



# Alcohol intake and cause-specific mortality: conventional and genetic evidence in a prospective cohort study of 512 000 adults in China

Iona Y Millwood\*, Pek Kei Im\*, Derrick Bennett, Parisa Hariri, Ling Yang, Huaidong Du, Christiana Kartsonaki, Kuang Lin, Canqing Yu, Yiping Chen, Dianjianyi Sun, Ningmei Zhang, Daniel Avery, Dan Schmidt, Pei Pei, Junshi Chen, Robert Clarke, Jun Lv, Richard Peto, Robin G Walters, Liming Li, Zhengming Chen, on behalf of the China Kadoorie Biobank Collaborative Group†



## Summary

**Background** Genetic variants that affect alcohol use in East Asian populations could help assess the causal effects of alcohol consumption on cause-specific mortality. We aimed to investigate the associations between alcohol intake and cause-specific mortality using conventional and genetic epidemiological methods among more than 512 000 adults in China.

**Methods** The prospective China Kadoorie Biobank cohort study enrolled 512 724 adults (210 205 men and 302 519 women) aged 30–79 years, during 2004–08. Residents with no major disabilities from ten diverse urban and rural areas of China were invited to participate, and alcohol use was self-reported. During 12 years of follow-up, 56 550 deaths were recorded through linkage to death registries, including 23 457 deaths among 168 050 participants genotyped for *ALDH2*-rs671 and *ADH1B*-rs1229984. Adjusted hazard ratios (HRs) for cause-specific mortality by self-reported and genotype-predicted alcohol intake were estimated using Cox regression.

**Findings** 33% of men drank alcohol most weeks. In conventional observational analyses, ex-drinkers, non-drinkers, and heavy drinkers had higher risks of death from most major causes than moderate drinkers. Among current drinkers, each 100 g/week higher alcohol intake was associated with higher mortality risks from cancers (HR 1.18 [95% CI 1.14–1.22]), cardiovascular disease (CVD; HR 1.19 [1.15–1.24]), liver diseases (HR 1.51 [1.27–1.78]), non-medical causes (HR 1.15 [1.08–1.23]), and all causes (HR 1.18 [1.15–1.20]). In men, *ALDH2*-rs671 and *ADH1B*-rs1229984 genotypes predicted 60-fold differences in mean alcohol intake (4 g/week in the lowest group vs 255 g/week in the highest). Genotype-predicted alcohol intake was uniformly and positively associated with risks of death from all causes (n=12939; HR 1.07 [95% CI 1.05–1.10]) and from pre-defined alcohol-related cancers (n=1274; 1.12 [1.04–1.21]), liver diseases (n=110; 1.31 [1.02–1.69]), and CVD (n=6109; 1.15 [1.10–1.19]), chiefly due to stroke (n=3285; 1.18 [1.12–1.24]) rather than ischaemic heart disease (n=2363; 1.06 [0.99–1.14]). Results were largely consistent using a polygenic score to predict alcohol intake, with higher intakes associated with higher risks of death from alcohol-related cancers, CVD, and all causes. Approximately 2% of women were current drinkers, and although power was low to assess observational associations of alcohol with mortality, the genetic evidence suggested that the excess risks in men were due to alcohol, not pleiotropy.

**Interpretation** Higher alcohol intake increased the risks of death overall and from major diseases for men in China. There was no genetic evidence of protection from moderate drinking for all-cause and cause-specific mortality, including CVD.

**Funding** Kadoorie Charitable Foundation, National Natural Science Foundation of China, British Heart Foundation, Cancer Research UK, GlaxoSmithKline, Wellcome Trust, Medical Research Council, and Chinese Ministry of Science and Technology.

**Copyright** © 2023 The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license.

## Introduction

The harmful use of alcohol worldwide accounted for an estimated 3 million deaths in 2016.<sup>1</sup> The main alcohol-attributed causes of death include liver cirrhosis, cardiovascular disease (CVD), some cancers (eg, mouth and throat, oesophagus, and liver), tuberculosis, pneumonia, alcohol use disorders, and injuries.<sup>1,2</sup> Estimates of the disease burden attributed to alcohol

intake have typically been based on risk estimates derived from observational studies of populations predominantly from high-income countries. A recent study highlighted the importance of evidence from diverse populations, with different region-specific and age-specific disease rates.<sup>3</sup> Furthermore, although the observational evidence is based on large sample sizes, the associations observed might not reflect causal

*Lancet Public Health* 2023; 8: e956–67

Published Online  
November 21, 2023  
[https://doi.org/10.1016/S2468-2667\(23\)00217-7](https://doi.org/10.1016/S2468-2667(23)00217-7)

\*Joint first authors

†Members of China Kadoorie Biobank Collaborative Group listed in the appendix (p 3)

Clinical Trial Service Unit and Epidemiological Studies Unit (IY Millwood DPhil, P K Im DPhil, D Bennett PhD, P Hariri PhD, L Yang PhD, H Du PhD, C Kartsonaki DPhil, K Lin PhD, Y Chen DPhil, D Avery MSc, D Schmidt MSc, Prof R Clarke MD, Prof R Peto FRS, R G Walters PhD, Prof Z Chen DPhil) and Medical Research Council Population Health Research Unit (IY Millwood, P K Im, D Bennett, L Yang, H Du, C Kartsonaki, R G Walters, Prof Z Chen), Nuffield Department of Population Health, University of Oxford, Oxford, UK; Turku PET Centre, Turku University Hospital and University of Turku, Turku, Finland (P Hariri); Department of Epidemiology and Biostatistics, School of Public Health, Peking University, Beijing, China (CYu PhD, D Sun PhD, Prof J Lv PhD, Prof L Li PhD); Peking University Center for Public Health and Epidemic Preparedness & Response, Beijing, China (CYu, D Sun, P Pei BSc, Prof J Lv, Prof L Li); Key Laboratory of Epidemiology of Major Diseases (Peking University), Ministry of Education, Beijing, China (CYu, D Sun, Prof J Lv, Prof L Li); NCD Prevention and Control Department, Sichuan Center for Disease Control and Prevention, Chengdu, Sichuan, China (N Zhang); China National Center for Food Safety Risk Assessment, Beijing, China (Prof J Chen PhD)

Correspondence to:  
 Assoc Prof Iona Y Millwood,  
 Clinical Trial Service Unit and  
 Epidemiological Studies Unit,  
 Nuffield Department of  
 Population Health, University of  
 Oxford, Oxford OX3 7LF, UK  
 iona.millwood@ndph.ox.ac.uk  
 See Online for appendix

## Research in context

### Evidence before this study

Moderate alcohol intake has been associated with lower risks of mortality overall and from specific diseases—in particular, ischaemic heart disease. However, these associations might be largely non-causal, as conventional observational studies of alcohol use are susceptible to bias from reverse causation and residual confounding. Genetic evidence from Mendelian randomisation studies, in particular using the *ALDH2*-rs671 and *ADH1B*-rs1229984 variants, which strongly affect alcohol intake and are common in East Asian populations, can help assess the causal relevance of alcohol intake for cause-specific mortality.

We searched PubMed from database inception to Feb 25, 2023 using the following search terms (title or abstract) for articles published in English: ([Alcohol AND Mendelian] or (ALDH2 or ADH1B or rs671 or rs1229984 or aldehyde dehydrogenase or alcohol dehydrogenase)) AND (mortality or death or fatal), and reviewed bibliographies within the identified publications.

Two previous Mendelian randomisation studies of alcohol and mortality in populations with European ancestry, and one in Chinese men, reported that higher alcohol intake was associated with higher risks of all-cause mortality. However, these studies did not assess causal relevance across a wide range of alcohol intakes and did not evaluate effects on cause-specific mortality.

### Added value of this study

The present prospective study used both conventional and genetic approaches within the same population. The genetic analyses minimised artefacts of confounding and reverse causation and assessed potential causal relevance across a wide range of alcohol intakes, from negligible to moderate and heavy intakes. Among Chinese men, conventional observational analyses showed characteristic J-shaped associations of self-reported alcohol intake categories with overall and cause-specific mortality, with highest risks among ex-drinkers, non-drinkers, and heavy drinkers, and lowest risks among moderate drinkers, consistent with findings from similar studies in populations of high-income countries.

Genetic analyses, using two genetic variants that predicted a 60-fold difference in mean intake (from 4 g/week in the lowest category to 255 g/week in the highest category) showed that higher alcohol intake was associated with a uniform dose-response increase in risks of death overall and from some cancers, CVD, and liver diseases. There were no genetic associations with respiratory or non-medical (mainly accidents and injuries) causes of death. There was no genetic evidence that moderate alcohol intake (ie, 10–20 g/day) had substantial protective effects for cause-specific or overall mortality, including for ischaemic heart disease deaths. In separate genetic analyses using a polygenic score to predict alcohol intake, there were similar associations with mortality overall, and from alcohol-related cancers and cardiovascular disease (CVD), across different alcohol intakes.

Alcohol intake was extremely low among women in the study, and the genetic variants had little effect on mortality overall or from specific causes, suggesting that the higher risks in men were chiefly mediated by alcohol, rather than by any pleiotropic effects of the genotypes studied.

### Implications of all the available evidence

Genetic studies, particularly in East Asian ancestry populations, have helped to reliably clarify the causal relevance of alcohol intake with mortality by accounting for biases in conventional observational approaches. The genetic evidence provides strong support for causal harmful effects of alcohol use, with risks of deaths from CVD, cancer, liver disease, and all-causes. There is no genetic evidence of any net beneficial effects of moderate drinking compared with not drinking for any causes of death, including CVD. These genetic studies that assess causal relevance have improved our understanding of the adverse effects of alcohol use on mortality, particularly at lower intakes. This knowledge can improve estimation of the regional and global burden of alcohol use, and inform public health policies to address the risks of moderate and heavy drinking.

effects, which could affect estimation of the adverse effects of alcohol use.<sup>4</sup>

Over the past few decades, meta-analyses of prospective studies have reported that the observed risks of mortality overall, and particularly from CVD, were lower among moderate drinkers, leading to widespread acceptance of approximately 1–2 drinks per day as a safe level of consumption in general populations.<sup>5–7</sup> However, systematic differences in health characteristics and behaviours (such as previous ill health, socioeconomic status, or smoking behaviours) between non-drinkers, moderate drinkers, and heavy drinkers, often influenced by selection into cohort studies and their demographics characteristics, can lead to reverse causation (whereby

health status affects drinking patterns), confounding, and other biases.<sup>7,8</sup> In China, where alcohol consumption has increased steadily in recent decades, there is little evidence available on alcohol drinking and cause-specific mortality in the general adult population.<sup>9,10</sup>

Mendelian randomisation uses genetic variants as instrumental variables to assess the causal relevance of alcohol intake while minimising the biases inherent to conventional observational studies.<sup>11</sup> A Mendelian randomisation study of people with European ancestry associated alcohol intake with a higher risk of all-cause mortality, but specific causes of death were not investigated, nor was the shape of the association across different levels of intake.<sup>12</sup> In East Asian populations,

two common genetic variants (*ALDH2*-rs671 and *ADH1B*-rs1229984) alter the function of enzymes involved in alcohol metabolism and strongly affect alcohol tolerability and alcohol intake.<sup>4</sup> These genetic variants have been used to assess the causal relevance of alcohol intake for incidence of CVD, other diseases, and overall mortality.<sup>4,13,14</sup> Ascertaining the causal relevance of alcohol for major causes of death (particularly CVD, for which associations of alcohol with fatal *vs* non-fatal events can differ) can improve estimations of the global burden of alcohol use and inform policies for prevention of alcohol-related harms.

This study investigated the associations between alcohol consumption and cause-specific mortality among more than 512 000 adult men and women from the prospective China Kadoorie Biobank (CKB). In addition to assessing conventional observational associations, we used a Mendelian randomisation approach to assess the strength, shape, and causal relevance of genotype-predicted alcohol intake with cause-specific mortality among a subset of more than 168 000 men and women with data on *ALDH2*-rs671 and *ADH1B*-rs1229984 genotypes. Additional analyses used a polygenic score to predict alcohol intake and evaluate associations with mortality.

## Methods

### Study design and participants

CKB is a prospective cohort of 512 724 adults aged 30–79 years who did not have a major disability at enrolment (response rate 28%) during 2004–08 from ten areas of China.<sup>15</sup> At baseline, participants attended survey clinics and completed an interviewer-administered laptop-based questionnaire covering sociodemographic and lifestyle characteristics (eg, smoking and alcohol drinking) and medical history. Physical measurements were taken (eg, blood pressure and anthropometry), and a 10 mL blood sample was collected. Resurveys of approximately 5% of surviving participants, following similar procedures, were done in 2008 (n=19 786), 2013–14 (n=25 041), and 2021–22 (n=25 087). Ethics approval was obtained from local, national, and international ethics committees and all participants provided written informed consent.

### Procedures

Alcohol drinking patterns were self-reported at baseline and resurveys.<sup>16,17</sup> Participants were classified as current drinkers (some alcohol use in most weeks of the past year), non-drinkers (no alcohol use in the past year and previously did not drink most weeks), occasional drinkers (occasional alcohol use in the past year and previously did not drink most weeks), and ex-drinkers (occasional or no alcohol use in the past year but previously drank most weeks). Current drinkers provided further details about their drinking patterns, including frequency, amount, and beverage type, and were further classified by weekly

alcohol intake (<140, 140–279, 280–419, or  $\geq$ 420 g/week for men; <70 or  $\geq$ 70 g/week for women). To account for measurement error and within-person variability in self-reported alcohol use over time, for each of these baseline-defined groups, the usual mean amount of alcohol intake of the group was estimated from the average of intakes at two resurveys (appendix p 8).<sup>18</sup>

Cause-specific mortality was ascertained through linkage via unique national identification numbers to local death registries managed by China Centre for Disease Control (CCDC). All deaths were reviewed by regional CCDC staff and the underlying cause of death was assigned using the International Classification of Diseases, tenth revision (ICD-10). By Jan 1, 2019, after median 12 years follow-up (IQR 11–13), 56 550 (11%) participants had died, and 4028 (1%) were lost to follow-up.

Deaths were grouped into broad categories (eg, CVD ICD-10 chapter I00–I99), specific causes (eg, ischaemic heart disease ICD-10 I20–I25), or by previously assigned relationship to alcohol (eg, cancers or other diseases and injuries designated as related to alcohol by the International Agency for Research on Cancer or WHO; appendix p 9).<sup>2,19</sup>

168 050 participants were genotyped for *ALDH2*-rs671 and *ADH1B*-rs1229984, including 151 347 randomly-selected individuals (included in all genetic analyses) and 16 703 people who had been selected for nested case-control studies of CVD or chronic obstructive pulmonary disease (only included as cases in analyses of relevant outcomes; appendix p 10).

Using a previously described approach, alcohol intake was predicted using a combination of genotype and study area, both of which had strong associations with alcohol intake, enabling a wide range of alcohol intakes to be assessed.<sup>4</sup> Mean alcohol intake was calculated among men within each of the 90 combinations of genotypes (*ALDH2*-rs671 and *ADH1B*-rs1229984 each AA, AG, or GG, resulting in nine combined genotypes) across the ten areas. Thresholds at 10, 25, 50, 100, and 150 g/week were applied to group the genotype-predicted mean alcohol intake into six categories (C1–C6) for genetic analyses among all genotyped participants. Combining genotype with study area enabled a reliable assessment of the shape and strength of associations with outcomes across a wide range of genotype-predicted mean alcohol intake, rather than the smaller range predicted by the genotypes alone.

Women were assigned into the same six categories as men based on their genotype and area (without reference to their mean alcohol intake) to assess potential pleiotropic effects of the genotypes studied—ie, effects of genotype not mediated by alcohol.

Supplemental analyses among 85 386 men and women used a weighted polygenic score of 825 alcohol-related variants from a multi-ancestry genome-wide meta-analysis to predict alcohol intake (appendix pp 4–7).<sup>20</sup>

### Statistical analysis

Analyses were conducted among men and women separately. In conventional observational analyses, Cox proportional hazards regression models were stratified for age-at-risk (5-year groups from ages 35 to 84 years) and ten geographical areas, adjusted for education, household income, smoking, physical activity, and fresh fruit intake. Participants reporting previous diseases at baseline were excluded. To allow comparisons in analyses involving more than two exposure groups, the variance of the log risk in each group, including the reference group, was calculated to obtain group-specific 95% CIs.<sup>21</sup> To account for measurement error and within-person variability in alcohol use over time (ie, regression dilution bias), the log HRs were plotted against usual alcohol intake for current drinkers.<sup>18</sup> The slope of a weighted linear regression through the plotted log HRs was used to estimate the HR per 100 g/week (approximately 1–2 drinks per day, assuming 1 drink=10 g alcohol) usual alcohol intake. Sensitivity analyses excluded the first 5 years of follow-up and additionally adjusted for red meat intake and self-rated health.

In genetic analyses, associations of genotype-predicted alcohol categories with self-reported alcohol intake, potential confounders, and risks of cause-specific mortality were assessed. Cox proportional hazards regression models were stratified for age-at-risk and ten areas, and adjusted for genomic principal components.<sup>22</sup> Log HRs were plotted against mean alcohol intake in each genotype-predicted alcohol intake category. To estimate the HR per 100 g/week, analyses were performed separately within each area with adjustment for age-at-risk and regional principal components. The slopes of a weighted linear regression within each area were meta-analysed with inverse-variance weighting. To assess potential pleiotropy of the genetic instrument, a heterogeneity test compared the meta-analysed slopes between men and women. Sensitivity analyses included adjusting for covariates, excluding previous diseases, using logistic regression or a two-stage least-square (2SLS) Mendelian randomisation approach, using the 90 genotype–area combinations as a continuous exposure, and excluding the highest category of predicted alcohol intake.<sup>23</sup> Analyses of the individual genetic variants included a comparison of GG and GA genotypes, and interaction between genotypes and self-reported alcohol intake.

Supplemental analyses with a polygenic score used a 2SLS approach within areas, followed by meta-analysis with inverse-variance weighting. Beta estimates from the regression of alcohol against the polygenic score in men were applied to the polygenic score values in women, to facilitate an assessment of pleiotropy.

Since all-cause mortality is a competing risk for cause-specific mortality, Cox regression models censored participants at death from any cause (or loss to follow-up, or the global censoring date of Jan 1, 2019) to estimate

cause-specific HRs, which compared event rates in participants who were alive and free of the event of interest. Comparing the HRs for the first 6 and subsequent years of follow-up showed no evidence of departure from the proportional hazards assumption, apart from liver disease deaths in genetic analyses, which had greater HRs in the earlier follow-up period ( $p$  heterogeneity=0.002). Analyses used R software (version 4.0.5).

### Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

### Results

Among 512724 study participants, the mean age at baseline was 52 years (SD 11). 210 205 (41.0%) were men and 302 519 (59.0%) were women, and 226 191 (44.1%) were from urban areas. Among men, 69 900 (33.3%) reported drinking alcohol most weeks (current drinkers), which varied across the ten study areas (table 1; appendix p 11). Non-drinkers and ex-drinkers were older than occasional and current drinkers, were more likely to live in rural areas, and had poorer health at baseline. Education and household income levels were highest among moderate drinkers (up to 140 g/week). Heavier drinkers were more likely to smoke, and consumed fresh fruit less frequently. Alcohol was consumed mainly as spirits and mainly with meals, and 17.9% of current drinkers reported flushing after drinking (appendix p 12). Among women, 101 285 (33.5%) drank alcohol occasionally, but only 6244 (2.1%) were current drinkers.

The A-alleles of *ALDH2*-rs671 (frequency 0.21, range by area 0.13–0.29) and *ADH1B*-rs1229984 (0.69, 0.64–0.74; appendix p 13) were both associated with lower alcohol intake (appendix p 14). For *ALDH2*-rs671, mean alcohol intakes among men were 2 g/week for AA genotype, 37 g/week for AG, and 162 g/week for GG. For *ADH1B*-rs1229984, mean alcohol intakes among men were 101 g/week for AA genotype, 109 g/week for AG, and 162 g/week for GG. Combining the two variants with area predicted 60-fold differences in mean alcohol intake in men, from 4 g/week in the lowest category to 255 g/week in the highest category, with the prevalence of ever-regular drinking ranging from 2.9% (124/4269) to 74.0% (11720/15 838; appendix pp 15–16). These categories were not associated with education, smoking, or other potential confounders, except for fresh fruit intake, which was lower in the higher alcohol intake categories. Among women, similar genotype–area categories were not associated with appreciable differences in mean alcohol intake (range 1–8 g/week) or potential confounders.

Of the 56 550 deaths recorded (31 956 in men and 24 594 in women), CVD (23 290 deaths) and cancers (17 691 deaths) together accounted for 72.5%, with

|                                  | Overall         | Non-drinkers    | Ex-drinkers    | Occasional drinkers | Current drinkers     |                                       |  |                      |                   |
|----------------------------------|-----------------|-----------------|----------------|---------------------|----------------------|---------------------------------------|--|----------------------|-------------------|
|                                  |                 |                 |                |                     | All current drinkers | <140 g/week (men), <70 g/week (women) | 140–279 g/week (men), ≥70 g/week (women) | 280–419 g/week (men) | ≥420 g/week (men) |
| Men                              | 210 205         | 42 779 (20.4%)  | 18 295 (8.7%)  | 79 231 (37.7%)      | 69 900 (33.3%)       | 25 093 (11.9%)                        | 18 907 (9.0%)                            | 12 832 (6.1%)        | 13 068 (6.2%)     |
| Sociodemographic characteristics |                 |                 |                |                     |                      |                                       |  |                      |                   |
| Age, years                       | 52.8 (10.9)     | 57.0 (11.1)     | 56.8 (10.3)    | 51.0 (10.8)         | 51.5 (10.2)          | 51.3 (10.9)                           | 51.9 (10.2)                              | 51.0 (9.6)           | 50.7 (9.5)        |
| Urban residence                  | 91 358 (43.5%)  | 13 974 (31.2%)  | 7714 (41.1%)   | 34 645 (44.1%)      | 35 025 (50.0%)       | 14 730 (58.6%)                        | 9967 (52.7%)                             | 6257 (48.5%)         | 4071 (31.2%)      |
| Education >6 years               | 121 429 (57.8%) | 18 770 (54.5%)  | 8720 (56.7%)   | 51 809 (60.5%)      | 42 130 (57.6%)       | 17 618 (63.9%)                        | 11 559 (60.1%)                           | 7259 (59.6%)         | 5694 (55.7%)      |
| Household income >20 000 ¥/year  | 95 937 (45.6%)  | 17 816 (42.0%)  | 7935 (44.9%)   | 34 181 (46.7%)      | 36 005 (46.8%)       | 13 489 (53.1%)                        | 9769 (51.3%)                             | 6737 (49.6%)         | 6010 (50.4%)      |
| Lifestyle risk factors           |                 |                 |                |                     |                      |                                       |  |                      |                   |
| Current smokers                  | 128 371 (61.1%) | 23 063 (52.3%)  | 10 531 (60.4%) | 44 943 (56.9%)      | 49 834 (71.7%)       | 15 849 (64.6%)                        | 13 632 (72.1%)                           | 9818 (76.1%)         | 10 535 (79.6%)    |
| Regular fresh fruit intake*      | 48 414 (23.0%)  | 8638 (24.9%)    | 4368 (25.3%)   | 19 467 (25.2%)      | 15 941 (21.1%)       | 7606 (28.0%)                          | 4241 (22.0%)                             | 2299 (18.9%)         | 1795 (16.4%)      |
| Physical activity, MET-h/d       | 22.0 (15.3)     | 21.1 (15.1)     | 20.3 (14.5)    | 22.5 (15.6)         | 22.2 (15)            | 22.6 (14.5)                           | 23.1 (14.9)                              | 23.1 (15.4)          | 22.5 (15.2)       |
| SBP, mm Hg                       | 132.8 (20.0)    | 132 (21.5)      | 134.1 (21.5)   | 131 (18.8)          | 134.3 (19.8)         | 131.8 (18.9)                          | 134.3 (19.8)                             | 136 (20.0)           | 137.7 (20.7)      |
| BMI, kg/m <sup>2</sup>           | 23.4 (3.2)      | 23.3 (3.2)      | 23.9 (3.4)     | 23.4 (3.2)          | 23.4 (3.2)           | 23.7 (3.2)                            | 23.7 (3.2)                               | 23.7 (3.2)           | 23.8 (3.2)        |
| Self-reported medical history    |                 |                 |                |                     |                      |                                       |  |                      |                   |
| Self-reported poor health        | 18 741 (8.9%)   | 4852 (12.8%)    | 3453 (17.1%)   | 6040 (7.7%)         | 4396 (5.9%)          | 1532 (6.5%)                           | 1185 (6.4%)                              | 764 (6.1%)           | 915 (7.1%)        |
| Previous chronic disease†        | 47 547 (22.6%)  | 11 540 (27.4%)  | 7441 (37.9%)   | 15 801 (21.2%)      | 12 765 (18.0%)       | 5234 (19.9%)                          | 3421 (18.0%)                             | 2108 (17.3%)         | 2002 (17.5%)      |
| Women                            | 302 519         | 192 333 (63.6%) | 2657 (0.9%)    | 101 285 (33.5%)     | 6244 (2.1%)          | 3224 (1.1%)                           | 3020 (1.0%)                              | NA                   | NA                |
| Sociodemographic characteristics |                 |                 |                |                     |                      |                                       |  |                      |                   |
| Age, years                       | 51.5 (10.5)     | 52.7 (10.7)     | 55.2 (9.4)     | 49.3 (9.9)          | 52.9 (10.3)          | 53.0 (10.7)                           | 52.8 (9.9)                               | NA                   | NA                |
| Urban residence                  | 134 833 (44.6%) | 83 001 (42.7%)  | 708 (30.2%)    | 48 305 (48.1%)      | 2819 (46.5%)         | 1977 (60.9%)                          | 842 (29.0%)                              | NA                   | NA                |
| Education >6 years               | 130 930 (43.3%) | 67 035 (41.2%)  | 799 (46.5%)    | 60 127 (49.0%)      | 2969 (48.2%)         | 1927 (49.6%)                          | 1042 (44.6%)                             | NA                   | NA                |
| Household income >20 000 ¥/year  | 123 095 (40.7%) | 82 323 (38.0%)  | 776 (44.8%)    | 37 538 (44.2%)      | 2458 (47.0%)         | 1593 (41.1%)                          | 865 (36.8%)                              | NA                   | NA                |
| Lifestyle risk factors           |                 |                 |                |                     |                      |                                       |  |                      |                   |
| Current smokers                  | 7151 (2.4%)     | 3131 (1.9%)     | 300 (5.4%)     | 2740 (2.8%)         | 980 (7.9%)           | 243 (10.0%)                           | 737 (20.5%)                              | NA                   | NA                |
| Regular fresh fruit intake*      | 96 133 (31.8%)  | 52 003 (30.0%)  | 851 (42.7%)    | 40 726 (36.9%)      | 2553 (39.1%)         | 1681 (44.1%)                          | 872 (35.2%)                              | NA                   | NA                |
| Physical activity, MET-h/d       | 20.4 (12.8)     | 20.1 (13.3)     | 20.2 (11.1)    | 20.6 (11.7)         | 20.5 (11.6)          | 20.0 (11.5)                           | 19.6 (11.7)                              | NA                   | NA                |
| SBP, mm Hg                       | 129.9 (22.0)    | 130.8 (22.5)    | 129.1 (23.2)   | 127.9 (20.5)        | 127.8 (21.6)         | 127.5 (20.9)                          | 129.3 (22.1)                             | NA                   | NA                |
| BMI, kg/m <sup>2</sup>           | 23.8 (3.5)      | 23.9 (3.5)      | 24.0 (3.5)     | 23.8 (3.4)          | 23.7 (3.4)           | 23.8 (3.4)                            | 23.8 (3.4)                               | NA                   | NA                |
| Self-reported medical history    |                 |                 |                |                     |                      |                                       |  |                      |                   |
| Self-reported poor health        | 34 350 (11.4%)  | 22 080 (12.6%)  | 648 (21.5%)    | 10 987 (9.6%)       | 635 (8.0%)           | 299 (10.8%)                           | 336 (9.8%)                               | NA                   | NA                |
| Previous chronic disease†        | 67 276 (22.2%)  | 44 474 (23.3%)  | 989 (33.1%)    | 20 417 (20.9%)      | 1396 (19.9%)         | 771 (23.9%)                           | 625 (21.9%)                              | NA                   | NA                |

Data shown are n (%) or mean (SD). Means and percentages are adjusted for the age and study area structure of the CKB population for the four drinking groups, and for the CKB drinker population for the weekly intake groups, using direct standardisation. CKB=China Kadoorie Biobank. MET-h/d=metabolic equivalent of task per hour per day. NA=not applicable. SBP=systolic blood pressure. \*≥4 days per week. †Chronic diseases included self-reported history of coronary heart disease, stroke, transient ischaemic attack, cancer, diabetes, tuberculosis, cirrhosis, hepatitis, rheumatoid arthritis, peptic ulcer, emphysema or chronic bronchitis, gallstone or gallbladder disease, rheumatic heart disease, and kidney disease.

Table 1: Baseline characteristics by alcohol drinking status for men and women

respiratory diseases (5362) and non-medical causes (3750) accounting for a further 16.1% (appendix p 17).

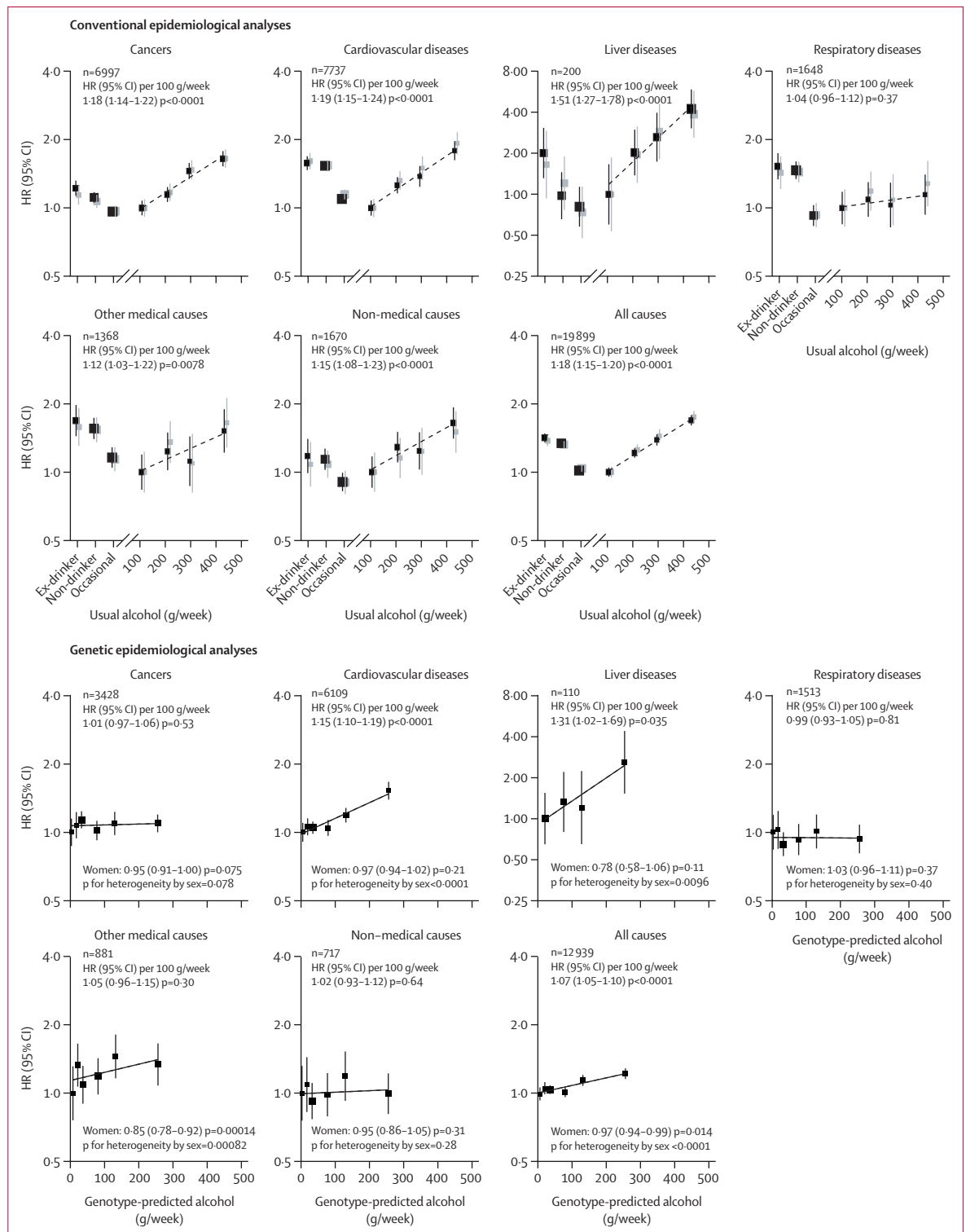
Among men, there were J-shaped or U-shaped associations between self-reported alcohol consumption and major causes of death, with higher risks in

ex-drinkers, non-drinkers, and heavier drinkers than occasional or moderate drinkers in analyses adjusted for age-at-risk, area, education, household income, smoking, physical activity, and fresh fruit intake (figure 1, appendix p 18). Consistent with the J-shaped



association with all-cause mortality, the estimated survival rate was higher in occasional and current drinkers, compared with non-drinkers and ex-drinkers (appendix p 19).

Among current drinkers, mortality risks increased with higher usual alcohol intake for CVD (HR per 100 g/week 1.19 [95% CI 1.15–1.24]), cancers (1.18 [1.14–1.22]), liver diseases (1.51 [1.27–1.78]), other (including ill-defined)



medical causes (1.12 [1.03–1.22]), non-medical causes (1.15 [1.08–1.23]), and all causes (1.18 [1.15–1.20]). The amount of alcohol intake was not associated with risks of death from infectious or respiratory diseases. With finer division of alcohol intake, the risk of all-cause mortality increased in a dose–response manner, with no evidence of a threshold at lower intakes (appendix p 20).

For specific causes of death, usual alcohol intake was associated with higher risks of ischaemic heart disease and stroke types (figure 2); cancers of the oesophagus, liver, and stomach; alcoholic liver disease and liver cirrhosis; and self-harm (appendix p 18). Associations were stronger for cancers pre-defined by the International Agency for Research on Cancer as alcohol-related (1.33 [95% CI 1.27–1.41]) than other cancers (1.09 [1.04–1.13]), and for causes pre-defined by WHO as alcohol-related (1.25 [1.21–1.28]) than other causes (1.09 [1.05–1.12]; figure 2, appendix p 21). The patterns of association were unaltered in sensitivity analyses to further address reverse causation and residual confounding (appendix p 22).

Among women, ex-drinkers and non-drinkers had higher risks of deaths from most causes than occasional or moderate drinkers, but among the few current drinkers, usual alcohol intake was only significantly associated with CVD mortality (1.50 [95% CI 1.06–2.13]; appendix p 23).

Among genotyped participants, there were 23457 deaths (13177 in men and 10280 in women; appendix p 17). In contrast with the J-shaped or U-shaped associations seen with self-reported alcohol consumption, mortality risks among men increased linearly across the range of genotype-predicted mean alcohol intake for CVD (HR per 100 g/week 1.15 [95% CI 1.10–1.19]), liver

diseases (1.31 [1.02–1.69]), and all causes (1.07 [1.05–1.10]) in pooled within-area analyses adjusted for age-at-risk and genomic principal components (figure 1, table 2). The genetic results were somewhat weaker than the corresponding estimates in the observational analyses (eg, 1.07 vs 1.18 per 100 g/week for all-cause mortality). There were no associations with respiratory, other medical, or non-medical causes of death. Although there was no association of genotype-predicted alcohol intake with overall cancer mortality (1.01 [0.97–1.06]), there was a positive association with the aggregated alcohol-related cancers (1.12 [1.04–1.21]; figure 2), including cancer of the oesophagus (1.16 [1.02–1.31]; table 2). In contrast to the positive association in conventional analyses, there was no associations with the aggregated other cancers (0.96 [0.91–1.01]).

Among types of CVD death, genotype-predicted alcohol intake was associated with higher risks of ischaemic stroke (1.12 [95% CI 1.00–1.25]), intracerebral haemorrhage (1.20, [1.13–1.28]), and total stroke (1.18 [1.12–1.24]). There were no significant associations with myocardial infarction (1.05 [0.96–1.15]) or overall ischaemic heart disease (1.06 [0.99–1.14]), although the trend was positive in both cases (figure 2). Genotype-predicted alcohol intake was associated with the aggregated WHO alcohol-related causes (1.13 [1.09–1.16]), but not with other causes of death (1.00 [0.97–1.04]; appendix p 21).

Sensitivity analyses, including those which excluded the highest category of genotype-predicted alcohol intake, did not materially alter the main genotypic findings, and although the magnitude of the excess risks varied (eg, 7–10% per 100 g/week for all-cause mortality), the 95% CIs all overlapped (appendix pp 24–26).

Compared with GA genotypes, *ALDH2*-rs671 GG was associated with higher risks of CVD and all-cause mortality, and *ADH1B*-rs1229984 GG with higher risks of alcohol-related cancer, CVD, and all-cause mortality (appendix pp 27–28).

There were interactions between *ALDH2*-rs671 genotype and self-reported alcohol intake for alcohol-related cancers, other cancers, and all-cause mortality, with higher risks for male drinkers with AG genotypes than those with GG genotypes (appendix p 29). When cancers were excluded, the interaction for all-cause mortality was null. There were no interactions with *ADH1B*-rs1229984 (appendix p 30). The HR per 100 g/week genotype-predicted alcohol intake for all-cause mortality excluding cancers was 1.10 (95% CI 1.07–1.13; appendix p 21).

Among women, using the same genotype-area categories as in men, there were no excess risks of cause-specific mortality (appendix p 31). However, there were lower risks of deaths from other medical causes (868 deaths; 0.85 [95% CI 0.78–0.92]), including diabetes (343 deaths; 0.80 [0.70–0.92]), colorectal cancer (234 deaths; 0.84 [0.71–1.00]), lung cancer (608 deaths; 0.89 [0.81–0.99]), and all causes (10057 deaths; 0.97

#### Figure 1: Conventional and genetic associations of alcohol intake with major cause-specific and all-cause mortality in men

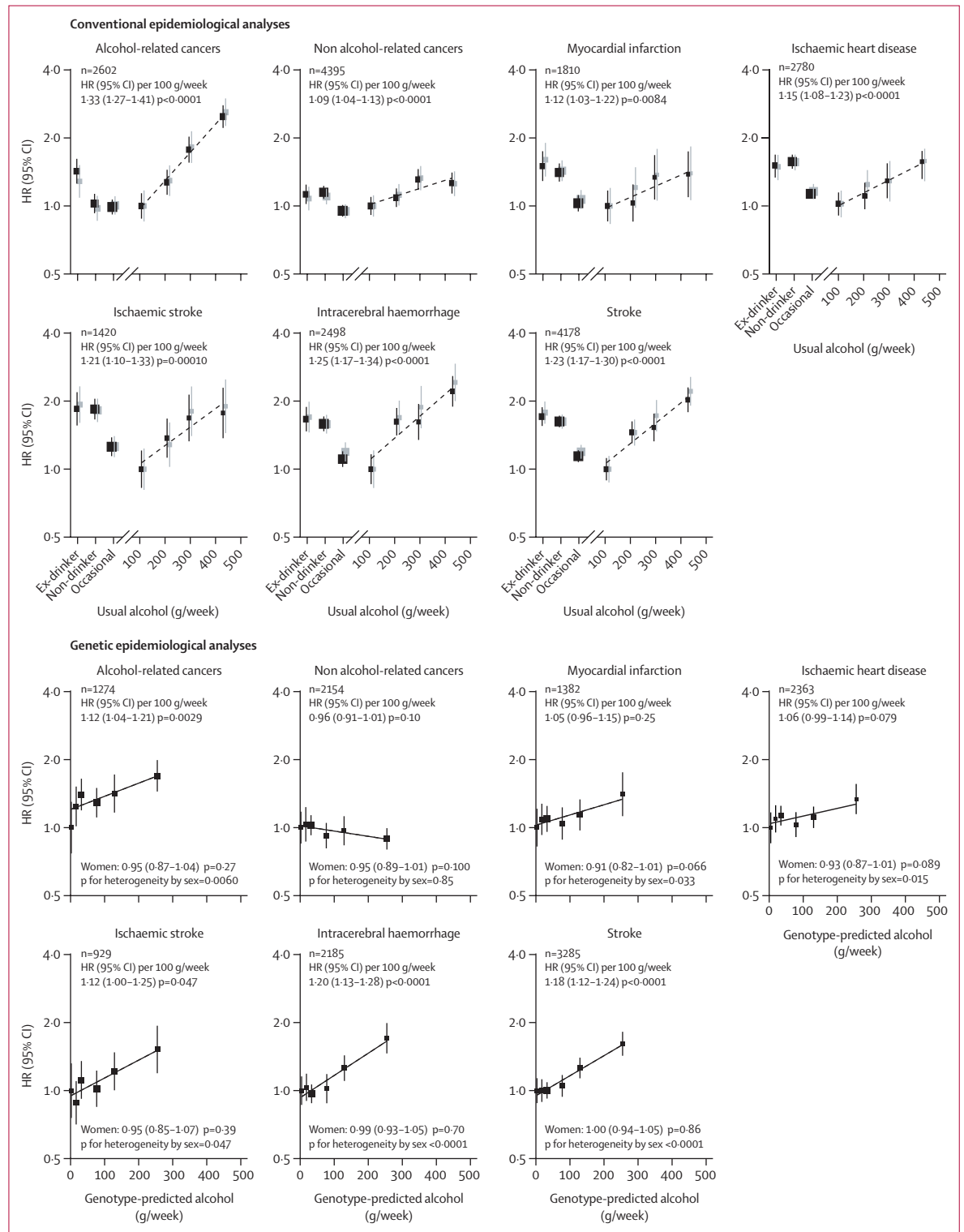
Conventional epidemiological analyses relate self-reported drinking patterns at baseline to mortality from major causes (all major causes are shown except for infectious diseases, for which the number of deaths was lower) and all-causes. Current drinkers with the lowest mean alcohol intake are the reference group.

The black squares represent findings from the main model adjusted for age-at-risk, area, education, household income, smoking, physical activity, and fresh fruit intake, with exclusion of participants with previous chronic disease. The HRs for current drinkers are plotted against usual alcohol intake, and a weighted linear regression through the plotted estimates gives the HR (95% CI) per 100 g/week (~1–2 drinks per day, assuming 1 drink contains 10 g alcohol).

The grey squares represent findings from sensitivity analyses that further exclude the first 5 years of follow-up. Genetic epidemiological analyses relate mean alcohol intake in six categories of genotype-predicted intake to mortality from major causes. The lowest mean intake group is the reference, and analyses are adjusted for age-at-risk, area, and genomic national principal components. HRs are plotted against the mean alcohol intake in each category. The HR (95% CI) per 100 g/week is the inverse-variance-weighted mean of a weighted linear regression through the plotted estimates within each study area, adjusted for age-at-risk and genomic regional principal components. The HR (95% CI) across six genetic categories in women applied the mean male intakes for each category, and the heterogeneity of effects was compared between men and women, to assess pleiotropy. The HR is plotted on a log scale. Each box represents HR with the area inversely proportional to the variance of the group-specific log hazard within each subplot. The vertical lines indicate group-specific 95% CIs. HR=hazard ratio.

[0.94–0.99]). Genotype-predicted risks differed substantially between men and women, with excess risks among men for alcohol-related cancer, CVD (including stroke types), liver, and all-cause mortality (figures 1, 2).

For both *ALDH2*-rs671 and *ADH1B*-rs1229984, there were excess risks among men for CVD and all-cause mortality when compared with women (appendix pp 27–28). Restricting the genetic analyses to





88080 women non-smokers did not alter the main findings (appendix p 32).

Among men, mean alcohol intake varied from 57 to 162 g/week across quintiles of a polygenic score, with no associations with potential confounders apart from small differences in fruit intake and smoking (appendix p 33). Polygenic score-predicted alcohol intake was associated with higher mortality risks from alcohol-related cancers (1.26 [95% CI 1.06–1.49]); CVD (1.18 [1.09–1.27]), including stroke (1.22 [1.11–1.35]) and ischaemic heart disease (1.16 [1.01–1.33]); and all causes (1.10 [1.05–1.16]; appendix pp 34–35). Risks of liver disease deaths (n=68) were higher (1.68 [0.99–2.86]), but not significantly. Among women, there were no associations, and the polygenic score only predicted small alcohol intake differences (3–6 g/week).

## Discussion

In this large prospective study of Chinese adults, using a strong genetic instrument to predict alcohol intake, we showed that genotype-predicted alcohol intake was associated with higher risks of mortality from CVD, particularly stroke, some cancers, liver diseases, and all-cause mortality. In contrast to the J-shaped associations seen in conventional observational analyses, there was no genetic evidence for a net protective effect of moderate drinking for major causes of death, including stroke and ischaemic heart disease, or overall mortality. For stroke, mortality risks increased linearly with amount of genotype-predicted alcohol intake, whereas for ischaemic heart disease mortality, there was a non-significant positive trend. Moreover, analyses among Chinese

women, who had very low intakes of alcohol, showed that the excess mortality hazards among men were probably due to alcohol itself, rather than to genetic pleiotropy.

Over the past few decades, numerous prospective studies have reported the lowest mortality risks among moderate drinkers (ie, 1–2 drinks per day), driven mainly by CVD deaths, in particular ischaemic heart disease.<sup>5,6,9,24</sup> In a combined analysis of 83 prospective studies, involving mainly populations in high-income countries and approximately 48000 deaths, the adjusted all-cause mortality risks were higher in ex-drinkers, non-drinkers, and heavier drinkers than moderate drinkers. Among current drinkers, risks did not increase until a threshold of approximately 2 drinks per day.<sup>5</sup> Although stroke mortality increased with higher alcohol intake, associations with ischaemic heart disease were less clear, with potentially different patterns for fatal and non-fatal events.<sup>5</sup> In the present study, with approximately 20000 deaths in Chinese men, we found similar lower risks among moderate drinkers, for mortality overall and for most major causes (including ischaemic heart disease and stroke), despite rigorous approaches to control for reverse causation and residual confounding. However, among male drinkers, there were continuous positive associations with major causes of death (apart from respiratory diseases) even at lower intakes, with no evidence of a threshold below which alcohol was unrelated to risk.

In recent years, Mendelian randomisation has been used to evaluate the causal relevance of alcohol for different diseases, but although a few previous Mendelian randomisation studies have reported higher risks of all-cause mortality with alcohol intake, they did not assess cause-specific mortality or evaluate causal relevance at different levels of intake.<sup>12,14,25</sup> A study including 13700 deaths in UK Biobank reported higher risks of all-cause mortality associated with each additional drink per day, using *ADH1B*-rs1229984 (OR 1.44 [95% CI 1.09–1.90]) or a score with 25 genetic variants (1.31 [1.08–1.59]).<sup>12</sup> A study of Australian men with 1329 deaths reported 47% higher all-cause mortality risk for *ADH1B*-rs1229984 GG genotypes compared with GA and AA genotypes (who drank less).<sup>25</sup> In the present genetic analyses, with 12939 deaths in men, there was a 17% (10–24%) higher risk of all-cause mortality for *ADH1B*-rs1229984 GG genotypes than GA genotypes.

In East Asian people, for whom the common *ALDH2*-rs671 variant is a strong determinant of alcohol intake, previous studies (including CKB) have assessed the causal relevance of alcohol in incident risk of CVD, cancer, and other diseases.<sup>4,13,26,27</sup> However, for cause-specific mortality, evidence from prospective studies is limited. A study with 2037 deaths in Chinese men reported a nominal trend for higher all-cause mortality with alcohol intake predicted by *ALDH2*-rs671.<sup>14</sup> In Biobank Japan participants (31403 deaths) both *ALDH2*-rs671 and *ADH1B*-rs1229984 A alleles were

### Figure 2: Conventional and genetic associations of alcohol intake with mortality from aggregated cancers and cardiovascular disease types in men

Conventional epidemiological analyses relate self-reported drinking patterns at baseline to mortality from aggregated cancers and cardiovascular disease types. Alcohol-related cancers include lip, oral cavity, pharynx, larynx, oesophagus, liver, and colon-rectum, defined as related to alcohol by the International Agency for Research on Cancer. Current drinkers with the lowest mean alcohol intake are the reference group. The black squares represent findings from the main model adjusted for age-at-risk, area, education, household income, smoking, physical activity, and fresh fruit intake, with exclusion of participants with previous chronic disease. The HRs for current drinkers are plotted against usual alcohol intake, and a weighted linear regression through the plotted estimates gives the HR (95% CI) per 100 g/week (~1–2 drinks per day, assuming 1 drink contains 10 g alcohol). The grey squares represent findings from sensitivity analyses that further exclude the first 5 years of follow-up. Genetic epidemiological analyses relate mean alcohol intake in six categories of genotype-predicted intake to mortality. The lowest mean intake group is the reference, and analyses are adjusted for age-at-risk, area, and genomic national principal components. HRs are plotted against the mean alcohol intake in each category. The HR (95% CI) per 100 g/week is the inverse-variance-weighted mean of a weighted linear regression through the plotted estimates within each study area, adjusted for age-at-risk and genomic regional principal components. The HR (95% CI) across six genetic categories in women applied the mean male intakes for each category, and the heterogeneity of effects was compared between men and women, to assess pleiotropy. The HR is plotted on a log scale. Each box represents HR with the area inversely proportional to the variance of the group-specific log hazard within each subplot. The vertical lines indicate group-specific 95% CIs. HR=hazard ratio.

|                                | n      | C1               | C2               | C3               | C4               | C5               | C6               | Per 100 g/week   | p value |
|--------------------------------|--------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|---------|
| Infectious diseases            | 163    | 1.00 (0.75-1.32) | NA               | NA               | 0.83 (0.50-1.38) | 1.38 (0.86-2.23) | 1.47 (0.88-2.46) | 1.22 (0.96-1.56) | 0.11    |
| Viral hepatitis                | 75     | 1.00 (0.69-1.45) | NA               | NA               | 0.30 (0.12-0.76) | 0.95 (0.43-2.09) | 1.26 (0.54-2.91) | 1.02 (0.64-1.62) | 0.94    |
| Cancers                        | 3428   | 1.00 (0.87-1.15) | 1.07 (0.94-1.23) | 1.13 (1.04-1.24) | 1.02 (0.92-1.12) | 1.09 (0.97-1.23) | 1.09 (1.00-1.20) | 1.01 (0.97-1.06) | 0.53    |
| Oesophageal cancer             | 406    | 1.00 (0.55-1.81) | 1.71 (1.27-2.31) | 1.45 (0.97-2.16) | 1.69 (1.37-2.08) | 2.00 (1.35-2.95) | 2.35 (1.70-3.23) | 1.16 (1.02-1.31) | 0.028   |
| Colorectal cancer              | 225    | 1.00 (0.50-2.02) | 1.71 (0.96-3.03) | 1.74 (1.23-2.47) | 2.05 (1.39-3.01) | 1.62 (1.04-2.52) | 2.40 (1.70-3.37) | 1.05 (0.88-1.26) | 0.59    |
| Liver cancer                   | 557    | 1.00 (0.71-1.40) | 0.92 (0.66-1.27) | 1.19 (0.96-1.47) | 1.02 (0.80-1.32) | 1.22 (0.92-1.60) | 1.29 (1.02-1.64) | 1.07 (0.95-1.19) | 0.28    |
| Stomach cancer                 | 498    | 1.00 (0.72-1.38) | 1.16 (0.88-1.53) | 1.05 (0.84-1.32) | 0.87 (0.68-1.12) | 0.73 (0.48-1.10) | 0.96 (0.75-1.23) | 0.94 (0.84-1.07) | 0.36    |
| Lung cancer                    | 1003   | 1.00 (0.79-1.27) | 0.99 (0.74-1.33) | 0.97 (0.84-1.11) | 0.83 (0.68-1.03) | 0.93 (0.77-1.13) | 0.82 (0.70-0.96) | 0.93 (0.86-1.01) | 0.091   |
| Alcohol-related cancers        | 1274   | 1.00 (0.77-1.30) | 1.24 (1.01-1.52) | 1.40 (1.19-1.65) | 1.29 (1.11-1.50) | 1.41 (1.17-1.72) | 1.69 (1.44-1.99) | 1.12 (1.04-1.21) | 0.0029  |
| Other (non-alcohol) cancers    | 2154   | 1.00 (0.85-1.18) | 1.03 (0.87-1.23) | 1.02 (0.93-1.13) | 0.92 (0.81-1.05) | 0.97 (0.84-1.12) | 0.89 (0.80-0.99) | 0.96 (0.91-1.01) | 0.10    |
| Cardiovascular diseases        | 6109   | 1.00 (0.91-1.10) | 1.06 (0.97-1.15) | 1.05 (0.98-1.11) | 1.05 (0.96-1.13) | 1.19 (1.10-1.28) | 1.52 (1.39-1.67) | 1.15 (1.10-1.19) | <0.0001 |
| Hypertensive heart disease     | 152    | 1.00 (0.80-1.24) | NA               | NA               | 0.69 (0.38-1.24) | 1.18 (0.67-2.06) | 1.44 (0.96-2.16) | 1.21 (0.98-1.51) | 0.082   |
| Ischaemic heart disease        | 2363   | 1.00 (0.85-1.17) | 1.10 (0.95-1.26) | 1.13 (1.02-1.25) | 1.03 (0.91-1.18) | 1.11 (1.00-1.24) | 1.35 (1.15-1.57) | 1.06 (0.99-1.14) | 0.079   |
| Myocardial infarction          | 1382   | 1.00 (0.82-1.21) | 1.09 (0.93-1.28) | 1.09 (0.96-1.25) | 1.04 (0.88-1.23) | 1.14 (0.97-1.33) | 1.40 (1.12-1.76) | 1.05 (0.96-1.15) | 0.25    |
| Stroke                         | 3285   | 1.00 (0.88-1.14) | 1.00 (0.89-1.12) | 1.00 (0.92-1.09) | 1.05 (0.94-1.17) | 1.26 (1.13-1.40) | 1.61 (1.43-1.82) | 1.18 (1.12-1.24) | <0.0001 |
| Ischaemic stroke               | 929    | 1.00 (0.76-1.32) | 0.88 (0.71-1.10) | 1.12 (0.92-1.35) | 1.02 (0.85-1.23) | 1.22 (1.00-1.48) | 1.52 (1.19-1.94) | 1.12 (1.00-1.25) | 0.047   |
| Intracerebral haemorrhage      | 2185   | 1.00 (0.86-1.16) | 1.03 (0.90-1.19) | 0.97 (0.88-1.07) | 1.02 (0.88-1.18) | 1.26 (1.10-1.43) | 1.71 (1.46-1.99) | 1.20 (1.13-1.28) | <0.0001 |
| Respiratory diseases           | 1513   | 1.00 (0.84-1.19) | 1.03 (0.85-1.24) | 0.89 (0.79-1.00) | 0.93 (0.79-1.09) | 1.01 (0.85-1.20) | 0.93 (0.81-1.08) | 0.99 (0.93-1.05) | 0.81    |
| Pneumonia                      | 176    | 1.00 (0.78-1.28) | NA               | NA               | 1.11 (0.78-1.59) | 0.97 (0.63-1.48) | 1.77 (1.01-3.12) | 1.11 (0.82-1.51) | 0.49    |
| COPD                           | 1214   | 1.00 (0.83-1.21) | 1.07 (0.85-1.33) | 0.91 (0.80-1.03) | 0.89 (0.74-1.08) | 1.05 (0.86-1.27) | 0.90 (0.76-1.05) | 0.99 (0.93-1.06) | 0.82    |
| Liver diseases                 | 110    | 1.00 (0.65-1.54) | NA               | NA               | 1.33 (0.80-2.20) | 1.20 (0.65-2.23) | 2.58 (1.52-4.38) | 1.31 (1.02-1.69) | 0.035   |
| ALD and liver cirrhosis        | 92     | 1.00 (0.60-1.66) | NA               | NA               | 1.45 (0.85-2.50) | 1.10 (0.54-2.24) | 3.12 (1.79-5.46) | 1.34 (1.02-1.76) | 0.034   |
| Other medical causes           | 881    | 1.00 (0.76-1.31) | 1.33 (1.07-1.65) | 1.09 (0.90-1.32) | 1.19 (0.99-1.43) | 1.45 (1.16-1.81) | 1.34 (1.08-1.66) | 1.05 (0.96-1.15) | 0.30    |
| Diabetes                       | 231    | 1.00 (0.57-1.75) | 1.68 (1.08-2.63) | 1.13 (0.79-1.60) | 1.67 (1.15-2.41) | 1.35 (0.86-2.13) | 1.82 (1.24-2.67) | 1.00 (0.84-1.18) | 0.96    |
| Renal diseases                 | 119    | 1.00 (0.79-1.27) | NA               | NA               | 0.67 (0.38-1.20) | 1.59 (1.01-2.51) | 0.90 (0.45-1.81) | 1.08 (0.81-1.43) | 0.61    |
| Ill-defined and unknown causes | 174    | 1.00 (0.79-1.26) | NA               | NA               | 1.03 (0.77-1.38) | 1.69 (1.04-2.74) | 0.82 (0.40-1.71) | 1.00 (0.72-1.41) | 0.99    |
| Non-medical causes             | 717    | 1.00 (0.76-1.32) | 1.09 (0.83-1.44) | 0.92 (0.77-1.11) | 0.99 (0.79-1.23) | 1.19 (0.93-1.53) | 1.00 (0.81-1.22) | 1.02 (0.93-1.12) | 0.64    |
| Transport accidents            | 281    | 1.00 (0.63-1.58) | 0.92 (0.55-1.53) | 1.04 (0.79-1.38) | 0.99 (0.69-1.41) | 0.91 (0.60-1.38) | 0.91 (0.67-1.23) | 0.96 (0.83-1.10) | 0.52    |
| Falls                          | 154    | 1.00 (0.76-1.32) | NA               | NA               | 1.00 (0.63-1.57) | 1.62 (0.90-2.90) | 0.99 (0.63-1.55) | 1.01 (0.81-1.24) | 0.95    |
| Self-harm                      | 86     | 1.00 (0.68-1.47) | NA               | NA               | 1.13 (0.53-2.44) | 1.35 (0.74-2.43) | 2.00 (1.05-3.81) | 1.13 (0.85-1.50) | 0.40    |
| All causes                     | 12 939 | 1.00 (0.94-1.07) | 1.06 (1.00-1.13) | 1.04 (1.00-1.09) | 1.02 (0.97-1.08) | 1.15 (1.08-1.21) | 1.23 (1.16-1.30) | 1.07 (1.05-1.10) | <0.0001 |
| WHO alcohol-related causes     | 8218   | 1.00 (0.92-1.09) | 1.06 (0.98-1.14) | 1.05 (0.99-1.11) | 1.06 (0.99-1.14) | 1.16 (1.09-1.24) | 1.44 (1.34-1.55) | 1.13 (1.09-1.16) | <0.0001 |
| Other causes                   | 4721   | 1.00 (0.90-1.11) | 1.10 (0.98-1.23) | 1.04 (0.97-1.11) | 0.97 (0.89-1.06) | 1.16 (1.05-1.28) | 1.01 (0.93-1.10) | 1.00 (0.97-1.04) | 0.91    |

Data shown are HR (95% CIs). Cox models were stratified by age-at-risk and study areas, and were adjusted for genomic national principal components. The slope per 100 g/week and p value was obtained within areas, adjusted for age-at-risk and genomic regional principal components, and meta-analysed with IVWMA. The partial F statistic for genotype-predicted alcohol intake categories within each area ranged from 34 to 783 (1752 overall), and the partial  $r^2$  ranged from 0.012 to 0.225 (0.136 overall). For cause-specific mortality with fewer than 200 deaths, C1-C3 were combined as one genetic category. ALD=alcoholic liver disease. COPD=chronic obstructive pulmonary disease. HR=hazard ratio.

Table 2: Genetic associations of alcohol intake with cause-specific mortality in men

weakly associated with lower all-cause mortality, but the findings were adjusted for alcohol so causal relevance could not be properly evaluated.<sup>28</sup> In the present Mendelian randomisation study of approximately 23 000 deaths (~13 000 in men) with a genetic instrument that predicted a 60-fold difference in mean alcohol intake in men, we found a uniform dose-response association of alcohol intake with risks of death from all causes, CVD (particularly stroke), some cancers (eg, oesophageal), and liver diseases, consistent with well established hazards considered by WHO to be alcohol-related.<sup>2</sup> For ischaemic heart disease mortality, there was no genetic evidence of any apparent protective effects of moderate drinking; if anything, there was a

positive trend towards higher risks with alcohol intake, which differs somewhat from the null association with non-fatal ischaemic heart disease.<sup>4,13</sup>

Given the very low alcohol consumption among women in the study, there was a unique opportunity to assess pleiotropy of the genetic variants, which provided strong support that the excess risks for CVD, some cancers, liver diseases, and overall deaths in men were due to alcohol itself. Although the *ALDH2-ADH1B* instrument had inverse associations in women for some outcomes, these were modest, and if anything would have attenuated the genetic associations in men towards the null. For causes pre-defined as unrelated to alcohol, the null genetic associations in men, in contrast to

positive associations with self-reported alcohol intake, indicate that the genetic approach is robust to confounding. Moreover, the lower genetic risk estimates, compared with the conventional dose–response estimates, also suggest potential uncontrolled residual confounding in the conventional analyses.

Estimation of the alcohol-attributable disease burden generally uses evidence from observational studies, which might not always reflect causal associations (eg, the apparently lower risks of CVD with moderate drinking), and large-scale randomised trial evidence is currently unavailable.<sup>1,3</sup> We show that alcohol itself is likely to be causally associated with deaths from several major causes in a linear and graded manner, with no apparent protective effects of moderate drinking for major causes of death, including CVD. Based on the approximately 7% excess risks for overall mortality per 100 g/week genotype-predicted alcohol intake, and the reported mean alcohol intake among men in the study, we estimate that alcohol drinking accounted for approximately 7–8% of male deaths in this Chinese population. This estimate is somewhat lower than that reported by other studies in China (eg, ~12% of male deaths at age 40–70 years in the 2016 Global Burden of Diseases, Injuries, and Risk Factors Study).<sup>1,16</sup> In addition to differences in relative risk estimates and sex-specific, region-specific, and age-specific alcohol intake, the proportions of deaths from different causes in different settings could greatly affect the estimation of alcohol-attributed mortality in China and elsewhere.<sup>3</sup>

Our study has several strengths, including incorporating a large number of deaths, use of strong genetic instruments, and the ability to assess genetic pleiotropy. However, it also has limitations. First, we lacked statistical power to study the effects of alcohol on less frequent causes of death (eg, tuberculosis); causes only affecting women (eg, breast cancer); or causes such as injuries, which could relate to alcohol differently among younger people or in different social contexts.<sup>1,3</sup> Second, our cohort study might have recruited disproportionately fewer heavy drinkers or more healthy people who had survived to middle age, leading to potential selection biases. Third, we did not assess associations of longitudinal drinking measurements with cause-specific mortality. Fourth, using the genetic methods available, we could not assess the causal relevance of drinking patterns (eg, heavy drinking episodes *vs* consumption with meals) and beverage types (eg, wine *vs* spirits) for cause-specific mortality. Finally, the genetic analyses estimates varied somewhat by the methods used, and were lower than the estimates in conventional analyses. However, this variation was small, and different methods (including use of an alternative polygenic score) gave generally consistent findings.

This study shows that alcohol use among Chinese men uniformly increases the risks of death overall and from

major causes (including CVD, some cancers, and liver diseases), with no evidence of protection conferred by moderate alcohol intake. Understanding the harms of alcohol use is important to inform and support public health strategies to reduce alcohol consumption at the population level. This information has started to be reflected in policy changes in some countries—for example, Canada has recently introduced guidance for low-risk drinking at a threshold of 1–2 drinks per week<sup>29</sup>—and new evidence from the present study could help accelerate such policy changes in other countries.

#### Contributors

IYM, PKI, LL, and ZC contributed to the conception of this paper. IYM, PKI, DB, PH, and ZC planned the statistical analyses. IYM drafted the manuscript. PKI analysed the data. IYM and PKI have accessed and verified the data reported in the manuscript. IYM, PKI, DB, and ZC contributed to the interpretation of the results and the revision of the manuscript. ZC, RP, JC, and LL designed the study. ZC, IYM, RGW, LY, YC, HD, CK, KL, RC, NZ, CY, PP, JL, DSu, and LL contributed to data acquisition and general study management. DA and DSc provided administrative and technical support. All authors critically reviewed the manuscript and approved the decision to submit for publication.

#### Declaration of interests

We declare no competing interests.

#### Data sharing

The CKB is a global resource for the investigation of lifestyle, environmental, blood biochemical, and genetic factors as determinants of common diseases. The CKB Collaboration Group is committed to making the cohort data available to the scientific community worldwide to advance knowledge about the causes, prevention, and treatment of disease. For detailed information on what data are currently available to open access users and how to apply for data, visit <https://www.ckbiobank.org/data-access>. Researchers who are interested in obtaining the raw data from the CKB study that underlies this paper should contact [ckbaccess@ndph.ox.ac.uk](mailto:ckbaccess@ndph.ox.ac.uk). A research proposal will be requested to ensure that any analysis is performed by bona fide researchers and, when data are not currently available to open access researchers, is restricted to the topic covered in this Article.

#### Acknowledgments

The chief acknowledgment is to the participants, the project staff, and the China Centre for Disease Control and its regional offices for assisting with the fieldwork. Members of the China Kadoorie Collaborative Group are listed in the supplementary material. The CKB baseline survey and the first resurvey were supported by the Kadoorie Charitable Foundation in Hong Kong. The long-term follow-up has been supported by Wellcome grants to Oxford University (212946/Z/18/Z, 202922/Z/16/Z, 104085/Z/14/Z, 088158/Z/09/Z), grants from the National Natural Science Foundation of China (82192900, 82192901, 82192904), and grants from the National Key Research and Development Program of China (2016YFC0900500). DNA extraction and genotyping was supported by grants from GlaxoSmithKline and the UK Medical Research Council (MC-PC-13049, MC-PC-14135). The UK Medical Research Council (MC\_UU\_00017/1, MC\_UU\_12026/2, MC\_U137686851), Cancer Research UK (C16077/A29186; C500/A16896), and the British Heart Foundation (CH/1996001/9454) provide core funding to the Clinical Trial Service Unit and Epidemiological Studies Unit at Oxford University for the project. PKI is supported by an Intermediate Career Research Fellowship from the Nuffield Department of Population Health, University of Oxford.

#### References

- 1 Griswold MG, Fullman N, Hawley C, et al. Alcohol use and burden for 195 countries and territories, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet* 2018; **392**: 1015–35.
- 2 WHO. Global status report on alcohol and health 2018. Geneva: World Health Organization, 2018.

- 3 GBD 2020 Alcohol Collaborators. Population-level risks of alcohol consumption by amount, geography, age, sex, and year: a systematic analysis for the Global Burden of Disease Study 2020. *Lancet* 2022; **400**: 185–235.
- 4 Millwood IY, Walters RG, Mei XW, et al. Conventional and genetic evidence on alcohol and vascular disease aetiology: a prospective study of 500 000 men and women in China. *Lancet* 2019; **393**: 1831–42.
- 5 Wood AM, Kaptoge S, Butterworth AS, et al. Risk thresholds for alcohol consumption: combined analysis of individual-participant data for 599 912 current drinkers in 83 prospective studies. *Lancet* 2018; **391**: 1513–23.
- 6 Di Castelnuovo A, Costanzo S, Bonaccio M, et al. Alcohol intake and total mortality in 142 960 individuals from the MORGAM Project: a population-based study. *Addiction* 2022; **117**: 312–25.
- 7 Zhao J, Stockwell T, Naimi T, Churchill S, Clay J, Sherk A. Association between daily alcohol intake and risk of all-cause mortality: a systematic review and meta-analyses. *JAMA Netw Open* 2023; **6**: e236185.
- 8 Emberson JR, Bennett DA. Effect of alcohol on risk of coronary heart disease and stroke: causality, bias, or a bit of both? *Vasc Health Risk Manag* 2006; **2**: 239–49.
- 9 Yang L, Zhou M, Sherliker P, et al. Alcohol drinking and overall and cause-specific mortality in China: nationally representative prospective study of 220,000 men with 15 years of follow-up. *Int J Epidemiol* 2012; **41**: 1101–13.
- 10 Manthey J, Shield KD, Rylett M, Hasan OSM, Probst C, Rehm J. Global alcohol exposure between 1990 and 2017 and forecasts until 2030: a modelling study. *Lancet* 2019; **393**: 2493–502.
- 11 Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet* 2014; **23**: R89–98.
- 12 Lankester J, Zanetti D, Ingelsson E, Assimes TL. Alcohol use and cardiometabolic risk in the UK Biobank: a Mendelian randomization study. *PLoS One* 2021; **16**: e0255801.
- 13 Im PK, Wright N, Yang L, et al. Alcohol consumption and risks of more than 200 diseases in Chinese men. *Nat Med* 2023; **29**: 1476–86.
- 14 Hu C, Huang C, Li J, et al. Causal associations of alcohol consumption with cardiovascular diseases and all-cause mortality among Chinese males. *Am J Clin Nutr* 2022; **116**: 771–79.
- 15 Chen Z, Chen J, Collins R, et al. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. *Int J Epidemiol* 2011; **40**: 1652–66.
- 16 Millwood IY, Li L, Smith M, et al. Alcohol consumption in 0.5 million people from 10 diverse regions of China: prevalence, patterns and socio-demographic and health-related correlates. *Int J Epidemiol* 2013; **42**: 816–27.
- 17 Im PK, Millwood IY, Guo Y, et al. Patterns and trends of alcohol consumption in rural and urban areas of China: findings from the China Kadoorie Biobank. *BMC Public Health* 2019; **19**: 217.
- 18 MacMahon S, Peto R, Cutler J, et al. Blood pressure, stroke, and coronary heart disease. Part 1, Prolonged differences in blood pressure: prospective observational studies corrected for the regression dilution bias. *Lancet* 1990; **335**: 765–74.
- 19 International Agency for Research on Cancer Working Group on the Evaluation of Carcinogenic Risks to Humans. A review of human carcinogens. Personal habits and indoor combustions. *IARC Monogr Eval Carcinog Risks Hum* 2012; **100**: 1–538.
- 20 Saunders GRB, Wang X, Chen F, et al. Genetic diversity fuels gene discovery for tobacco and alcohol use. *Nature* 2022; **612**: 720–24.
- 21 Plummer M. Improved estimates of floating absolute risk. *Stat Med* 2004; **23**: 93–104.
- 22 Walters RG, Millwood IY, Lin K, et al. Genotyping and population characteristics of the China Kadoorie Biobank. *Cell Genomics* 2023; **3**: 100361.
- 23 Burgess S, Davey Smith G, Davies NM, et al. Guidelines for performing Mendelian randomization investigations: update for summer 2023. *Wellcome Open Res* 2023; **4**: 186.
- 24 Saito E, Inoue M, Sawada N, et al. Impact of alcohol intake and drinking patterns on mortality from all causes and major causes of death in a Japanese population. *J Epidemiol* 2018; **28**: 140–48.
- 25 Almeida OP, McCaul K, Hankey GJ, Yeap BB, Golledge J, Flicker L. Excessive alcohol consumption increases mortality in later life: a genetic analysis of the health in men cohort study. *Addict Biol* 2017; **22**: 570–78.
- 26 Im PK, Yang L, Kartsonaki C, et al. Alcohol metabolism genes and risks of site-specific cancers in Chinese adults: an 11-year prospective study. *Int J Cancer* 2022; **150**: 1627–39.
- 27 Liu Z, Song C, Suo C, et al. Alcohol consumption and hepatocellular carcinoma: novel insights from a prospective cohort study and nonlinear Mendelian randomization analysis. *BMC Med* 2022; **20**: 413.
- 28 Sakaue S, Akiyama M, Hirata M, et al. Functional variants in ADH1B and ALDH2 are non-additively associated with all-cause mortality in Japanese population. *Eur J Hum Genet* 2020; **28**: 378–82.
- 29 Paradis C, Butt P, Shield K, et al. Canada's guidance on alcohol and health: final report. Ottawa, ON: Canadian Centre on Substance Use and Addiction, 2023.