Responses to editorial committee comments. Comments / queries in <> with our response immediately underneath. Please note that we have added other BMJ requirements (What this paper adds / Ethical improvement and informed consent / Copyright statement / Patient involvement statement).

We carefully reviewed the Methods and have simplified the description:
1. Deleted instances of un-necessary detail (e.g. genotyping platforms for the various cohorts) and un-necessary detail(s) (e.g., no need to state that separate male-female analyses were done).
2. Shifted description of calculation of partial R2 values to the legend of Table 1.
3. Shifted description of imputation of genetic variants in the genetic risk score to the legend of Table 2.
4. Shifted the description of study power to legend of Figure S4 and beginning of Results.
5. Various other edits.

Regarding lack of clarity as to where results mentioned in text were appearing in Tables we have:
1. Reduced the cross-referencing to Supplementary Tables
2. Increased referencing to main text Tables, of which there are only two.
3. We have also reformatted Table 2.
4. The Supplemental Tables are now in a multi-sheet Excel file, in a more presentable form.

Regarding methods, in response to reviewers we have added two further dietary scores to the revised manuscript (Mediterranean and data-driven dietary scores). This has necessitated addition of two further paragraphs to the Methods. While these paragraphs are not necessarily tricky to follow (aside perhaps from some of the data-driven methodology), they do add to the length of the manuscript. While we would prefer to leave the dietary score construction methodology in the main manuscript, we are very prepared to take editorial advice on this.

We hope that the methodology is easier to follow as a result of our response to Dr Kirkham (above).

People with gout were excluded from the analysis (refer Fig S1), from the ARIC, CARDIA, FHS and NHANES III cohorts. However gout was not ascertained from the CHS cohort, meaning that there will be some people with gout included in the analysis (~63 assuming a 3% prevalence). Some of these patients may have been following a diet but, as a proportion of the overall cohort, it will represent a fraction of a percent.

The study sample was selected for a variety of criteria (Fig S1 and described in the second paragraph of the Methods; e.g. European ancestry, not having gout, not on urate-lowering therapy, not taking diuretics). Given the gout exclusion we had previously included a statement (last sentence of the penultimate paragraph of the Discussion) that these data couldn't be generalized to people with gout. We have now shifted this statement to the final paragraph of the Discussion, for better emphasis.

While we didn't address the dietary advice to gout patients issue in the manuscript, given that gout patients were excluded, we note that there are no robust data demonstrating effectiveness of dietary advice for urate-lowering in the management of gout in the primary care setting (Holland and McGill "Comprehensive dietary education in treated gout patients does not further improve serum urate" 2015 PMID 25495503 & Moi et al. Cochrane review "Lifestyle interventions for chronic gout" 2013 PMID 23728699). We did discuss in paragraph 5 of the Discussion that clinical trials demonstrate the DASH diet to be effective in lowering urate, however the barriers to implementing this diet both at the population level and primary care setting are yet to be overcome. There is a growing school of thought within the rheumatology setting that time spent on giving dietary advice in the primary care setting is

time taken away from establishing patients on urate-lowering therapy, an effective means of lowering urate.

We reordered the items in the table according to the strength of effect of urate-associated foods. For the full cohort we also dedicated a single column to the beta value and bolded the beta values. We also added this sentence to the second paragraph of the Results (page 12): "The food with the strongest urate-raising effect (shellfish) associated with a 2.49 μmol/L increase in serum urate per serving per week, equating to a 17.43 μmol/L (0.28 mg/dL) increase per daily serving."

That diet plays only a minor role is not the widespread perception of the general public or most medical professionals who view hyperuricaemia and gout as conditions of dietary excess, and often believe that serum urate levels can be reduced substantially by diet management. Focus on diet also reinforces negative beliefs about hyperuricemia and gout that can be stigmatizing (Duyck et al. ""You don't have to be a drinker to get gout, but it helps"": a content analysis of the depiction of gout in popular newspapers". 2016 PMID 27134185). We did not directly address this, or the below, editorial comment in the revised paper because our study focused on urate control in the non-gout population.

As outlined above, we believe our study is important because it challenges the wide-held belief that diet plays a central role in regulating serum urate levels. There are widely held beliefs within the community, which often lead to people with hyperuricaemia feeling blamed and stigmatised (Chandratre et al. ""You want to get on with the rest of your life": a qualitative study of health-related quality of life in gout" 2016 PMID 26245722), or receiving advice which is unlikely to alter the urate concentration.

We have now included a Patient Involvement Statement. While we did not directly involve patients in the writing of the article ND and TRM have been involved in the Auckland-based Counties Manukau District Health Board Maaori Gout Action Group (Winnard et al. "Debunking the myths to provide 21st Century management of gout." 2008; PMID 18535649) which informs our wider research program in control and management of serum urate levels and gout. TJM (ne Flynn) has done qualitative research with gout patients about beliefs of their causes of gout, leading to a diet-related scientific publication (Flynn et al. "Positive association of tomato consumption with serum urate levels: support for tomato consumption as an anecdotal trigger of gout flares" 2015 PMID 26286027) that strongly stimulated the diet-wide association study currently under review.
PIS: "No patients were asked for input in creation of this article. ND and TRM are founding members of the Auckland-based Counties Manukau District Health Board Maaori Gout Action Group [73] that has identified that "to achieve modern management of gout, those with gout need to be supported by primary care practitioners who are aware of the need for early intervention with allopurinol, as well as whaanau/families and communities who understand the impact and causes of gout" [73]. This group includes patient advocacy from Arthritis New Zealand. The views of the Maaori Gout Action Group inform our research into the causes and management of gout and have influenced our design and conclusions of this study. In qualitative research with gout patients during her PhD research at the University of Otago (awarded 2016) TJM was informed by the lived experiences of gout patients that led to publication of the role of tomato consumption as a trigger of gout [74]. This work was a key genesis for the work presented here."

Responses to reviewers. Comments / queries in <> with our response immediately underneath.

Reviewer 1

It is important to note that that the analyses presented in Table 2 had a consistent number of participants (n=16,760), with the dietary score analyses missing (at most) 108 participants (0.6%), meaning that any non-random missingness will be a small effect. Please also refer to Figure S1 which describes the exclusion criteria that were applied consistently to the various cohorts, with the resultant 16,760 participants used in all the analyses. We have added a line to this figure for the exclusion of participants who did not have data for all the covariate variables. This includes 1,348 participants for whom menopausal status was unknown, along with 240 participants who did not have data for the additional adjustors added in response to reviewer 2. The text has been edited accordingly, as have the cohort summary statistics presented in Tables S1 and S4 and Figures S2, S3, S4 and S6.

For some of the food-specific analyses presented in Table 1 and Tables S5-S7, an entire cohort would require exclusion, dependant on the nature of the food frequency questions asked in that specific study. Thus this extra exclusion of participants is non-random. This occurred for 26 of the 63 food items (cream, flavoured milk, candy, apple/pear, banana, melon, peach, cooked cereal, rice, nuts, peanuts, bacon, hamburger, hotdog, sausage/lunch meat, cabbage/cauliflower, broccoli, carrot, corn, lettuce, peas, string beans, creamer, mayonnaise/dressing, condiments, and table sugar). Differences in specific questions asked did not have an effect on the construction of the diet quality scores, as in general these were concerned with consumption of broader food groups, and as such all 16,760 participants were included in the diet quality score analyses.
To clarify we have added these sentences / phrases to the Methods section: "If an identical question could not be created the non-matching information was excluded, either from only the cohort with non-matching data (eg. NHANES III asked about consumption of peanuts, peanut butter, nuts, and seeds in a single question, making this non-comparable to either the nuts or peanuts questions of the other four studies) ..." (page 7) and we have also now added 'n' to all relevant analysis tables (Tables 1-2 and Tables S5-S7).

The purpose was primarily to account for cryptic relatedness which could lead to shared diet and less to correct for population stratification (as the study was performed in European only). We had mentioned population stratification in the Methods, and have now edited the sentence (page 7) to read: "Analyses in ARIC, CARDIA, CHS, and FHS were additionally adjusted for whole-genome principal component vectors one to four to account for cryptic relatedness (especially within FHS) that may cause inflation of test statistics owing to possible shared diet or heritability of serum urate levels."

We are not fully certain of the reviewer's point here. Perhaps the reviewer would like to see the SNPs analysed individually and then the individual variances explained summed? We had previously done this but not presented the data. The summed individual variances explain 8.7% of variance in serum urate levels, compared to 7.9% explained by the weighted genetic risk score. We have now included the data as a Supplemental table (Table S8) in the revised manuscript.

This is a good suggestion. We have now tested for interaction between dietary score and genetic risk score in determining serum urate levels with no evidence for interaction. The data are included in Table S9 and we added this text as the last sentence of the Results: "There was evidence for interaction between any of the four diet scores and the weighted genetic risk score only for the DASH diet in the female-only cohort (P=0.04), for all others P was $P \geq 0.21$ (Table S9)."

Reviewer 2

We have removed reference to 'healthy diet' in the abstract, replacing with 'diet quality'. We also altered the first sentence of the abstract to reflect that we evaluated the relative contributions of diet and genetics: "To systematically test dietary components for association with serum urate and to evaluate the relative contributions of percent variance in serum urate explained by estimates of diet quality and inherited genetic variants."

The five studies did not collect consistent information about portion size, ARIC and FHS both specified portion in weight per food item (eg. whole milk, 8 oz. glass), CARDIA allowed participants to specify the usual portion they ate of each food item ("how much do you usually have?"), CHS interviewers asked participants to estimate whether they usually have a small, medium or large portion per item (eg. a medium apple was ~2 ½" in diameter), and NHANES III did not ask about portion size. Thus we did not consider portion size in our study. This was mentioned in the footnote of Table S2, but we have also added a sentence to the methods section (page 5):"Average consumption was not able to be adjusted for portion size in the aggregated data, as the NHANES III study did not specify portion size, the CHS study only specified a relative portion size (small / medium / large), and the portion sizes specified by the ARIC, CARDIA, and FHS studies were not consistent."

For the purposes of our regression analyses whether analyzing as per day or per week makes no difference to the P-values supporting the final outcome measures (the beta value for an analysis of serves per day is 7-times that of a serves per week analysis). We analysed the data in serves per week as both the ARIC and FHS cohorts had suggested conversion values in serves per week for the questionnaire categories they used, and the CARDIA data were supplied as serves per week, converted by the CARDIA study investigators from an open-ended frequency question per food.

We have constructed a Mediterranean diet score based on Panagiotakos et al (2006) and tested for association with serum urate levels, these data have been added to Table 2 and we have also added a new Supplementary Figure (S4). Similar to the Healthy Eating and DASH diet scores the Mediterranean diet score associates strongly with serum urate levels but explains very little variance. We added the Kontogianni publication to the Introduction and added additional text describing construction of the Mediterranean diet score to the Methods (pp 8-9).

We constructed the Harvard 'Healthy eating' score according to the Harvard recommendations and nescient to the foods previous associations with serum urate levels. While alcohol is not in the Harvard pyramid, it is recommended to have the same servings as those foods in level 2 (in moderation and not sparingly as for those in level 1). High- and low-fat dairy products were distinguished according to the particular diet score – not for the Harvard score but were distinguished for the DASH score. (In response to reviewer 2 we have now included alcohol as an adjustor in the DASH score regression analysis.)

These foods were not excluded from level 4. We were deliberately nescient to the known effects of particular foods on serum urate levels. The aim of our study was to examine the effects of common 'healthy' diets, but not a 'gout diet', on serum urate levels.

Healthy fats and oils are included in level 4 of the Harvard pyramid. We did not include healthy fats and oils, used for cooking and dressings etc, as this information was not collected in a consistent manner across the five studies used in our analysis.

We have now used the word 'unfavourable' in preference to 'minimised' throughout.

We have checked the assumptions of a linear regression for linearity, as well as multivariate normality and homoscedasticity by generating several plots. We have reviewed these plots for a non-normal pattern, which would indicate a violation of one of these assumptions. For the reviewer there are plots @ http://bit.ly/DWASPlots which comprises 63x4 = 252 plots (per full, male-only, and female-only cohorts) for each of the foods from the diet-wide association study. While there are deviations at the outer ends of the qq-plot distributions in the adjusted models these are considerably lessened without adjustment, indicating that it is the adjustors (some of which explain greater amounts of variance than the serum urate-associated foods) causing the deviations. There were additional deviations in the homoscedasticity plots of the full cohort plots, these deviations were entirely explained by sex and were not present in the sex-specific plots.

We added these sentences to the Methods section of the revised manuscript (page 7): "For each food item the four basic statistical assumptions of a linear regression (linear relationship, multivariate normality, multi-colinearity, and homoscedasticity) were assessed by generating plots of the food item versus serum urate, the standardised residuals versus the normal distribution (quantile-quantile plot), and the standardised residuals versus the predicted values from the linear regression. These plots (not shown) did not indicate any substantial deviation from these basic statistical assumptions."

We have included years of education, exercise levels, and smoking status as adjustors in the revised manuscript. This has resulted in altered diet-wide significant individual food items in the DWAS (Table 1) and altered diet score associations (Table 2). There are now eighteen diet-wide significant associations (new urate-raising ones are beef / pork / lamb, potato, poultry, table sugar) and the new urate-lowering one is cold cereal. We have altered the text in the Results and Discussion accordingly. (Given that we have added the Mediterranean diet score and published beta values in Table 2 we have also reformatted Table 2, including reporting genome-wide heritability estimates only in the main text (page 16)).

This has been rectified.

This has been done in the data presented in Table 2 of the revised manuscript. We mention this in the Methods (page 9): "The diet quality scores were included in separate multivariate linear regression of serum urate levels adjusted for……. and alcohol for the DASH diet (the DASH score does not include a separate component for alcohol)."

A new column showing the median (minimum; maximum) has now been included for each food (per study) in Table S4.

We apologise, the data had inadvertently been presented for CARDIA as serves per day and has been rectified in Table S4. Please note that the CARDIA data for decaffeinated coffee and fried food had a strange distribution. This was due to < 100 people reporting consumption of either of these food items, as such the data for these two foods in CARDIA was considered unreliable and excluded. This resulted in only two of the five cohorts having data for these two foods, and as such they were excluded from the entire analysis – the text throughout the manuscript has been edited to reflect this reduction from 65 to 63 food items.

Each study collected serving size information differently, for the studies that specified a serving size for each food these serving sizes were not necessarily consistent between foods (eg. 4oz glass of wine, 1 bottle / can of beer, 1oz of cream cheese in FHS). As portion size was not considered in this study (now mentioned in the methods) adding portion sizes to Table S4 would unnecessarily complicate the table, therefore we have not done this.

This has been addressed in response to reviewer 1.

This has been rectified.

The nut category was missing in the CHS cohort. This, and including '*' for candy, has been addressed in the revised manuscript.

This has been rectified.

Reviewer 3

In the analyses presented in Table 2 the genetic risk score and dietary scores had been included in the same models. Thus the influence of one on the other, in terms of variance explained, had been reported (there was very little change).
Regarding the point on heritability of food preference we added these sentences to the Introduction: "In addition, food consumption is heritable, for example heritability of coffee consumption is estimated to be between 36 to 58% [25], alcohol consumption to be between 43 to 53% [26] and sugar-sweetened beverage consumption to be 48% [27], and genome-wide association studies have identified genetic associations with coffee and alcohol consumption habits [28,29]. It is therefore possible that the heritable component of specific food consumption contributes to the heritability of serum urate levels (e.g. signals in the ABCG2, GCKR, and MLXIPL genes are common to coffee consumption and serum GWAS) [14,28]."
We added this text to the Discussion in the penultimate study limitations paragraph (page 24): "Finally, there is a heritable component to food preferences including food consumption and alcohol [26,27], implying non-independence between dietary scores and the genetic risk score."

We have tested the association of the food groups fruit, meat, vegetables, dairy with serum urate levels and evaluated the amount of variance in serum urate levels explained. We added this sentence to page 15 of the revised manuscript: "Food groups (fruit, vegetables, meat, and dairy) explained between 0.16% and 0.52% of variation in serum urate levels (Table S5)."

We have done this using factor analysis. In the per-study parallel factor analysis all of them agree there should be between 9 & 12 vectors retained in the factor analysis, but as the actual dietary patterns vary between studies we chose to run the factor analysis on everyone together to get consistent loadings for each food. We ran this analysis with 11 factors and kept those factors with a sum of squared loadings <1 (indicating a large amount of dietary variability is being captured by the factor). Thus only only one data-driven diet pattern was assessed. It's made up of non-citrus juice, soft drinks, butter, white bread, pasta, beef / pork / lamb, and chips / popcorn. This diet score has been included in the revised manuscript. In contrast to the other scores it is positively associated with urate levels with broadly equivalent effect sizes and amount of variance explained. The data are included in Table 2 and the

study-specific distributions are in Figure S6. We have added text to the Methods (page 9), Results (page 13) and Discussion (page 19).

We are unaware of other sufficiently powered cohorts that could be used for independent replication as per other "WAS" studies. Therefore, as the reviewer suggests, for established associations we have included in Table 1 beta values from statistically significant associations with serum urate levels of food items from other studies that presented data on men and women combined. However some of these data are derived from NHANES III in studies led by Hyon Choi (except the Gao soft drinks study on NHANES 2001-2002 and the Scottish Zgaga et al study), so are not truly independent of our data. We added this sentence to Results (page 11): "The effect sizes for shellfish, skim milk, and non-citrus fruit were similar to those reported in previous studies, while the effect sizes for soft drink, beer, liquor, and beef / pork / lamb were within the range of previously reported values."

Regarding the point on replication we have now merged paragraphs 2 and 3 of the Discussion to incorporate more information from the literature on the less established associations. We have now focused this paragraph on association with serum urate and therefore deleted previous text concerning gout and CVD associations of fruit and peanut consumption. The substantive new text in this paragraph is: "Of the eleven novel / less established associations we found some evidence within the literature to support the tea, cheese, non-citrus fruit, egg, brown bread and cold cereal associations. ……Similarly, tea consumption was associated with increased serum urate levels in our study, contradictory to a recent meta-analysis that found no evidence for association of tea consumption with serum urate,[66] although the meta-analysis did provide weak evidence for association of green tea consumption with increased serum urate levels. Our study did not distinguish between black and green tea. ….. That coarse bread and cheese associate with reduced urate levels in two cohorts of European ancestry and cereal in one of two cohorts of European ancestry [65] provides support for our data associating brown bread, cold cereal, and cheese with reduced serum urate levels (Table 1). We are not aware of other studies specifically testing for association of potato, table sugar, peanut, and margarine consumption with serum urate levels thus these findings require replication before being claimed as genuine urate-raising or urate-lowering foods."