Alcohol Consumption and the Risk of Type 2 Diabetes: A Systematic Review and Dose-Response Meta-analysis of More Than 1.9 Million Individuals From 38 Observational Studies

Craig Knott, Steven Bell, and Annie Britton

Diabetes Care 2015;38:1804–1812 | DOI: 10.2337/dc15-0710

SYSTEMATIC REVIEW

OBJECTIVE

Observational studies indicate that moderate levels of alcohol consumption may reduce the risk of type 2 diabetes. In addition to providing an updated summary of the existing literature, this meta-analysis explored whether reductions in risk may be the product of misclassification bias.

RESEARCH DESIGN AND METHODS

A systematic search was undertaken, identifying studies that reported a temporal association between alcohol consumption and the risk of type 2 diabetes. No restrictions were placed upon the language or date of publication. Non-English publications were, where necessary, translated using online translation tools. Models were constructed using fractional polynomial regression to determine the best-fitting dose-response relationship between alcohol intake and type 2 diabetes, with a priori testing of sex and referent group interactions.

RESULTS

Thirty-eight studies met the selection criteria, representing 1,902,605 participants and 125,926 cases of type 2 diabetes. A conventional noncurrent drinking category was reported by 33 studies, while five reported a never-drinking category. Relative to combined abstainers, reductions in the risk of type 2 diabetes were present at all levels of alcohol intake <63 g/day, with risks increasing above this threshold. Peak risk reduction was present between 10–14 g/day at an 18% decrease in hazards. Stratification of available data revealed that reductions in risk may be specific to women only and absent in studies that adopted a never-drinking abstention category or sampled an Asian population region.

CONCLUSIONS

Reductions in risk among moderate alcohol drinkers may be confined to women and non-Asian populations. Although based on a minority of studies, there is also the possibility that reductions in risk may have been overestimated by studies using a referent group contaminated by less healthy former drinkers.

Research Department of Epidemiology and Public Health, University College London, London, U.K.

Corresponding author: Craig Knott, craig.knott.10@ucl.ac.uk.

Received 6 April 2015 and accepted 9 June 2015.

This article contains Supplementary Data online at http://care.diabetesjournals.org/lookup/suppl/doi:10.2337/dc15-0710/-/DC1.

Type 2 diabetes is associated with substantial increases in the risk of vascular morbidities, such as coronary heart disease and stroke (1), as well as health complications ranging from kidney failure and incontinence to limb loss and blindness (2). Collectively, ~12% of global health expenditure was spent on diabetes in 2010, or 376 billion USD, rising to 490 billion USD over the next two decades (3). Such figures ignore the indirect costs of diabetes, including sickness absence, early retirement, and demand for social care.

Alongside established lifestyle factors, such as smoking (4), adiposity (5), and diet (6,7,8), alcohol consumption is also thought to play a role in the development of type 2 diabetes. The most recent meta-analysis to have explored the alcohol-diabetes relationship was undertaken by Baliunas et al. (9) in 2009. Pooling data from 20 observational studies, they identified peak risk reduction at 24 g/day (relative risk [RR] 0.60, 95% CI 0.52–0.69) among women and 22 g/day (RR 0.87, 95% CI 0.76–1.00) among men, relative to never drinkers, with risk increasing in a dose-dependent manner above these levels.

Several biological mechanisms have been proposed to explain the apparent reduction in risk of type 2 diabetes among moderate drinkers. These include the anti-inflammatory hypothesis, which posits that alcohol may beneficially alter the expression of inflammatory proteins implicated in metabolic processes (10), including adiponectin (11) and interleukin-1β (12), and a possible stimulatory effect of alcohol upon the synthesis of HDL (11). However, studies investigating such mechanisms are subject to notable limitations, including short follow-up periods and small sample sizes, limiting the generalizability of findings both at the population level and over the long term (13).

It is possible that reductions in risk identified between moderate alcohol exposure and incident type 2 diabetes may occur partly as an artifact of referent group selection, particularly where confounder adjustment is weak (14,15). To date, observational studies have commonly adopted pooled nondrinkers as the unexposed referent category. However, nondrinkers are far from homogeneous, comprising both never and former drinkers. Former drinkers are particularly notable, displaying poorer health and higher levels of mortality than moderate and never drinkers (16). Many existing alcohol-diabetes studies may have therefore overestimated the degree of risk reduction among moderate consumers of alcohol by comparing drinkers to a less healthy nondrinking referent category (17). Indeed, in a meta-analysis exploring the relationship between alcohol consumption and all-cause mortality, reductions in risk were attenuated when data were restricted to studies that excluded former drinkers from the referent category (18). Similar findings have been identified elsewhere (14,19).

Although a preceding meta-analysis (9) attempted to overcome the methodological shortcoming of calculating risks relative to pooled nondrinkers, they did so only by weighting studies with nondrinking referent categories according to the sex-specific proportions of former drinkers reported by five studies for which such data were available. Of these five studies, just two had strictly defined never drinking as lifelong abstention. It was unclear whether proportions of never drinkers drawn from five studies could be reliably applied to a multitude of disparate study populations.

A new meta-analysis was thus undertaken. In addition to updating the pool of selected studies, this meta-analysis explicitly sought to test for differences in the dose-response relationship according to the choice of referent group and reports referent-specific associations between average daily alcohol consumption and incident cases of type 2 diabetes.

RESEARCH DESIGN AND METHODS

Data Sources and Searches

PubMed (MEDLINE), Embase, the Cumulative Index to Nursing and Allied Health Literature (CINAHL), and the Alcohol and Alcohol Problems Science (ETOH) databases were searched for relevant studies.

Where possible, searches identified publications with titles or abstracts containing an alcohol-related term ("alcohol," "ethanol," or "drink*"), plus a diabetes-related term ("diabet*", "NIDDM," or "T2D*"), plus a term indicative of longitudinal observational data ("cohort," "inciden*", "prospective," "longitudinal," "case," or "retrospective"). No limits were placed upon the language or date of publication, and searches were undertaken on 18 February 2014. Unpublished literature, including conference abstracts and working papers, was not included.

Of publications included in the final meta-analysis, referenced and referencing publications were searched for additional literature not captured by initial electronic searches.

Study Selection

Types of Study

Cohort, case-cohort, case-control, and nested case-control designs were eligible, and both community and occupational data sets were considered. Participants had to be adults aged ≥16 years.

Types of Exposure

Sex-specific self-reported alcohol consumption was selected as the exposure of interest. With nonlinear relationships having previously been identified between alcohol consumption and type 2 diabetes (9), consumption needed to be reported across three or more categories, inclusive of a never- or nondrinking group. Studies were excluded if consumption could not be converted into grams per day and if any abstention category was contaminated by current drinkers.

Types of Outcome

Incident type 2 diabetes was selected as the outcome. Diagnostic tests and their respective thresholds have varied over time (20). Restricting selection to publications that defined type 2 diabetes according to current recommendations would unnecessarily exclude earlier publications, which adopted the gold standard of the period. Such an approach would also exclude self-reported outcome data. An inclusive range of measures were thus considered: all historic World Health Organization recommendations, self- or doctor-reported diagnosis, or anti-diabetes medication prescription or linkage to clinical registry data.

Short-listing Against Selection Criteria

Duplicate publications were omitted and remaining publications screened to remove any that did not report a temporal association between alcohol exposure and type 2 diabetes. Screened publications were then independently short-listed against study selection criteria by two authors, with one-third reviewed by all three authors. Differences of opinion were resolved via the input of the third reviewer, and the majority decision was upheld where a publication was reviewed by all three reviewers. The degree of agreement between reviewers was determined using the Cohen and Fleiss κ (21) statistics. In all cases, agreement was high ($\kappa \geq 0.815$).

Data Requests

To limit the number of excluded publications, we contacted authors of studies that reported an alcohol-diabetes relationship but did not meet selection criteria and requested revised analyses modified in accordance with selection criteria.

Duplicate Studies

Duplicate studies were identified among short-listed entries and omitted with consideration for the type and number of confounding factors, sample size, and length of follow-up. Decisions were reached by consensus.

Data Extraction and Quality Assessment

Data Extraction

Once eligible studies had been short-listed, relevant characteristics and results were extracted and independently verified by a second reviewer. Extracted data included sample size, country, baseline age, sex, confounder adjustment, length of follow-up, and risk estimates for each exposure category.

Measures of Exposure

Exposure reported in number of drinks was converted to grams per day assuming country-specific standard drinks (22). Exposures categorized according to periods longer than a day were converted into daily estimates assuming an even distribution of consumption over the reference period. Where averages were not reported for each exposure category, the medians of the lower and upper limits were selected. For categories with no upper limit, median values were defined as 1.5 times the lower limit of the category (9).

Measures of Effect

As odds ratios approximate RRs only when the incidence of an outcome is low, published odds ratios and their respective CIs were adjusted according to the Zhang and Yu method (23). With hazard ratios being a form of RR that is independent of study length (24), hazard ratios were considered equivalent to RRs for the purpose of the meta-analysis.

Where publications reported a referent group other than never or noncurrent drinking, risk estimates were recalculated to ensure that risk estimates were each relative to the reference group of interest (25). Using the Hamling method accounted for the nonindependence present between estimates that share the same reference category, thereby reducing any underestimation of variance during their recalculation (25,26). Adjustment for this covariance was also undertaken during the calculation of meta-analytic models.

Estimates were extracted from models that reported sex-specific risk across three or more categories of exposure and incorporated the maximum number of confounding variables without adjustment for potential mediators, i.e., markers of insulin, glucose, or triglycerides.

Quality Assessment

Study quality was assessed using the Newcastle-Ottawa quality assessment scale (27). It comprises eight questions grouped under three broad dimensions: selection of groups under study, comparability of groups under study, and outcome ascertainment. Questions range from the representativeness of the sample to the method of case ascertainment. A single point is awarded for each question bar concerning the comparability of the groups under study, for which up to two points can be awarded. Study quality was thus determined on a scale from 0 to 9 points. A full list of questions and criteria used for determining study quality is provided in Supplementary Table 1.

The effect of study quality was explored by stratifying data according to whether studies were scored below the median value.

Data Synthesis and Analysis

Model Selection

Models were constructed using fractional polynomial regression, which permitted the expression of nonlinear relationships (28). Building on a null model containing only a constant parameter, first-order and second-order polynomials were fitted for each analysis according to a restricted range of fractional powers.

Fit for each analysis was determined according to the deviance statistic, equivalent to the sum of squared residuals under OLS regression, such that the best-fitting model was that which reported deviance closest to zero.

Random Effects

All analyses were undertaken using random effects (29). The overall degree of heterogeneity present between studies was quantified using the I^2 index (30).

Small-Study Effects

As asymmetry cannot be explored using continuous dose-response data, alcohol consumption was recoded into a drinking/nondrinking binary variable and risk estimates were recalculated accordingly. The log of these new estimates was then plotted against the log SE, with a summary estimate calculated according to a standard fixed-effect meta-analysis (31). For the purpose of identifying small-study effects, the use of a random-effects weighting component is not recommended. Doing so would provide a greater relative weight to smaller studies and may mask any underlying asymmetry where sample size and the direction of a point estimate are associated (30).

All analyses were performed using Stata, version 13 (StataCorp, College Station, Texas).

Additional Analyses

In addition to the primary analysis of all pooled data combined, a priori consideration was given to the effect of sex and referent group, stratifying data by these explanatory factors where significant to the 0.05 level.

Upon identifying a single study that contributed a substantial proportion of sampled data, an a posteriori sensitivity analysis was undertaken. This explored the effect of excluding the large study from the pooled analysis.

A further a posteriori sensitivity analysis was undertaken to explore why male dose-response data appeared to differ from that reported previously (9). Male dose-response data were stratified according to whether they had been extracted from publications included in the 2009 meta-analysis (n = 17) or new publications sampled as part of this current meta-analysis (n = 20). Although the 2009 meta-analysis sampled 20 publications, only 17 of these were included in this current meta-analysis. Of the three that were omitted, one did not appear to report sex-specific risk estimates, while two concerned studies for which newer data were available that benefited from increased sampled size, longer follow-up, or greater confounder adjustment.

Finally, factors potentially contributing to any observed heterogeneity were investigated. These were thought to include participant age, method of case ascertainment, degree and type of confounder adjustment, follow-up duration, the healthy worker effect (32), and population region (33). Due to the risk of aggregation bias, only a subset of factors could be explored in the absence of individual-level data (34). Data were stratified on each appropriate factor, with differences explored visually after

adjustment for the effect of sex and reference group.

RESULTS

Of an initial 2,704 results, 38 studies met a priori selection criteria: 33 used a conventional noncurrent drinking category and 5 included a never-drinking category, strictly defined as zero consumption across the life course (Fig. 1). Selected study characteristics are summarized in Supplementary Tables 2 and 3. Aggregate data were available for 1,082,639 male and 819,966 female participants, among whom 79,633 and 46,293 cases of type 2 diabetes were reported, respectively. Crude or age-adjusted estimates were provided by 15 studies. Of the remaining 23 studies, covariate adjustment was variable: adiposity (n = 17), smoking (n = 16), physical activity (n = 15), heritability (n = 10), education (n = 9), dietary variables (n = 6), blood pressure (n = 5), ethnicity (n = 3), and occupation (n = 3).

All Data

Data from all 38 studies are plotted in Fig. 2. Studies each contributed at least three data points, inclusive of reference category, which were all plotted of a size inversely proportional to their SE. Visual inspection suggested considerable between-study heterogeneity: an observation corroborated by an I^2 of 75% (95% CI 67–80) along the first-order polynomial and 50% (95% CI 31–63) along the second-order polynomial.

Relative to all abstainers (current nondrinkers and never drinkers), a reduction in the risk of type 2 diabetes appeared present at all levels of alcohol intake <63 g/day, with risks increasing above this threshold. Peak risk reduction was present between 10–14 g/day, with an 18% decrease in risk relative to combined abstainers. The nonlinear model offered a better parameterization of the dose-response relationship than a linear regression ($P \leq 0.001$).

Sex-Specific Data

A sex-stratified scatter diagram of extracted data indicated a difference in the dose-response relationship by sex. A sex-interaction term was found to be significant ($P \leq 0.001$) and improved the fit of the model ($P \leq 0.001$).

Sex-stratified results are presented in Fig. 3 and indicate that any reduction in risk may be specific to women, who exhibited a decreased risk of type 2 diabetes at <71 g/day and peak reduction of 34% at 31–37 g/day relative to combined abstainers. This equated to any level of alcohol consumption below approximately four pints of 4% ABV lager per day, with peak reduction at almost two pints of 4% ABV lager per day. For men, a shallow increase in risk appeared to be present from very low levels of consumption.

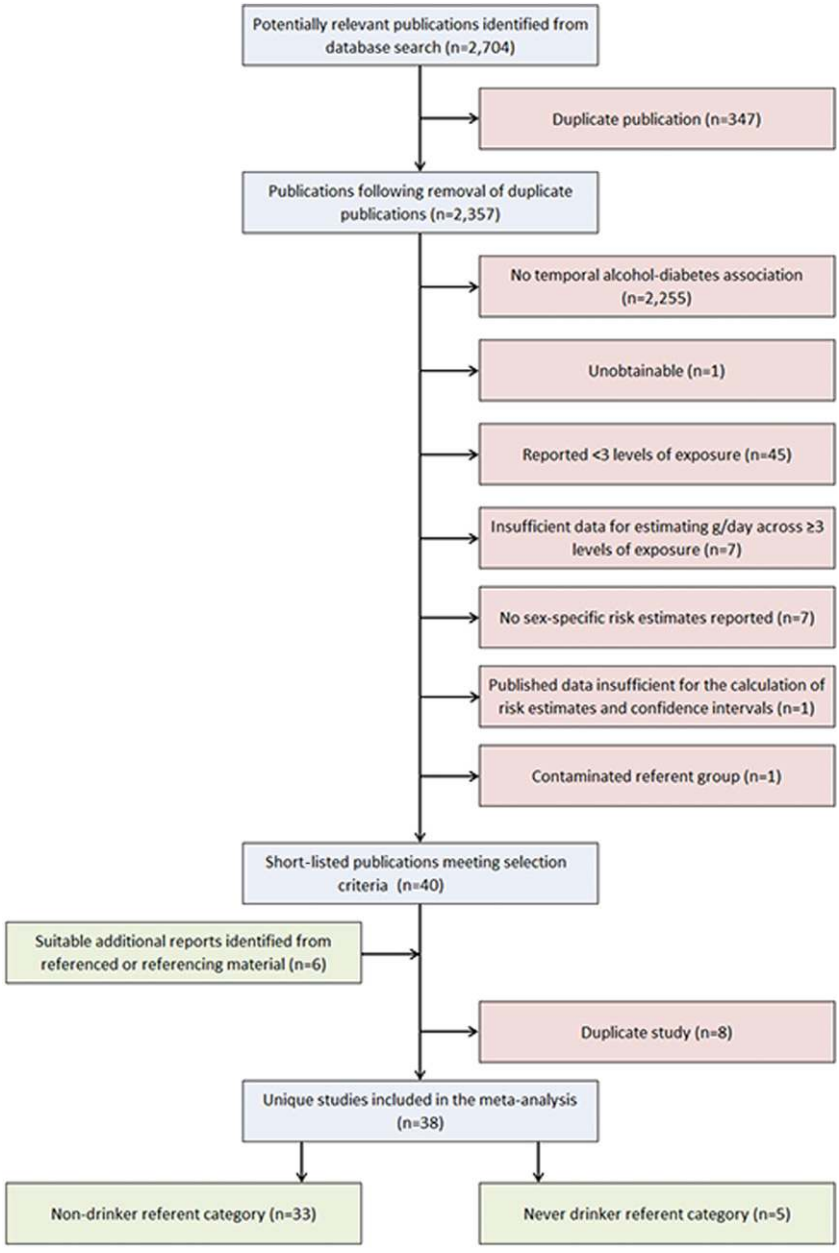

Figure 1—Study flow diagram.

Referent-Specific Data

Few studies used a strictly defined never-drinking category (men: four studies, n = 15,766 [35–38]; women: four studies, n = 98,521 [35–37,39]). The referent interaction was significant (P = 0.005) and improved the fit of the model (P = 0.02). Sex-adjusted, referent-stratified results are displayed in Fig. 4. Consumption relative to never drinkers was associated with no reduction in the risk of type 2 diabetes at any level. By comparison, consumption of <59 g/day showed a reduction in risk relative to noncurrent drinkers.

Sex-Specific Data (Never-Drinking Studies Only)

Having identified significant differences in dose-response by both sex and referent group, sex-specific data from the five studies using a strictly defined

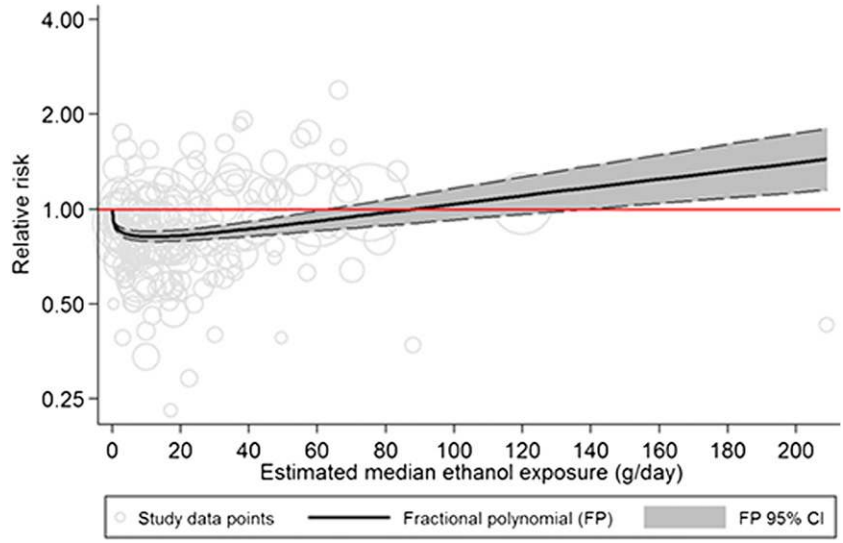

Figure 2—Dose-response relationship between average daily alcohol consumption and incident type 2 diabetes.

never-drinking abstention category are reported in Supplementary Fig. 1.

Compared with the model reporting all sex-specific data combined (Fig. 3), restricted analyses showed similar results but with greater imprecision. Consumption among men showed no reduction in risk at any level of exposure, with decreases specific to women and present across a narrower range of exposure (<61 g/day).

Small-Study Effects

Funnel plots showed notable asymmetry among female data points, with the majority of smaller studies reporting a greater degree of risk reduction than the summary estimate, relative to pooled nondrinkers (data not shown). Given the recommendation that only a simple inverse variance weight be used when deriving the summary estimate, asymmetry was likely the product of a large Korean study, which provided 65% of participant data and reported a lower degree of risk reduction than most other studies (40). The impact of the Korean study upon modeled dose-response curves was diminished after the addition of a random-effects weighting component in the primary analyses undertaken for this article (Supplementary Fig. 2).

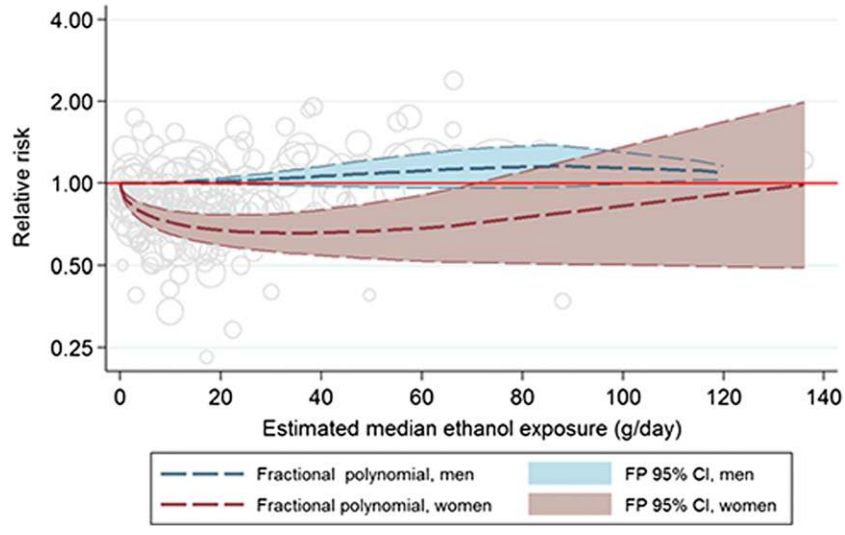

Figure 3—Dose-response relationship between average daily alcohol consumption and incident type 2 diabetes, stratified by sex. FP, fractional polynomial.

Quality Assessment

The quality of selected studies ranged from three to nine points out of nine, with a median score of six (Supplementary Table 3). Such a finding indicated broad discrepancies in study quality, with studies being of moderate quality on average. Sex- and referent-adjusted stratification according to whether data were derived from a study with a score below the median value showed little difference in the dose-response relationship between both groups (Supplementary Fig. 3).

Putative Sources of Heterogeneity

Method of case ascertainment was summarized as participant self-report (n = 11), objective ascertainment (n = 21), or a combination thereof (n = 6). Given the small number of studies to have used both methods, attention was focused upon the subset using either self-reported or objective outcome data. The sex- and referent-adjusted dose-response relationship of the 32 applicable studies was stratified according to these two categories of case ascertainment. Stratified sex- and referent-adjusted analyses showed a less pronounced reduction in risk among studies using objective outcome data compared with those that used self-reported case ascertainment (Supplementary Fig. 4).

The next factor thought to explain some degree of the observed between-study heterogeneity was whether data were extracted from an occupational (n = 12) or general population (n = 26) cohort. Although CIs overlapped along the length of the fitted curves, effect estimates extracted from occupational cohorts appeared to show greater levels of risk reduction (Supplementary Fig. 5).

A total of 15 studies reported crude or age-adjusted estimates (n = 14), with 23 studies providing multivariable-adjusted data (n = 24). Compared with a model based on crude or age-adjusted data, multivariable-adjusted data appeared to show less reduction in risk at moderate levels of alcohol consumption but with reductions in risk present across a broader range of exposure (Supplementary Fig. 6). This relationship was little changed when using an alternative confounding variable that defined studies according to whether their degree of adjustment was above or below the mean of four confounding factors.

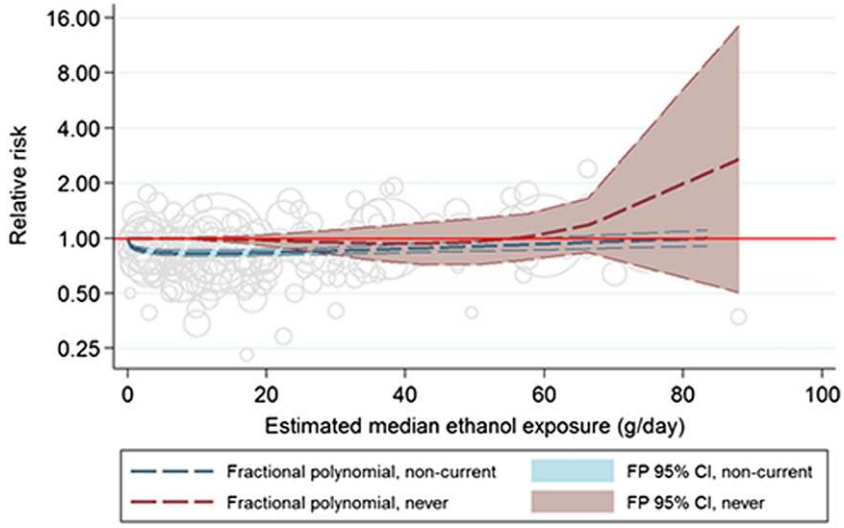

Figure 4—Dose-response relationship between average daily alcohol consumption and incident type 2 diabetes, stratified by referent category and adjusted for sex. FP, fractional polynomial.

Finally, data were stratified according to whether effect estimates were extracted from an Asian (n = 13) or non-Asian (n = 25) population. No reduction in risk was found within data drawn from Asian populations, with reductions in risk specific to participants from non-Asian regions (Supplementary Fig. 7).

CONCLUSIONS

The updated and expanded meta-analysis showed no reduction in type 2 diabetes risk at any level of alcohol consumption among men, regardless of reference group. This is in contrast to a 2009 meta-analysis, which reported peak reduction in risk among men at 22 g/day (RR 0.87, 95% CI 0.76–1.00), relative to quasi-never drinkers (9). For exploration of this discrepancy, male data were stratified according to whether they had been included in the 2009 meta-analysis (Supplementary Fig. 8). These stratified dose-response data indicate that reductions in risk among lighter drinkers were particular to the studies sampled by the 2009 meta-analysis; among the 20 new studies added as part of the updated meta-analysis, no reduction in risk was present at any level of alcohol consumption, relative to pooled nondrinkers. Such a finding hints at marked heterogeneity between the two groups of publications. On the basis of supplementary analyses that investigated potential sources of heterogeneity (Supplementary Figs. 3–7), the absence of any reduction in risk among newly sampled studies would be expected were they more likely to have sampled data from Asian populations or used objective methods of case ascertainment.

Reductions in risk appeared to be specific to women, who exhibited a decreased risk of type 2 diabetes at <71 g/day and peak reduction of 34% at 31–37 g/day, relative to combined abstainers (current nondrinkers and never drinkers).

A reduction in risk being specific to female drinkers may be attributable to a number of factors. Firstly, female never drinkers may be less healthy than their male equivalents. Although research concerning the health status of never drinkers is lacking, a recent study analyzing data from the 1958 National Child Development Study found that, of participants to consistently report long-standing illness from the age of 23 years, women were significantly more likely to report being never drinkers at ages 33 and 42 years (41). Such data hint at the possibility that risk factors for type 2 diabetes may be disproportionately distributed between the sexes—a problem particularly pronounced for any estimates drawn from poorly adjusted studies. However, no sex-specific differences were identified in the average number of covariates adjusted for among selected studies.

Secondly, exposure data analyzed as part of this meta-analysis concerned average volume intake over a given time and therefore did not capture the effect of episodic drinking behaviors upon the risk of type 2 diabetes. The importance of such a consideration is well illustrated, such as in a recent meta-analysis of ischemic heart disease (42). While a 36% reduction in risk was identified among moderate drinkers (<30 g/day), no reduction was evident among moderate drinkers who also reported heavy episodic consumption. This analysis mirrored findings from earlier studies (43,44) and suggests that a higher degree of heavy episodic drinking among men may go some way toward explaining observed sex-specific differences in the alcohol-diabetes relationship. Data collated from 172 European general practices appear to support such a possibility, with the multivariable-adjusted odds of heavy episodic drinking being more than four times greater among men than women (45).

Thirdly, putative biological pathways may operate differently between men and women, such as the effect of alcohol consumption on insulin sensitivity. After an analysis of results reported by 14 intervention studies, alcohol consumption was associated with reduced fasting insulin concentrations and improved insulin sensitivity among women only (13). However, findings from such intervention studies should be interpreted with caution owing to their small size, heterogeneous designs and populations, and often conflicting results (46).

Fourthly, sex-specific differences in the dose-response relationship may have been attributable in part to disparities in the characteristics of studies from which male and female data were extracted, with 84.1% of male participants and 57.6% of female participants having been sampled from studies of Asian populations and 13.6% and 34.1% of male and female participants having been sampled from studies using self-reported methods of case ascertainment. Supplementary analyses reported as part of this meta-analysis indicate that such factors may have an effect upon degree of observed risk reduction. For instance, reductions in risk were found to be particular to non-Asian populations (Supplementary Fig. 7), which might be expected given impairments to alcohol metabolism (47) and a heightened genetic susceptibility among Asian populations to the development of type 2 diabetes (48). Furthermore, relative to studies using objective measures of case ascertainment, reductions in risk were greatest among those that relied upon self-reported measures (Supplementary Fig. 4). However, although

the data presented in Supplementary Fig. 4 suggest that self-reported data may introduce an underestimation of diabetes risk (49), recent studies have found self-reported methods of case ascertainment to be valid and appropriate for use in epidemiological studies (50,51).

Strengths

This meta-analysis benefited from the addition of 18 studies published since 2008 or otherwise missed or discounted during previous meta-analyses. This equated to an additional 1,425,356 participants and 113,370 cases, relative to the last published meta-analysis in 2009 (9).

While the previous meta-analysis may have adopted a never-drinking referent category for the determination of risk among exposed participants, it afforded only an approximation of risk by weighting effect estimates relative to nondrinkers according to the sex-specific prevalence of former drinkers reported by five of the 20 selected studies to have reported never and former drinkers separately. This approach assumed that the proportion of former drinkers contained within a nondrinking category could be reliably estimated according to those reported by five studies, with differences in the proportion of former drinkers explained by sex alone. Furthermore, just two of the five selected studies had strictly defined never drinking as lifelong abstention.

Contrary to this approach, we explicitly tested for a referent group interaction and, having identified a significant difference in the dose-response relationship according to the choice of referent group, sought to stratify risk estimates by abstention category (Fig. 4).

Limitations

Heterogeneity between sampled studies was high, complicating interpretation. Factors likely to have contributed to between-study differences in dose response were thought to include participant age, method of exposure and case ascertainment, follow-up duration, the healthy worker effect of occupational cohorts, ethnicity, and both the degree and type of confounder adjustment. For instance, more than one-third (39%) of selected studies provided crude or age-adjusted data, while just six studies (16%) gave consideration to the effect of dietary factors.

Where the risk of aggregation bias was low in the absence of individual-level data (34), these likely sources of heterogeneity were explored visually via the stratification of dose-response curves. The resulting supplementary analyses (Supplementary Figs. 3–7) confirmed that reductions in risk were lowest among studies with greater levels of confounder adjustment and suggest that future studies exploring the alcohol-diabetes relationship should give greater consideration to the role of confounding factors.

The use of meta-regression to formally and jointly test of differences in dose response according to putative sources of heterogeneity was avoided owing to the potential for low statistical power relative to regressions of individual-level data, even when effect sizes and the number of studies are large (34,52,53). While it has been suggested that statistical power may be sufficient in instances where the number of covariates does not exceed a ratio of 1 to every 10 studies (54), simulations suggest that power is especially low when heterogeneity is high (55).

Although the quality of selected studies was assessed using the Newcastle-Ottawa assessment scale (27), such tools are subject to notable limitations. For instance, while a wide range of instruments have thus far been devised for the assessment of nonrandomized studies, each comprises assessment criteria that are disparate in both number and nature (56). In addition to the use of different rating scales or summary scores that risk weighing the importance of component items in ways not directly related to their impact upon the internal validity of a given study (57,58), their contrasting construction is such that the choice of tool may have a large bearing upon the assessment of study quality. Alongside the effect of such factors upon the interpretation of results derived from a quality assessment instrument, the Newcastle-Ottawa tool has received particular criticism. These criticisms range from the tool's focus upon the generalizability of a given sample to the general population as opposed to its internal validity (59) to the arbitrary nature of some questions that appear to weaken interrater reliability (60,61). With these limitations in mind, the Newcastle-Ottawa quality assessment tool should be considered only as a rough guide for readers as opposed to a definitive measure of study quality.

A further shortcoming rests with the limited number of studies to have explicitly separated former drinkers from strictly defined never drinkers. Caution should be applied when drawing inferences based upon analyses from just five unique studies that reported the dose-response relationship by referent group.

Regardless of the referent category selected, sampled studies consistently relied upon self-reported alcohol consumption data, which is known to substantially underreport the amount of alcohol sold owing to factors such as questionnaire design (62) and a range of cognitive biases (63). In addition, by relying upon only a single cross-sectional self-report of alcohol consumption, sampled studies did not consider the effect of temporal changes in alcohol consumption both during the length of study and prior to study initiation. The assumption of stable temporal consumption is likely to be invalid, with disparate trajectories of alcohol consumption consistently identified regardless of the length of follow-up or the age of the cohort under study (64,65).

Conclusion

Dose-response analyses exploring the association between alcohol consumption and incident type 2 diabetes have typically identified a reduction in risk at relatively moderate levels of exposure among both men and women. By contrast, the primary analyses undertaken as part of this meta-analysis suggest that reductions in risk at moderate levels of alcohol consumption drinkers may be confined to women, with a series of sex-adjusted supplementary analyses indicating that reductions in risk may be greatest among studies that used self-reported methods of case ascertainment or sampled individuals from non-Asian populations.

In addition, the analyses also hinted at the possibility that many existing studies may have overestimated the degree to which the risk of type 2 diabetes is reduced among moderate consumers of alcohol, with reductions in risk appearing to be specific to studies using a noncurrent drinking referent category. Unfortunately, very few studies have excluded less healthy former drinkers from the abstention category, limiting the inferences than can be drawn from the stratification of data by abstention group.

Further research is now required to better understand sex-specific differences in the dose-response relationship between alcohol consumption and type 2 diabetes. Such research will be aided by the application of detailed trajectory-based analyses

capable of modeling the effect of changes to alcohol exposure as a function of time. Until then, however, policy-makers, medical professionals, and the general public should apply caution before considering moderate alcohol consumption as conferring individuals with a reduction in metabolic risk.

Funding. C.K., S.B., and A.B. are funded by the European Research Council (ERC-StG-2012-309337_AlcoholLifecourse; principal investigator A.B. [http://www.ucl.ac.uk/alcohol-lifecourse]) and the U.K. Medical Research Council/Alcohol Research UK (MR/M006638/1).

The funders had no role in study design, data collection or analysis, decision to publish, or preparation of the manuscript. The views expressed are those of the authors and not necessarily those of the funders.

Duality of Interest. No potential conflicts of interest relevant to this article were reported.

Author Contributions. C.K., S.B., and A.B. designed the study. C.K. undertook the literature search. C.K., S.B., and A.B. were involved in the short-listing of identified studies. Data extraction and analysis were undertaken by C.K. with input from S.B. and A.B. All authors contributed to the final manuscript.

References

1. Sarwar N, Gao P, Seshasai SR, et al.; Emerging Risk Factors Collaboration. Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: a collaborative meta-analysis of 102 prospective studies. Lancet 2010;375:2215–2222
2. Vinik AI, Maser RE, Mitchell BD, Freeman R. Diabetic autonomic neuropathy. Diabetes Care 2003;26:1553–1579
3. Zhang P, Zhang X, Brown J, et al. Global healthcare expenditure on diabetes for 2010 and 2030. Diabetes Res Clin Pract 2010;87:293–301
4. Willi C, Bodenmann P, Ghali WA, Faris PD, Cornuz J. Active smoking and the risk of type 2 diabetes: a systematic review and meta-analysis. JAMA 2007;298:2654–2664
5. Kodama S, Horikawa C, Fujihara K, et al. Comparisons of the strength of associations with future type 2 diabetes risk among anthropometric obesity indicators, including waist-to-height ratio: a meta-analysis. Am J Epidemiol 2012;176:959–969
6. Salmerón J, Hu FB, Manson JE, et al. Dietary fat intake and risk of type 2 diabetes in women. Am J Clin Nutr 2001;73:1019–1026
7. Willett W, Manson J, Liu S. Glycemic index, glycemic load, and risk of type 2 diabetes. Am J Clin Nutr 2002;76:274S–280S
8. Meyer KA, Kushi LH, Jacobs DR Jr, Slavin J, Sellers TA, Folsom AR. Carbohydrates, dietary fiber, and incident type 2 diabetes in older women. Am J Clin Nutr 2000;71:921–930
9. Baliunas DO, Taylor BJ, Irving H, et al. Alcohol as a risk factor for type 2 diabetes: a systematic review and meta-analysis. Diabetes Care 2009;32:2123–2132
10. Akash MS, Rehman K, Chen S. Role of inflammatory mechanisms in pathogenesis of type 2 diabetes mellitus. J Cell Biochem 2013;114:525–531
11. Brien SE, Ronksley PE, Turner BJ, Mukamal KJ, Ghali WA. Effect of alcohol consumption on biological markers associated with risk of coronary heart disease: systematic review and meta-analysis of interventional studies. BMJ 2011;342:d636
12. Szabo G, Mandrekar P, Catalano D. Inhibition of superantigen-induced T cell proliferation and monocyte IL-1 beta, TNF-alpha, and IL-6 production by acute ethanol treatment. J Leukoc Biol 1995;58:342–350
13. Schrieks IC, Heil AL, Hendriks HF, Mukamal KJ, Beulens JW. The effect of alcohol consumption on insulin sensitivity and glycemic status: a systematic review and meta-analysis of intervention studies. Diabetes Care 2015;38:723–732
14. Fillmore KM, Stockwell T, Chikritzhs T, Bostrom A, Kerr W. Moderate alcohol use and reduced mortality risk: systematic error in prospective studies and new hypotheses. Ann Epidemiol 2007;17(Suppl.):S16–S23
15. Shaper AG. Alcohol and mortality: a review of prospective studies. Br J Addict 1990;85:837–847; discussion 849–861
16. Tsubono Y, Yamada S, Nishino Y, Tsuji I, Hisamichi S. Choice of comparison group in assessing the health effects of moderate alcohol consumption. JAMA 2001;286:1177–1178
17. Chikritzhs T, Fillmore K, Stockwell T. A healthy dose of scepticism: four good reasons to think again about protective effects of alcohol on coronary heart disease. Drug Alcohol Rev 2009;28:441–444
18. Di Castelnuovo A, Costanzo S, Bagnardi V, Donati MB, Iacoviello L, de Gaetano G. Alcohol dosing and total mortality in men and women: an updated meta-analysis of 34 prospective studies. Arch Intern Med 2006;166:2437–2445
19. Knott CS, Coombs N, Stamatakis E, Biddulph JP. All cause mortality and the case for age specific alcohol consumption guidelines: pooled analyses of up to 10 population based cohorts. BMJ 2015;350:h384
20. World Health Organization. *Use of Glycated Haemoglobin (HbA1c) in the Diagnosis of Diabetes Mellitus. Abbreviated Report of a WHO Consultation*. Geneva, World Health Org., 2011
21. Fleiss JL. *Statistical Methods for Rates and Proportions*. New York, John Wiley & Sons, 1981, p. 229–232
22. House of Commons Science and Technology Committee. *Alcohol Guidelines. Eleventh Report of Session 2010–12: Report, Together with Formal Minutes, Oral and Written Evidence*. Vol. 1. London, The Stationery Office, 2012
23. Zhang J, Yu KF. What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. JAMA 1998;280:1690–1691
24. A'Court C, Stevens R, Heneghan C. Against all odds? Improving the understanding of risk reporting. Br J Gen Pract 2012;62:e220–e223
25. Hamling J, Lee P, Weitkunat R, Ambühl M. Facilitating meta-analyses by deriving relative effect and precision estimates for alternative comparisons from a set of estimates presented by exposure level or disease category. Stat Med 2008;27:954–970
26. Greenland S, Longnecker MP. Methods for trend estimation from summarized dose-response data, with applications to meta-analysis. Am J Epidemiol 1992;135:1301–1309
27. Wells GA, Shea B, O'Connell D, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomized studies in meta-analyses [Internet], 2014. Available from http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp. Accessed 14 March 2015
28. Royston P, Altman DG. Regression using fractional polynomials of continuous covariates: parsimonious parametric modelling (with discussion). Appl Stat 1994;43:425–467
29. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR, Eds. *Introduction to Meta-Analysis*. Chicheser, U.K., John Wiley & Sons, 2009
30. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327:557–560
31. Sterne JA, Harbord RM. Funnel plots in meta–analysis. Stata J 2004;4:127–141
32. Eisen EA, Robins JM, Picciotto S. Healthy worker effect in occupational studies. In *Encyclopedia of Environmetrics*. El-Shaarawi AH, Piegorsch W, Eds. Chichester, U.K., John Wiley & Sons, 2012, p. 1269–1272
33. Oldroyd J, Banerjee M, Heald A, Cruickshank K. Diabetes and ethnic minorities. Postgrad Med J 2005;81:486–490
34. Berlin JA, Santanna J, Schmid CH, Szczech LA, Feldman HI; Anti-Lymphocyte Antibody Induction Therapy Study Group. Individual patient- versus group-level data meta-regressions for the investigation of treatment effect modifiers: ecological bias rears its ugly head. Stat Med 2002;21:371–387
35. Kao WH, Puddey IB, Boland LL, Watson RL, Brancati FL. Alcohol consumption and the risk of type 2 diabetes mellitus: atherosclerosis risk in communities study. Am J Epidemiol 2001;154:748–757
36. Carlsson S, Hammar N, Grill V, Kaprio J. Alcohol consumption and the incidence of type 2 diabetes: a 20-year follow-up of the Finnish twin cohort study. Diabetes Care 2003;26:2785–2790
37. Burke V, Zhao Y, Lee AH, et al. Predictors of type 2 diabetes and diabetes-related hospitalisation in an Australian Aboriginal cohort. Diabetes Res Clin Pract 2007;78:360–368
38. Heianza Y, Arase Y, Saito K, et al. Role of alcohol drinking pattern in type 2 diabetes in Japanese men: the Toranomon Hospital Health Management Center Study 11 (TOPICS 11). Am J Clin Nutr 2013;97:561–568
39. Wannamethee SG, Camargo CA Jr, Manson JE, Willett WC, Rimm EB. Alcohol drinking patterns and risk of type 2 diabetes mellitus among younger women. Arch Intern Med 2003;163:1329–1336
40. Jee SH, Foong AW, Hur NW, Samet JM. Smoking and risk for diabetes incidence and mortality in Korean men and women. Diabetes Care 2010;33:2567–2572
41. Ng Fat L, Cable N, Marmot MG, Shelton N. Persistent long-standing illness and non-drinking over time, implications for the use of lifetime abstainers as a control group. J Epidemiol Community Health 2014;68:71–77
42. Roerecke M, Rehm J. Alcohol consumption, drinking patterns, and ischemic heart disease:

a narrative review of meta-analyses and a systematic review and meta-analysis of the impact of heavy drinking occasions on risk for moderate drinkers. BMC Med 2014;12:182
43. Roerecke M, Rehm J. Irregular heavy drinking occasions and risk of ischemic heart disease: a systematic review and meta-analysis. Am J Epidemiol 2010;171:633–644
44. Ruidavets JB, Ducimetière P, Evans A, et al. Patterns of alcohol consumption and ischaemic heart disease in culturally divergent countries: the Prospective Epidemiological Study of Myocardial Infarction (PRIME). BMJ 2010;341:c6077
45. Nazareth I, Walker C, Ridolfi A, et al. Heavy episodic drinking in Europe: a cross section study in primary care in six European countries. Alcohol Alcohol 2011;46:600–606
46. Hendriks HFJ. Moderate alcohol consumption and insulin sensitivity: observations and possible mechanisms. Ann Epidemiol 2007; 17(Suppl. 1):S40–S42
47. Edenberg HJ. The genetics of alcohol metabolism: role of alcohol dehydrogenase and aldehyde dehydrogenase variants. Alcohol Res Health 2007;30:5–13
48. Cho YS, Lee JY, Park KS, Nho CW. Genetics of type 2 diabetes in East Asian populations. Curr Diab Rep 2012;12:686–696
49. Molenaar EA, Van Ameijden EJ, Grobbee DE, Numans ME. Comparison of routine care self-reported and biometrical data on hypertension and diabetes: results of the Utrecht Health Project. Eur J Public Health 2007;17:199–205
50. Schneider AL, Pankow JS, Heiss G, Selvin E. Validity and reliability of self-reported diabetes in the atherosclerosis risk in communities study. Am J Epidemiol 2012;176:738–743
51. Margolis KL, Lihong Qi, Brzyski R, et al.; Women Health Initiative Investigators. Validity of diabetes self-reports in the Women's Health Initiative: comparison with medication inventories and fasting glucose measurements. Clin Trials 2008;5:240–247
52. Lambert PC, Sutton AJ, Abrams KR, Jones DR. A comparison of summary patient-level covariates in meta-regression with individual patient data meta-analysis. J Clin Epidemiol 2002; 55:86–94
53. Schmid CH, Stark PC, Berlin JA, Landais P, Lau J. Meta-regression detected associations between heterogeneous treatment effects and study-level, but not patient-level, factors. J Clin Epidemiol 2004;57:683–697
54. Baker WL, White CM, Cappelleri JC, Kluger J, Coleman CI; Health Outcomes, Policy, and Economics (HOPE) Collaborative Group. Understanding heterogeneity in meta-analysis: the role of meta-regression. Int J Clin Pract 2009; 63:1426–1434
55. Higgins JP, Thompson SG. Controlling the risk of spurious findings from meta-regression. Stat Med 2004;23:1663–1682
56. Sanderson S, Tatt ID, Higgins JP. Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. Int J Epidemiol 2007;36:666–676
57. Jüni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. JAMA 1999;282:1054–1060
58. Greenland S, O'Rourke K. On the bias produced by quality scores in meta-analysis, and a hierarchical view of proposed solutions. Biostatistics 2001;2:463–471
59. Stang A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. Eur J Epidemiol 2010;25:603–605
60. Hartling L, Milne A, Hamm MP, et al. Testing the Newcastle Ottawa Scale showed low reliability between individual reviewers. J Clin Epidemiol 2013;66:982–993
61. Oremus M, Oremus C, Hall GB, McKinnon MC; ECT & Cognition Systematic Review Team. Inter-rater and test-retest reliability of quality assessments by novice student raters using the Jadad and Newcastle-Ottawa Scales. BMJ Open 2012;4:2
62. Stockwell T, Donath S, Cooper-Stanbury M, Chikritzhs T, Catalano P, Mateo C. Under-reporting of alcohol consumption in household surveys: a comparison of quantity-frequency, graduated-frequency and recent recall. Addiction 2004;99: 1024–1033
63. Del Boca FK, Darkes J. The validity of self-reports of alcohol consumption: state of the science and challenges for research. Addiction 2003;98(Suppl. 2):1–12
64. Sher KJ, Jackson KM, Steinley D. Alcohol use trajectories and the ubiquitous cat's cradle: cause for concern? J Abnorm Psychol 2011;120:322–335
65. Britton A, Ben-Shlomo Y, Benzeval M, Kuh D, Bell S. Life course trajectories of alcohol consumption in the United Kingdom using longitudinal data from nine cohort studies. BMC Med 2015;13:47