum urate concentration, urate excretion and gout. Nat Genet. 2008; 40 (4):437–42. https://doi.org/10.1038/ng.106 PMID: 18327257

11. White J, Sofat R, Hemani G, Shah T, Engmann J, Dale C, et al. Plasma urate concentration and risk of coronary heart disease: a Mendelian randomisation analysis. Lancet Diabetes Endocrinol. 2016; 4 (4):327–36. https://doi.org/10.1016/S2213-8587(15)00386-1 PMID: 26781229

12. Palmer TM, Nordestgaard BG, Benn M, Tybjaerg-Hansen A, Davey Smith G, Lawlor DA, et al. Association of plasma uric acid with ischaemic heart disease and blood pressure: mendelian randomisation analysis of two large cohorts. BMJ. 2013; 347:f4262. https://doi.org/10.1136/bmj.f4262 PMID: 23869090

13. Kleber ME, Delgado G, Grammer TB, Silbernagel G, Huang J, Kramer BK, et al. Uric Acid and Cardiovascular Events: A Mendelian Randomization Study. J Am Soc Nephrol. 2015; 26(11):2831–8. https://doi.org/10.1681/ASN.2014070660 PMID: 25788527

14. Jordan DM, Choi HK, Verbanck M, Topless R, Won HH, Nadkarni G, et al. No causal effects of serum urate levels on the risk of chronic kidney disease: A Mendelian randomization study. PLoS Med. 2019; 16(1):e1002725. https://doi.org/10.1371/journal.pmed.1002725 PMID: 30645594

15. Li X, Meng X, Spiliopoulou A, Timofeeva M, Wei WQ, Gifford A, et al. MR-PheWAS: exploring the causal effect of SUA level on multiple disease outcomes by using genetic instruments in UK Biobank. Ann Rheum Dis. 2018; 77(7): 1039–47. https://doi.org/10.1136/annrheumdis-2017-212534 PMID: 29437585

16. Cortes A, Dendrou CA, Motyer A, Jostins L, Vukcevic D, Dilthey A, et al. Bayesian analysis of genetic association across tree-structured routine healthcare data in the UK Biobank. Nat Genet. 2017; 49 (9):1311–8. https://doi.org/10.1038/ng.3926 PMID: 28759005

17. Denny JC, Bastarache L, Ritchie MD, Carroll RJ, Zink R, Mosley JD, et al. Systematic comparison of phenome-wide association study of electronic medical record data and genome-wide association study data. Nat Biotechnol. 2013; 31(12):1102–10. https://doi.org/10.1038/nbt.2749 PMID: 24270849

18. Benjamini Yoav HY. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc. 1995; 72(4):405–16.

19. Hemani G, Zheng J, Wade KH, Laurin C, Elsworth B, Burgess S, et al. MR-Base: a platform for systematic causal inference across the phenome using billions of genetic associations. bioRxiv 078972 [preprint]. 2016; https://doi.org/10.1101/078972.

20. Hemani G, Bowden J, Haycock P, Zheng J, Davis O, Flach P, et al. Automating Mendelian randomization through machine learning to construct a putative causal map of the human phenome. bioRxiv 173682 [preprint]. 2017; https://doi.org/10.1101/173682.

21.   Feig DI, Soletsky B, Johnson RJ. Effect of allopurinol on blood pressure of adolescents with newly diagnosed essential hypertension: a randomized trial. JAMA. 2008; 300:924–32. https://doi.org/10.1001/jama.300.8.924 PMID: 18728266

22.   Wang J, Qin T, Chen J, Li Y, Wang L, Huang H, et al. Hyperuricemia and risk of incident hypertension: a systematic review and meta-analysis of observational studies. PLoS ONE. 2014; 9(12):e114259. https://doi.org/10.1371/journal.pone.0114259 PMID: 25437867

23.   Billiet L, Doaty S, Katz JD, Velasquez MT. Review of hyperuricemia as new marker for metabolic syndrome. ISRN Rheumatol. 2014; 2014:1–7.

24.   Peng TC, Wang CC, Kao TW, Chan JY, Yang YH, Chang YW, et al. Relationship between hyperuricemia and lipid profiles in US adults. Biomed Res Int. 2015; 2015:1–7.

25.   Li C, Hsieh MC, Chang SJ. Metabolic syndrome, diabetes, and hyperuricemia. Curr Opin Rheumatol. 2013; 25(2):210–6. https://doi.org/10.1097/BOR.0b013e32835d951e PMID: 23370374

26.   Raimondo A, Rees MG, Gloyn AL. Glucokinase regulatory protein: complexity at the crossroads of triglyceride and glucose metabolism. Curr Opin Lipidol. 2015; 26(2):88–95. https://doi.org/10.1097/MOL.0000000000000155 PMID: 25692341

27.   Bi M, Kao WH, Boerwinkle E, Hoogeveen RC, Rasmussen-Torvik LJ, Astor BC, et al. Association of rs780094 in GCKR with metabolic traits and incident diabetes and cardiovascular disease: the ARIC Study. PLoS ONE. 2010; 5(7), e11690. https://doi.org/10.1371/journal.pone.0011690 PMID: 20661421

28.   Mazharian A, Mori J, Wang YJ, Heising S, Neel BG, Watson SP, et al. Megakaryocyte-specific deletion of the protein-tyrosine phosphatases Shp1 and Shp2 causes abnormal megakaryocyte development, platelet production, and function. Blood. 2013; 121(20):4205–20. https://doi.org/10.1182/blood-2012-08-449272 PMID: 23509158

29.   Combe B, Landewe R, Daien CI, Hua C, Aletaha D, Alvaro-Gracia JM, et al. 2016 update of the EULAR recommendations for the management of early arthritis. Ann Rheum Dis. 2017; 76(6):948–59. https://doi.org/10.1136/annrheumdis-2016-210602 PMID: 27979873