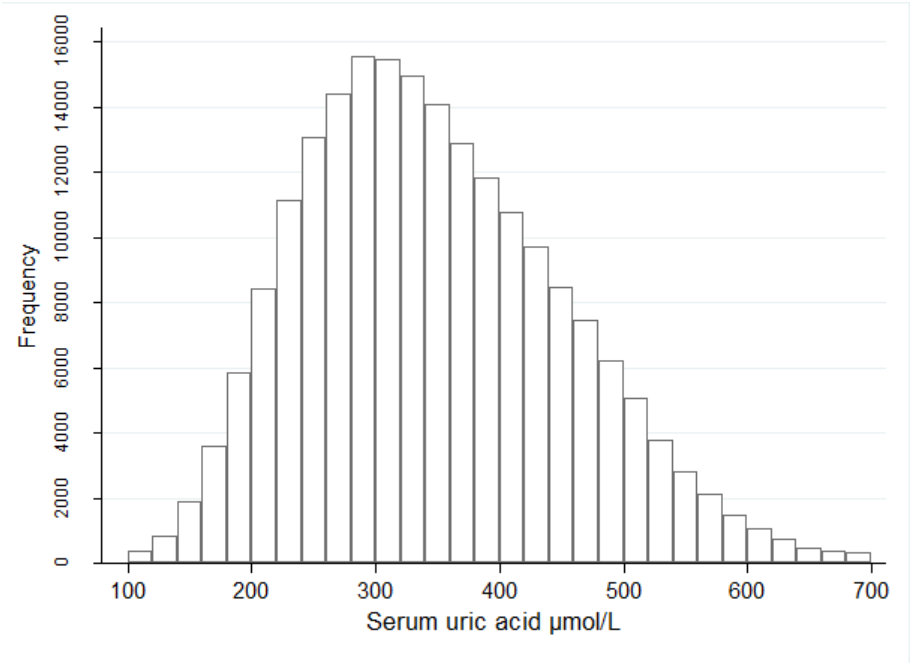
Supplementary Information

In addition to the potential confounders included in the primary models, we also examined the effect of including other variables that may influence uric acid levels or are important predictors of disease events of interest based on the literature or biological plausibility. A creatinine measure is often ordered by the GP at the same time as uric acid and therefore we also extracted these test results to control for any subclinical renal disease not captured by diagnostic Read codes. Oral corticosteroids are prescribed for certain respiratory conditions and can moderately lower uric acid. It is difficult to determine whether such prescriptions meet the classical definition of a confounder or play a role on the causal pathway. Similarly, raised antioxidant capacity may protect against a range of diseases and thus co-morbidity may not be a true confounder. As a sensitivity analysis, we identified prescriptions in the year prior to the uric acid test for oral corticosteroids, cardioprotective drugs that can influence uric acid levels (diuretics, aspirin, antihypertensives, beta-blockers), and uric acid-lowering drugs used to treat gout. Patients with prevalent type II diabetes and cardiovascular disease were also identified using Read codes and relevant prescriptions (e.g. insulin). There was no pre-existing evidence that statins or inhalers directly affect uric acid but due to the high frequency of prescription in the UK and potential relationships with outcomes, we decided to investigate these drugs in the regression models. Including these additional variables in the models had no meaningful impact on the overall results (Supplementary Figure S2). Adjusting for prevalent COPD had very little impact on the associations for lung cancer.

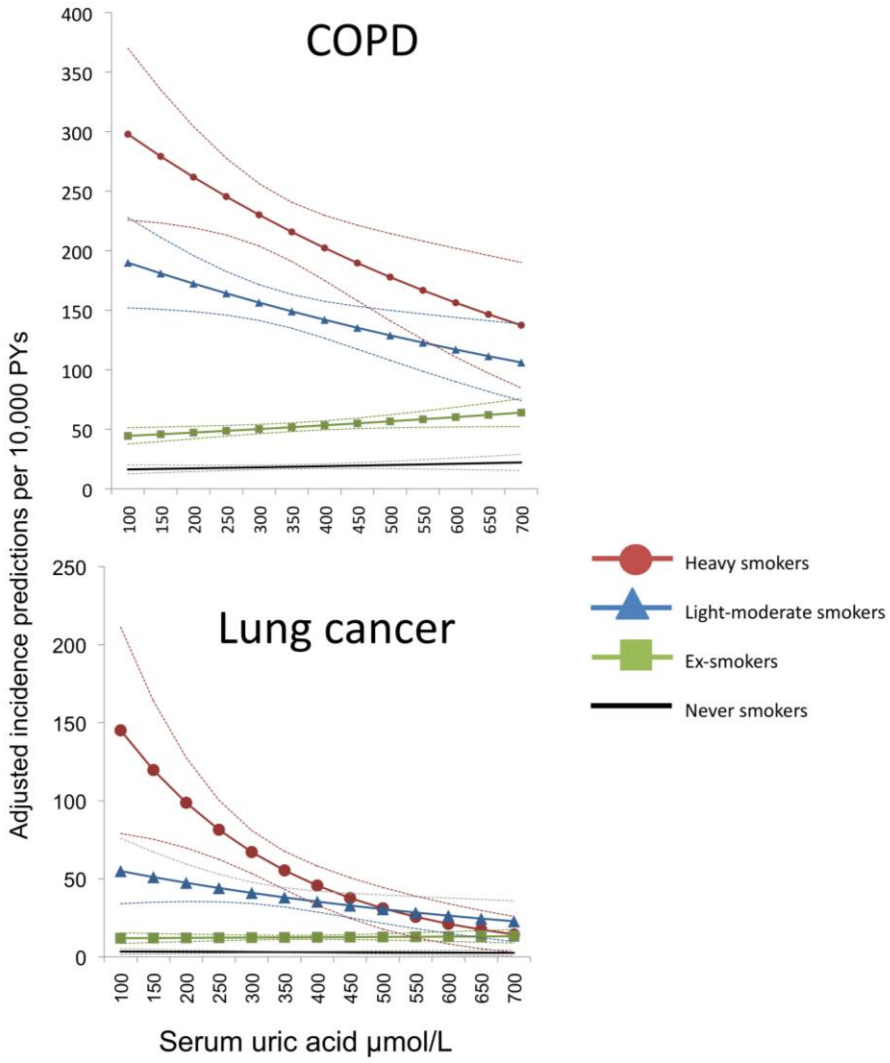
Adjusting for smoking intensity (number of cigarettes per day) as a continuous variable within the current smoking strata to account for any residual confounding also had no meaningful effect on the results for any disease events.

Using cubic spline transformations of uric acid levels (3, 4 and 5 knots at tertile, quartile and quintile values of serum uric acid) and comparing these models to those assuming simple linear relationships using the Akaike Information Criterion showed no evidence that non-linear functions better described the relationships with disease incidence ($p>0.05$).

There is some evidence that uric acid levels may increase in response to acute events such as exacerbation of COPD. To investigate this possibility in a supportive analysis, we plotted mean uric acid levels over time leading up to and following the diagnosis of event of interest. There was no evidence that uric acid was changing as a function of time to or since diagnosis and thus we assumed any effects of reverse causation to be negligible.



Supplementary Figure S1: Histogram showing the distribution of serum uric acid levels in a cohort of patients from the Health Improvement Network database. Conversion factor: µmol/L x 0.01681 mg/dL.



Supplementary Figure S2: Adjusted incidence predictions for chronic obstructive pulmonary disease (COPD) and lung cancer as a function of serum uric acid levels in a cohort of primary care patients showing an interaction with smoking status.

Adjusted for age, sex, social deprivation, year of uric acid test, alcohol, height, weight, blood pressure, ethnicity, area pollution, creatinine levels, comorbidity for cardiovascular disease or diabetes, prescriptions for oral corticosteroids, inhaler medication, cardioprotective drugs, and uric acid lowering agents. Conversion factor: µmol/L x 0.01681 mg/dL.