Annals of Epidemiology 23 (2013) 193–197

Contents lists available at SciVerse ScienceDirect

Annals of Epidemiology

journal homepage: www.annalsofepidemiology.org

Estimating the mediating effect of different biomarkers on the relation of alcohol consumption with the risk of type 2 diabetes

Joline W.J. Beulens PhD [a,*], Yvonne T. van der Schouw PhD [a], Karel G.M. Moons PhD [a], Hendriek C. Boshuizen PhD [b], Daphne L. van der A PhD [c], Rolf H.H. Groenwold MD, PhD [a]

[a] Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, The Netherlands
[b] Expertise Centre for Methodology and Information Services, National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands
[c] Centre for Nutrition and Health, National Institute of Public Health and the Environment (RIVM), Bilthoven, The Netherlands

ARTICLE INFO

Article history:
Received 14 September 2012
Accepted 25 December 2012
Available online 29 January 2013

Keywords:
Diabetes
Mediation
Alcohol consumption
Biomarkers

ABSTRACT

Purpose: Moderate alcohol consumption is associated with a reduced type 2 diabetes risk, but the biomarkers that explain this relation are unknown. The most commonly used method to estimate the proportion explained by a biomarker is the difference method. However, influence of alcohol–biomarker interaction on its results is unclear. G-estimation method is proposed to accurately assess proportion explained, but how this method compares with the difference method is unknown.
Methods: In a case–cohort study of 2498 controls and 919 incident diabetes cases, we estimated the proportion explained by different biomarkers on the relation between alcohol consumption and diabetes using the difference method and sequential G-estimation method.
Results: Using the difference method, high-density lipoprotein cholesterol explained the relation between alcohol and diabetes by 78% (95% confidence interval [CI], 41–243), whereas high-sensitivity C-reactive protein (−7.5%; −36.4 to 1.8) or blood pressure (−6.9; −26.3 to −0.6) did not explain the relation. Interaction between alcohol and liver enzymes led to bias in proportion explained with different outcomes for different levels of liver enzymes. G-estimation method showed comparable results, but proportions explained were lower.
Conclusions: The relation between alcohol consumption and diabetes may be largely explained by increased high-density lipoprotein cholesterol but not by other biomarkers. Ignoring exposure–mediator interactions may result in bias. The difference and G-estimation methods provide similar results.

Introduction

Moderate alcohol consumption is consistently associated with a reduced risk of type 2 diabetes [1]. Short-term intervention studies have shown that moderate alcohol consumption improves blood lipid profiles and insulin sensitivity and reduces inflammatory factors [2–5]. These biomarkers might explain the inverse association between moderate alcohol consumption and risk of diabetes. One study has investigated to what extent different biomarkers mediate the effects of alcohol consumption on diabetes [6]. This study showed that inflammatory factors, endothelial dysfunction, and fasting insulin did not explain the relation between alcohol consumption and diabetes, whereas adiponectin concentrations explained about 25% of this relation. These results suggest that other mechanisms like blood lipid profile could be involved.

Like most epidemiologic studies, the study by Beulens et al. [6] used logistic regression to estimate the proportion explained by biomarkers by means of a method called the difference method. In this method, adiponectin is simply added to the model regressing outcome on exposure. Hence, two regression coefficients of the exposure–outcome relation are estimated, that is, one with and one without including the biomarker in the regression model. The (relative) difference between the estimated regression coefficients of the exposure–outcome relation based on the models with and without the biomarker included indicates to what extent the association is mediated by that biomarker. Studies have shown that omitting interaction between exposure and mediators in this analysis may yield biased estimates [7–9]. However, it is uncommon to evaluate such interactions, and the impact of omitting exposure–mediator interaction from the model is typically unknown.

In addition, G-estimation models [9,10] have been proposed to accurately assess the proportion explained. These models can be

Conflict of interest: None.

* Corresponding author. Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, PO Box 85500, 3508 GA Utrecht, The Netherlands.
E-mail address: J.Beulens@umcutrecht.nl (J.W.J. Beulens).

1047-2797/$ – see front matter
http://dx.doi.org/10.1016/j.annepidem.2012.12.014

applied to validly estimate controlled direct and indirect effects, that is, direct effects of the exposure on the outcome and indirect effects through mediators while fixing (or controlling) the value of the mediator [8–11]. How this method compares with the most commonly used difference method in empirical data is not extensively studied.

In this study, we will therefore investigate the biomarkers that explain the relation between alcohol consumption and risk of diabetes and evaluate the impact of exposure–mediator interaction using the difference method. Additionally, we will apply sequential G-estimation method to compare its results with the difference method. We have included biomarkers that have previously been associated with alcohol consumption: blood lipid profile [12], high-sensitivity C-reactive protein (hsCRP) [12], liver enzymes [13,14], and blood pressure [15].

Materials and methods

Design and study population

EPIC-NL comprises the two Dutch contributions to the European Prospective Investigation into Cancer and Nutrition (EPIC) study: Prospect-EPIC and MORGEN-EPIC set up simultaneously in 1993–1997. Its design and rationale are described elsewhere [16]. The Prospect-EPIC study includes 17,357 women, aged 49–70 years at baseline, who participated in the national breast cancer screening program and were living in the city of Utrecht and its surroundings [17]. The MORGEN-EPIC cohort consists of 22,654 men and women aged 21–64 years selected from random samples of the Dutch population in three towns in the Netherlands (Amsterdam, Doetinchem, and Maastricht). All participants gave written informed consent. The study complied with the Declaration of Helsinki and was approved by the local ethical committees.

A 6.5% random sample of the baseline cohort was taken for the measurement of biomarkers, using the efficient case–cohort design. This study is performed using the baseline random sample ($n = 2604$) and all incident cases of type 2 diabetes ($n = 924$; of which 79 were included in the random sample). After exclusion of prevalent diabetes cases ($n = 43$), individuals with missing alcohol consumption ($n = 10$) or follow-up information ($n = 58$), 2419 controls from the random sample, and 919 incident diabetes cases were used for the analysis.

Assessment of alcohol intake, type 2 diabetes, mediators, and other covariates

At baseline, participants filled in a general questionnaire and food frequency questionnaire. Alcohol consumption was assessed by the general questionnaire and the food frequency questionnaire. Details and validity of this assessment have been described elsewhere [18,19]. The occurrence of diabetes during follow-up was ascertained mainly through self-report via follow-up questionnaires, a urine dipstick test, and linkage to a hospital discharge diagnoses registry. Follow-up was completed until January 1, 2006. Potential diabetes cases were verified against medical or pharmacy records, and 72% of these cases were verified as type 2 diabetes and used for the analysis [20].

At baseline, blood samples were drawn and stored for future use. HbA1c, blood lipid profile (total cholesterol, high-density lipoprotein [HDL] and low-density lipoprotein [LDL] cholesterol, and triglycerides), hsCRP, and liver enzymes (aspartate aminotransferase [AST], alanine aminotransferase [ALT], and γ-glutamyl transferase [GGT]) were assessed in the random sample and in all incident cases of type 2 diabetes as previously described [16].

Other covariates were obtained from the general questionnaire. At baseline, body weight, height, and waist and hip circumferences were measured. Physical activity was assessed using a questionnaire, and the Cambridge Physical Activity Index was used to categorize participants as follows: inactive, moderately inactive, moderately active, and active [21]. Because we could not calculate a total physical activity score for 14% of all participants, we imputed missing scores by single imputation using linear regression modeling (SPSS MVA procedure; SPSS, Inc., Chicago, IL) with other lifestyle factors (e.g., smoking and body mass index [BMI]) and the outcome (type 2 diabetes). In the Prospect-EPIC study, systolic and diastolic blood pressures (DBPs) were measured twice by a trained observer with an automated and calibrated Oscillomat (Bosch & Son, Jungingen, Germany) in supine position, and the mean was calculated. In the MORGEN-EPIC study, the measurement of systolic and DBPs was performed twice by a trained observer using a Random Zero Sphygmomanometer in supine position, and the mean was calculated.

Statistical analyses

Missing data for general covariates were lower than 5%. For most biomarkers, missing data were approximately 6%, except for hsCRP (12.0%) and HDL and LDL cholesterol (8.5%). We used single imputation modeling to impute these missing data (SPSS MVA procedure; SPSS, Inc., Chicago, IL).

Using backward selection (based on the Akaike information criterion), we selected covariates for adjustment. All models were adjusted for the following covariates: age, sex, BMI, smoking status, educational status, family history of diabetes, and energy and fat intake. We only evaluated mediation effects for those potential mediators that were related to alcohol intake in a linear regression: ALT, AST, GGT, hsCRP (log transformed because of deviation from the normal distribution), triglycerides, HDL cholesterol, DBP, and HbA1c. We estimated the effect of these potential mediators on the association between alcohol consumption and diabetes cumulative incidence using logistic regression (primary end point, dichotomous). Alcohol consumption was modeled linearly per 10-g increment. We investigated whether a significant deviation from the linear model was present by including the quadratic term of alcohol consumption. Because this term was not significant ($P = .57$), alcohol consumption was modeled linearly per 10-g increment in all further analyses. We further investigated whether an interaction between alcohol consumption and gender was present by including the interaction term. As this did not suggest a differential effect ($P = .19$), men and women were combined in further analyses.

Mediation effects were estimated using two different approaches. In the first approach, two regression models were fitted: (1) a model in which the outcome was regressed on alcohol consumption and the abovementioned covariates and (2) a model in which the outcome was regressed on alcohol consumption, the abovementioned covariates, and one of the potential mediators. We applied a logistic regression model for the binary outcome diabetes. Model (1) yields the overall (or total) effect of alcohol consumption on diabetes, whereas model (2) yields the direct effect of alcohol consumption on diabetes (i.e., the conditional direct effect). The proportion of the effect (PE) of alcohol consumption that is explained by the mediator was defined as PE = (total effect − direct effect)/total effect, in which the effect is the log(odds ratio [OR]) of the association between alcohol consumption and diabetes as estimated by the two aforementioned models. In epidemiology, this approach is often referred to as the "difference method," and under a rare outcome assumption, this method yields approximately the same results as the Baron–Kenny method [9].

However, in the case of interaction between exposure and mediator, the difference method may yield biased estimates of the