Is serum urate causally associated with incident cardiovascular disease?

Simon Thornley¹, Roger J. Marshall¹, Rod Jackson¹, Dudley Gentles¹, Nicola Dalbeth², Sue Crengle³, Andrew Kerr¹,⁴ and Sue Wells¹

Abstract

Objective. With studies reporting both positive and negative associations, the influence of serum urate on incident cardiovascular disease (CVD) is uncertain. We sought to determine whether serum urate is causally associated with incident CVD.

Methods. Participants were aged 30–80 years and were screened for CVD risk in primary care between 2006 and 2009. Participants had blood pressure, lipids, age and ethnic group recorded at assessment, with record linkage providing drug dispensing, hospital diagnoses and laboratory test results. Outcomes were derived from hospital diagnoses and mortality records until December 2009. Cox models were used to assess the influence of exposures on outcomes.

Results. A total of 78,707 people, free of CVD, were enrolled, and 1,328 CVD events occurred during follow-up. Serum urate was recorded before baseline assessment in 43% (34,008/78,707) of participants. After adjustment for confounding factors, a 2 s.d. difference in serum urate (0.45 vs 0.27 mmol/l) was associated with a hazard ratio (HR) of 1.56 (95% CI 1.32, 1.84). This was more than double that of the equivalent distributional change in high-density lipoprotein cholesterol (adjusted HR 1.22) and one-third greater than that for HbA1c (adjusted HR 1.41).

Conclusion. Serum urate is likely to be causally associated with CVD. This supports public health action to reduce urate levels in populations with significant burdens of the disease.

Key words: cardiovascular diseases, risk factors, epidemiology.

Introduction

‘Measurement of serum uric acid levels is unlikely to enhance usefully the prediction of [coronary heart disease] CHD, and this factor is unlikely to be a major determinant of the disease in general populations’ [1]. ‘These results showed that hyperuricaemia has a strong association with…death in all causes, coronary heart disease,…and indicated that serum uric acid seems to be a considerable risk factor for reduced life expectancy’ [2]. Uncertainty characterizes research on the effect of serum urate on the incidence of cardiovascular disease (CVD), reflected in these two contrasting quotes from European and Japanese researchers, respectively. Meta-analyses of observational studies [3, 4] have reported some association between hyperuricaemia (serum urate >400 μmol/l) and incident CHD. One study [3], for example, reported an increase in risk of CVD (pooled relative risk 1.46; 95% CI 1.20, 1.73), but the association diminished with adjustment for established risk factors (pooled relative risk 1.09; 95% CI 1.03, 1.16). Such factors included age, systolic blood pressure, cholesterol, smoking and either a diagnosis or other laboratory indices of diabetes. However, generally it is not clear which factors to adjust for. Some, such as systolic blood pressure, may be on the causal pathway between serum urate and CVD, thus mediating the relationship rather than acting as a confounder, since there is evidence that high levels of serum urate causally influence the onset of hypertension [4, 5].
Thus whether serum urate is causally associated with CVD remains uncertain.

The prevalence of gout is extremely high, by international standards, among Māori (~6%) and Pacific people (~8%) living in New Zealand [6, 7]. Together these ethnic groups comprise between 20% and 25% of the total population, so serum urate is measured frequently in New Zealand. We have been conducting a large cohort study of CVD risk factors that involves linking risk factor data derived from primary care-based screening to laboratory results, drug dispensing, hospital admission diagnoses, and mortality records. Here we explore the relationship between urate levels and the incidence of CVD.

### Methods

#### Study design

The PREDICT cohort has been described elsewhere [8, 9]. Briefly, PREDICT is a web-based program for assessing and managing CVD risk that is integrated with commonly used general practice electronic medical records in New Zealand. PREDICT generates a standardized cardiovascular risk profile on each patient and calculates a 5-year percentage CVD risk based on the patient’s age, gender, ethnic group, blood pressure, total to high-density lipoprotein (HDL) cholesterol ratio, family history of CVD, smoking and diabetes status, using Framingham study-derived equations [10]. PREDICT clinical information is either entered by the general practitioner at the time of risk assessment (blood pressure, diabetes status, family history of CVD, smoking and family history of CVD) or derived from practice management software from information stored about the patient (age, gender, ethnic group and total to HDL cholesterol ratio). A risk profile is stored for each patient and is linked to an encrypted, unique, National Health Index. The index enables linkage with national health records, such as hospital discharge codes, mortality, community drug dispensing and laboratory test results. Enrollees between January 2006 and 15 October 2009 were selected for this analysis since full dispensing and coded mortality records were available for this period. Those aged <30 or >80 years at the time of screening were excluded. We also excluded people with a history of prior CVD or heart failure, identified by a general practitioner diagnosis of CVD or hospitalization with CVD in the past 5 years, or those dispensed a loop diuretic in the 6 months before assessment who were assumed to have heart failure.

#### Serum urate

The urate concentration of enrollees in PREDICT were obtained by anonymized linkage to community laboratory data in the Diagnostic Medlab database. The database was the sole provider of laboratory tests, outside hospital, to the greater Auckland region until late 2009. If a person had had many serum urate measurements, we used the one prior to their first PREDICT assessment. Serum urate and other laboratory variables (HbA1c and HDL cholesterol) were only recorded if a patient had had a laboratory test in the 5-year period before their PREDICT assessment; otherwise data for these variables were missing.

#### Outcomes

CVD events (acute coronary syndromes, coronary procedures, stroke, transient ischaemic attack, haemorrhage stroke, peripheral vascular disease, peripheral arterial procedures and congestive heart failure) were identified from hospital discharge and mortality records using International Classification of Disease codes, as previously described [11]. They covered the period from January 2006 to December 2009. The analysis was of the outcome of any CVD event, including immediate death that was coded as caused by such conditions.

#### Other data

Patient medications, specifically anti-hypertensive and gout medication, were obtained from the New Zealand Pharmaceutical Information Database [12], a register of community dispensing, and were merged with PREDICT records. Anti-hypertensive drugs included ß-blockers, calcium channel blockers, angiotensin converting enzyme (ACE) inhibitors, ß-blockers and thiazide diuretics. Gout medication included allopurinol, colchicine and probenecid.

Demographic information was taken from the PREDICT database and, in the case of ethnicity, the variable was recorded in both PREDICT and national data sources. Ethnicity data were classified according to a standardized data protocol [13], in which Māori and then Pacific ethnic groups are prioritized if more than one ethnic group was recorded. Individuals were coded as either Māori, Pacific, Indian or Other. The Other group were mainly of New Zealand European descent.

#### Analysis

Statistical analysis centred on developing a survival model for incidence of any CVD event, taking serum urate as a potential predictor. We used a Cox regression model, with age at event as the time variable rather than time to event from baseline enrolment. This approach is increasingly recommended for observational epidemiological analysis of CVD [14, 15]. Time, therefore, was considered as left truncated (patients are observed conditional on survival up until that point) and right censored. Proportional hazards assumptions were checked using scaled Schoenfeld residuals [16]. Restricted cubic splines were used to investigate the relationship between continuous variables and time to event as a check on the modelling assumption of linearity. HRs for continuous variables were reported by comparing the relative hazard between the 16th and 84th centiles of the variable (1 s.d. either side of the mean for a normal distribution) [17]. This allows direct comparison with relative hazards of binary variables, since the 16th and 84th centiles are roughly equivalent to the one-unit difference between 0 and 1 of a binary variable. That is, 2 s.d.s on a binary scale with a mean of ~0.5 is 1.
Fig. 1 A DAG, characterizing the influences on risk of developing CVD.

Dark font variables are those that are observed, whereas grey font variables represent those that are unobserved. 

To help decide model structure and which variables to include in a model, we sketched a diagram of likely causal mechanisms and pathways [a directed acyclic graph (DAG)] [18], and used some of the ideas proposed by Pearl to examine causality. The DAG we constructed included measured and unmeasured influences on CVD risk that we felt were important (Fig. 1). Lines with an end arrow represent a causal link acting in the direction of the arrow. Those lines that are solid represent a link that can potentially be observed from our data, dashed lines are causal mechanisms we believe to hold but cannot be observed in our data.

From Fig. 1, we considered a diagnosis of gout (derived from dispensing data) as likely to be a mediator of the influence of urate on CVD risk. We assumed that lipoprotein concentrations are causally associated with CVD, and so they are included in the model. Blood pressure (BP), measured before baseline (BP_{t-1} in Fig. 1) and measured at baseline (BP_{t}), were considered mediators of the influence of urate on CVD risk, and were therefore not controlled for in the model. Blood pressure lowering and statin therapies were considered to represent colliders [18], which may introduce bias from an unobserved variable that would otherwise not influence our analysis, so these variables were excluded. Although certain blood pressure medications, such as thiazide diuretics, are known to raise serum urate, we believe that this is a relatively small influence on the causal paths compared with those otherwise identified in the DAG. Furthermore, blood pressure and treatment of the condition were regarded as mediators (caused by high urate levels), as was a diagnosis of gout. In view of these considerations, to block backdoor paths, using Pearl’s terminology (adjust for confounding variables), the regression analysis adjusted for gender, HbA1c, ethnic group, smoking status and lipoprotein concentrations (HDL cholesterol). Because post-treatment variables are likely to be strongly influenced by baseline recordings and other unobserved influences (colliders), we believed that including such variables was more likely to result in biased effect estimates than excluding them from the analysis [19].

We also tested for interactions, in particular, between serum urate and gender. From other studies of CVD using similar risk factor profiles, effect estimates have been observed to change with age. These were tested using the likelihood ratio ($P < 0.05$). When using age at event as the time variable, such interactions may manifest as violations of the proportional hazards assumption. We tested for such violations using the cox.zph function, which calculates tests of the proportional hazards assumption for each covariate by correlating the corresponding set of scaled Schoenfeld residuals with a suitable transformation of time, based on the Kaplan–Meier estimate of the survival function [20]. Stratification allowed different baseline hazards for exposure groups, where the proportional hazards assumption was likely to be contravened.

Because many serum urate and other laboratory values were missing, we used multiple imputation, so that the modelling results reported were based on averaging model parameters over 10 random imputations [16]. Variables that were used in the multiple imputation modelling were gender, serum metabolic markers (creatinine, HDL cholesterol, triglycerides, HbA1c and urate), age at enrolment, systolic blood pressure, smoking status, diabetes, ethnic group, death and CVD event, along with an interaction between serum urate and sex.

All analyses were done using R software (version 2.14.1) [21]. The package rms [22] was used for the Cox regression analysis and the functions aregImpute and fit.mult.impute for multiple imputation. The aregImpute function finds transformations that optimize how each variable may be predicted from every other variable, using additive semiparametric models. The fit.mult.impute function was then used to average sets of regression coefficients and compute variance and covariance, adjusted for the error derived from the uncertainty from imputation of missing data. This project was approved by the national Multi Region Ethics Committee in 2007 (MEC/07/19/EXP).

Results

Table 1 shows the baseline characteristics of the PREDICT cohort, that is, when first enrolled in the PREDICT database. Men (56%) outnumbered women, and the mean age was 55 years, with women rather older. Pacific people accounted for 23% of the cohort, Māori 16% and Indian 7%. The remainder of the cohort and majority ethnic group, Other, was 86% European, 6% Chinese and 8% composed of a large variety of ethnic groups. More women than men were diagnosed with diabetes; however, HbA1c levels were equivalent. Men had higher levels of serum urate, higher lipid ratios and lower levels of HDL. Women were less likely than men to smoke cigarettes.

Urate values were approximately normally distributed, with a lower mean in women compared with men.
There were, however, 57% missing serum urate levels, 58% missing HbA1c and 13% missing HDL cholesterol. No other data items had missing values. Mean observed urate values varied substantially by ethnic group, but varied little by diabetes or smoking status (Table 2). Māori and Pacific ethnic groups were highest, whereas mean levels among the Other and Indian ethnic groups were about half to two-thirds of an S.D. lower.

A total of 1328 CVD events occurred during follow-up, with 167 patients dying during this period. Median follow-up time was 538 days, with a maximum of 1424 and a minimum of 1. The distributions of observed and imputed urate values (Fig. 2) show that the variance is reduced among imputed compared with observed values, because imputed values are regressed to the mean.

Initial model checking revealed that the proportional hazard assumption was not supported for gender and ethnicity. For this reason the model was stratified on these factors, meaning that separate baseline hazards were assumed for all permutations of gender and ethnic group. A gender by urate and ethnic group by urate interaction was tested for, but neither was found to be statistically significant.

Table 3 shows the results of the Cox regression analysis. In the model, comparing measures at the 16th and 84th centiles showed an HR of 1.56 (95% CI 1.32, 1.84) for serum urate. The relative hazard was linear (on the logarithmic scale), which indicated that an increase in serum urate of 0.29 mmol/l \[=\ln(2)/(\ln(urate))\] doubled the relative hazard of CVD across the distribution of observed urate values. In the study population, the association between a 2 S.D. difference in urate level (67% relative increase) on incident CVD risk was higher than adjusted effects of the
equivalent change in the distribution of HbA1c (37% relative increase) and HDL cholesterol (155% relative increase).

An illustration of the extent to which a change in serum urate affects survival is shown in Fig. 3. It gives the Cox-modelled survival plot for a man in the Other ethnic group, with HDL cholesterol = 1.34 mmol/l and HbA1c = 6%.

Cox-modelled survival estimates (with 95% confidence intervals) by time for serum urate at the 16th centile (0.27 mmol/l) compared with the 84th centile (0.45 mmol/l) for a non-smoking male in the Other ethnic group, with HDL cholesterol = 1.34 mmol/l and HbA1c = 6%.

An illustration of the extent to which a change in serum urate affects survival is shown in Fig. 3. It gives the Cox-modelled survival plot for a man in the Other ethnic group, with HDL cholesterol of 1.34 mmol/l, HbA1c of 6% and serum urate of 0.27 mmol/l, the 5-year risk of CVD equates to $1 - (0.813/0.871) = 6.7\%$. If his serum urate is about 2 S.D. greater, that is 0.45 mmol/l, his 5-year cumulative incidence of CVD will increase to 10.1% (about a one-third increase). This calculation is based on an estimate of the baseline survival ($S_0$) of 88.9% at 60 years and 83.8% at 65 years.

**Discussion**

We found that, given plausible assumptions, raised serum urate is likely to have a substantial causal effect on the incidence of CVD, of a similar or greater magnitude to the effects of high HbA1c or low HDL cholesterol levels. A strength of the study was that it was based on a large cohort, $n = 79,707$, in which the information had been collected in a systematic and standardized way. An associated weakness, however, is the incompleteness of the laboratory data, and in particular of the serum urate values. To deal with these missing values, we used multiple imputation, which is recommended to reduce bias in effect estimates when compared with complete-case analysis [16]. In a sensitivity analysis, using complete-case analysis, our observed effect estimates were similar to those derived from multiple imputation (the adjusted HR of the effect of urate was 1.59 (95% CI 1.32, 1.90), compared with 1.56 in Table 2]. We can only speculate that if serum urate had been measured in everyone, the effect estimates would not have differed substantially.

We have also explored sensitivity analyses of multivariate effect estimates (see supplementary data, available at *Rheumatology* Online). These included carrying out separate analyses by gender, with and without imputation of missing values, varying the time scale chosen, including those who had used loop diuretics at baseline assessment, and dividing urate values by quintile. Generally these analyses were congruent with the effect estimates reported in Table 3.

A positive aspect of the study was that we explicitly considered causal paths to determine variables to adjust in the model [18]. This process is conceptually appealing and less likely to underestimate the effect of serum urate.

---

**Table 3** Measures of association with CVD from a Cox proportional hazards model

<table>
<thead>
<tr>
<th>Factor</th>
<th>Comparisona</th>
<th>Crude HRb (95% CI)</th>
<th>Adjusted HRc (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serum urate, mmol/l</td>
<td>High</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>0.45</td>
<td>0.27</td>
<td>1.76 (1.56, 1.99)</td>
<td>1.56 (1.32, 1.84)</td>
</tr>
<tr>
<td>Non-smoker</td>
<td>1.00</td>
<td>0.00</td>
<td>1.71 (1.50, 1.95)</td>
</tr>
<tr>
<td>HbA1c, %</td>
<td>8.00</td>
<td>5.50</td>
<td>1.41 (1.29, 1.54)</td>
</tr>
<tr>
<td>Serum HDL, mmol/l</td>
<td>1.03</td>
<td>1.79</td>
<td>1.75 (1.49, 2.04)</td>
</tr>
</tbody>
</table>

$n = 78,707$. aComparisons for the continuous variables are taken at the 84th and 16th centiles ($z = 1$) for comparability with binary variables. The binary variables are, by definition, set at 0 and 1. bStratified by sex and ethnic group. cAdjusted for all variables in the table and stratified by gender and ethnic group.
or to introduce inappropriate confounder adjustments. Whether our DAG is an accurate reflection of the causal interaction between the variables in our study may be debated; for example, whether systolic blood pressure is a mediator of the effect of urate or whether it should be included as a confounder in regression equations. To explore the effects of variable selection based on the DAG, we carried out a sensitivity analysis, altering the model by (i) adding a diagnosis of diabetes in the model; (ii) inter-changing HDL cholesterol with the total cholesterol/HDL ratio and (iii) adding blood pressure as a potential con-founder. The first two changes to the model resulted in a <2% change in the adjusted HR for serum urate. After inclusion of diabetes, the effect of HbA1c was reduced from 1.41 to 1.36 (comparing individuals with baseline measures of 8.00 and 5.50%). If systolic blood pressure is included, our adjusted effect of urate diminished slightly [for 0.45 compared with 0.27 mmol/l; the HR reduced from 1.56 (95% CI 1.32, 1.84) to HR = 1.48; (95% CI 1.25, 1.75)].

Family history of ischaemic heart disease may also be considered a confounder of the relationship between urate and CVD. For this study, family history of CVD was recorded as a history of CHD or ischaemic stroke in a first-degree relative (father or brother <55 years, mother or sister <65 years). When this variable was added, however, the point and interval estimate for serum urate remained unchanged (see supplementary data, available at Rheumatology Online).

The nature of the cohort, composed of a generally high CVD risk population rather than a designed sample, is also a potential weakness. The high prevalence of diabetes (~20%) reported among the cohort along with the gender distribution favouring men was expected. Nevertheless, there is substantial heterogeneity in the risk profile of the study population, so there is no reason to believe that the lack of representativeness was an important problem for these analyses.

Using age as the outcome variable in the Cox model imposes an assumption that entry into the study is independent of age. This is unlikely to be true here, given that national CVD screening guidelines recommend initiation of screening at specified ages, which differ for varying ethnic groups. However, use of age in survival modelling is now the recommended approach for cohort studies that estimate CVD incidence [15]. This choice of time scale is justified because the use of age at event results in less bias of effect estimates compared with time to event [16]. This method is increasingly recommended if the period of observation is the time from birth to event, as it is when modelling CVD risk, rather than time from intervention to event, as it is in a randomized trial [15].

There are some differences and similarities between our results and other studies. Some studies report the association between raised serum urate in a binary fashion (presence or absence of hyperuricaemia) [3]. In our study we found evidence of a log-linear relationship between urate concentration and CVD risk, so studies that employed a cut-off are likely to underestimate the effect of the exposure. Some studies have reported effects of serum urate on CVD as a continuous variable. For example, the Framingham study [23] (n = 6763) found no significant association between serum urate and CHD incidence in their adjusted Cox model among men, but a significant increase in risk with women. We found a statistically significant association between urate and CVD, without a gender × urate interaction, and so our measures of association were aggregated over the whole population. Our study reached similar conclusions to the National Health and Nutrition Examination Survey (NHANES)-I [24] cohort study (n = 5926), which found that serum uric acid levels were significantly associated with CVD mortality after adjustment for known confounders. The NHANES study used CVD mortality as its outcome event while our study used hospital admission or death from CVD. Its advantages were that it had long follow-up time (16.4 years) and included both white and black participants (not present in the Framingham cohort). The authors reported age-adjusted HRs for cardiovascular mortality in participants aged 45–54 years that were similar to ours: 1.28 (95% CI 1.08, 1.52) in men and 1.43 (95% CI 1.16, 1.37) in women for each 0.06 mmol/l rise in serum urate level. Our study, using age as the time-to-event variable, is likely to adjust for the effect of age more accurately than the more traditional methods used in the NHANES study, which aggregated age into 15-year intervals. Despite these differences in methods, this study was concordant with our findings.

Our study provides statistical evidence that elevated serum urate is likely to cause CVD. Biological evidence also supports our findings, highlighting the role of serum urate in the pathogenesis of endothelial injury [25]. If this causal relationship is true, attention turns to what determines serum urate, and how serum urate may be lowered, either by lifestyle modification, or through drug treatment. In addition to the role of purine-rich foods, fructose intake may also contribute to hyperuricaemia through increased hepatic synthesis of purines [26, 27]. This gives impetus for public health action to limit intake of sugar, the dominant source of fructose in the diet, as well as purine-rich foods. In addition, this finding raises the question of whether pharmaceutical reduction of urate levels, with urate-lowering therapy such as allopurinol, can reduce CVD risk. Such an idea is strengthened by the results of a small (n = 65) cross-over trial of allopurinol therapy in patients with symptomatic, effort-dependent angina. Treatment with 600 mg of allopurinol per day for 6 weeks increased the time to S-T segment depression by 43 s (95% CI 31, 58) compared with otherwise equivalent treatment with placebo [28].

Māori and Pacific people in this study had mean levels of serum urate ~15% higher than Others, Indians and other Asians. We speculate that dietary rather than genetic differences are more likely to account for these differences, because New Zealand Māori almost all share significant ancestry with New Zealand Europeans due to widespread intermarriage, yet their mean urate values vary substantially. In our regression analysis, to obtain a
model with good fit, we stratified by ethnicity rather than evaluating the effects of this variable. However, because CVD incidence and prevalence are significantly higher in Māori and Pacific peoples compared with Europeans [29–31], future work will evaluate whether variations in serum urate levels account for ethnic variations in the prevalence of CVD. In conclusion, serum urate is associated with incident CVD and is likely to be a significant factor in the disease’s causal pathway.

Rheumatology key messages

- Substantial uncertainty exists about the effect of serum urate on CVD.
- After making causal assumptions, we found a strong association between urate and CVD.
- Urate differs substantially between ethnic groups, and may explain differences in CVD burden.

Acknowledgements

The authors would like to thank the National Health Board Analytic Services and Diagnostic Medlab Ltd for supplying anonymized data. They would also like to thank affiliated general practitioners and practice nurses and patients belonging to the ProCare Network, Auckland PHO Ltd, HealthWest, East Health Services, National Māori Hauora Coalition, Alliance, Te Tai Tokerau and Manaia PHOs.

PREDICT was developed by a collaboration of clinical epidemiologists at the University of Auckland, IT specialists at Enigma Publishing Ltd (a private provider of online health knowledge systems), primary health care organizations, non-governmental organizations (New Zealand Guidelines Group, National Heart Foundation, Diabetes New Zealand, Diabetes Auckland), several district health boards and the Ministry of Health. The PREDICT software platform is owned by Enigma Publishing Ltd (PREDICT is a trademark of Enigma Publishing Ltd).

Funding: This research was supported by a grant from the National Heart Foundation of New Zealand and a Health Research Council Clinical Research Training Fellowship (grant number 11/145).

Disclosure statement: N.D. has received consultancy fees from Takeda, Ardea Biosciences, Metabolex and Novartis. All other authors have declared no conflicts of interest.

Supplementary data

Supplementary data are available at Rheumatology Online.

References

5 Feig DI, Soletsky B, Johnson RJ. Effect of allopurinol on blood pressure of adolescents with newly diagnosed essential hypertension. JAMA 2008;300:924–32.
19 Gelman A, Hill J. Causal inference using regression on the treatment variable. In: Data analysis using regression and


