ORIGINAL RESEARCH

Urinary Ethyl Glucuronide as Measure of Alcohol Consumption and Risk of Cardiovascular Disease: A Population-Based Cohort Study

Inge A. T. van de Luitgaarden, MD; Ilse C. Schrieks, PhD; Lynanne M. Kieneker, PhD; Daan J. Touw, PharmD, PhD; Adriana J. van Ballegooijen, PhD; Sabine van Oort, MD; Diederick E. Grobbee, MD, PhD; Kenneth J. Mukamal, MD, MPH; Jenny E. Kootstra-Ros, PhD; Anneke C. Muller Kobold, PhD; Stephan J. L. Bakker, PhD; Joline W. J. Beulens, PhD

BACKGROUND: Moderate alcohol consumption has been associated with a lower risk of cardiovascular disease (CVD) and all-cause mortality compared with heavy drinkers and abstainers. To date, studies have relied on self-reported consumption, which may be prone to misclassification. Urinary ethyl glucuronide (EtG) is an alcohol metabolite and validated biomarker for recent alcohol consumption. We aimed to examine and compare the associations of self-reported alcohol consumption and EtG with CVD and all-cause mortality.

METHODS AND RESULTS: In 5676 participants of the PREVEND (Prevention of Renal and Vascular End-Stage Disease) study cohort, EtG was measured in 24-hour urine samples and alcohol consumption questionnaires were administered. Participants were followed up for occurrence of first CVD and all-cause mortality. Cox proportional hazards regression models, adjusted for age, sex, and CVD risk factors, were fitted for self-reported consumption, divided into 5 categories: abstention, 1 to 4 units/month (reference), 2 to 7 units/week, 1 to 3 units/day, and ≥4 units/day. Similar models were fitted for EtG, analyzed as both continuous and categorical variables. Follow-up times differed for CVD (8 years; 385 CVD events) and all-cause mortality (14 years; 724 deaths). For both self-reported alcohol consumption and EtG, nonsignificant trends were found toward J-shaped associations between alcohol consumption and CVD, with higher risk in the lowest (hazard ratio for abstention versus 1–4 units/month, 1.42; 95% CI, 1.02–1.98) and highest drinking categories (hazard ratio for ≥4 units/day versus 1–4 units/month, 1.11; 95% CI, 0.68–1.84). Neither self-report nor EtG was associated with all-cause mortality.

CONCLUSIONS: Comparable associations with CVD events and all-cause mortality were found for self-report and EtG. This argues for the validity of self-reported alcohol consumption in epidemiologic research.

Key Words: alcohol consumption ■ biomarker ■ cardiovascular disease ■ epidemiologic research ■ ethyl glucuronide

Alcohol consumption is among the most frequently studied risk factors for the development of chronic diseases.1–3 Observational research suggests that the relation between alcohol consumption and cardiovascular disease (CVD) follows a J-shaped curve, indicating that moderate alcohol consumers have a lower cardiovascular risk compared with both abstainers and heavier drinkers.2,4–7 To date, the cardioprotective effects of moderate alcohol consumption remain debated, mainly because the data stem almost exclusively from observational studies, and a long-term randomized controlled trial is lacking. Even mendelian...
randomization studies have failed to provide a single clear answer.\(^8\)\(^{-10}\)

All observational studies of the association between alcohol and CVD have relied on self-report to estimate alcohol consumption. Self-report is a potentially unreliable source of information, with a tendency to underestimate or misclassify consumption.\(^11\) Whether objectively measured alcohol consumption would yield similar results is unknown because reliable objective markers of habitual alcohol consumption are scarce, as most biomarkers either reflect short time periods\(^12\) or are not sufficiently specific.\(^13\) Urinary ethyl glucuronide (EtG) is a relatively new biomarker of alcohol consumption. It is a direct metabolite of ethanol, and thus a specific marker of alcohol consumption, with a detection time up to 72 hours after consumption.\(^14\) EtG has been validated as a marker for alcohol consumption in controlled experiments.\(^15\)\(^{-17}\) Moreover, a previous analysis of our cohort indicated that EtG appears to be linearly associated with self-reported habitual consumption,\(^18\) with particularly high sensitivity for heavier drinking. Specificity was 92% and sensitivity was 66%, increasing up to 93% in the heavier drinking categories.\(^18\) Hence, EtG appears to be a suitable marker to detect abstention and moderate to heavy drinking, in contrast with markers like carbohydrate-deficient transferrin (CDT), which tend to be elevated only in heavy drinking.\(^19\)

In this study, we compared EtG as objective measure of habitual alcohol consumption with self-reported alcohol consumption in the association between alcohol consumption and CVD and all-cause mortality in a prospective population-based cohort. Moreover, by combining information on EtG, CDT, and self-report, we excluded participants with apparently misreported consumption to enhance the validity of their alcohol assessment.

**METHODS**

**Study Population**

The PREVEND (Prevention of Renal and Vascular End-Stage Disease) study cohort is a Dutch cohort drawn from the general population of Groningen, The Netherlands, in 1997, originally established to monitor the long-term development of cardiovascular and renal diseases in participants with microalbuminuria. Details of this study have been published elsewhere.\(^20\)

In short, after exclusion of insulin-dependent subjects and pregnant women, the cohort included 6000 participants with a urinary albumin concentration >10 mg/L. A random sample of 2592 subjects without microalbuminuria was also included. During the study period (1997–2013), participants attended 5 follow-up visits. Follow-up data on mortality were available up until January 2017. The PREVEND study was conducted in accordance with the Declaration of Helsinki guidelines and was approved by the Medical Ethics Committee of the University Medical Center Groningen. All participants gave written informed consent. The data that support the findings of this study are available from the corresponding author upon reasonable request.

In the present study, we included participants who attended the second follow-up visit (N=6894; April 2001–December 2003), as urinary EtG concentrations were measured in urine samples that were collected during this period. The study period comprised the time from this visit until end of follow-up: from April 2001 until January 2017. Follow-up data for CVD events were only available until January 2011, whereas information on all-cause mortality covered the entire study period. Participants without
EtG measurements (N=60) or self-reported alcohol consumption (N=64) were excluded. Moreover, participants were excluded when urinary leukocyte measurements performed with Nephur-test+leuco sticks (Boehringer Mannheim, Mannheim, Germany) showed evidence for a urinary tract infection, defined as the presence of ≥75 leukocytes/μL (N=363) or ≥50 erythrocytes/μL (N=196). Previous research has shown that bacterial contamination can influence EtG concentrations, which can lead to both false-positive and false-negative results.21,22

Finally, participants with prevalent CVD at baseline were excluded (N=443), as well as participants who did not contribute any follow-up time after the baseline visit (N=48) or had missing values for ≥1 of the covariates (N=44). The analytical sample included 5676 participants.

Assessment of Alcohol Consumption
Participants were asked to collect two 24-hour urine samples up to a maximum of 4 days before the baseline visit after thorough oral and written instruction. Participants were asked to avoid heavy exercise and to postpone the urine collection in case of urinary tract infection, menstruation, or fever. Participants stored the samples temporarily at home at a temperature of 4°C before the visit. At the visit, aliquots of these urine specimens were stored at −20°C. EtG concentrations were measured in the second 24-hour urine sample using the Thermo Scientific DRI Ethyl Glucuronide assay. It has a detection limit of 100 ng/mL and has shown good agreement with established liquid chromatography/mass spectrometry methods in detecting EtG.23 Intra-assay and interassay coefficients of variation were previously established at <1.7% and <2.2%, respectively.23 In accordance with previous research,24–26 we used a cutoff value of ≥100 ng/mL to define positivity for intentional alcohol consumption.

Self-reported alcohol consumption was measured with a single question assessing the combined quantity-frequency consumption on the participants’ average usual alcohol consumption at baseline and the first 2 follow-up visits. Participants were asked to choose 1 of the following categories: abstention (no alcohol consumption), 1 to 4 units/month, 2 to 7 units/week, 1 to 3 units/day, or ≥4 units/day. In The Netherlands, a standard serving of an alcoholic beverage contains approximately 10 g of alcohol.27 We assessed whether alcohol consumption remained stable over time, comparing self-reported alcohol consumption at baseline with self-reported consumption at the second follow-up visit. Alcohol consumption was considered stable if a participant did not shift >1 category during total follow-up.

Transferrin and CDT concentrations were measured in serum. Transferrin was analyzed by immunoturbidimetric assay on a Cobas analyzer (Roche Diagnostics GmbH, Mannheim Germany), whereas CDT was analyzed on a BNII nephelometer (Siemens Healthcare GmbH, Marburg, Germany). The transferrin assay is standardized against the reference preparation of the Institute for Reference Materials and Measurements BCR470/CRM470. The obtained intra-assay and interassay coefficients of variation were 1.4 to 1.9 at a level of 1.8 g/L and 1.8% to 1.8% at a level of 2.8 g/L. The detection limit of the assay is 0.1 g/L.

Reference values for CDT were 28.1 to 76.0 mg/L CDT (1st–99th percentile). Intra-assay and interassay coefficients of variation were 2.8% to 4.9% and 1.5% to 7.6%, respectively, depending on the level measured. The detection limit for CDT was 20 mg/L. The percentage CDT was calculated by dividing the CDT concentration on the total transferrin concentration. Reference values for percentage CDT there were 1.19% to 2.47% CDT (1st–99th percentile).28

Primary and Secondary End Points
The primary end point was time to first CVD event. This was composed of cardiac events, cerebrovascular events, and peripheral vascular events. Cardiac events included myocardial infarction, ischemic heart disease, coronary artery bypass grafting, percutaneous transluminal coronary intervention, and death from previously mentioned conditions. We included the following cerebrovascular events: intracranial hemorrhage, subarachnoid hemorrhage, ischemic stroke, transient cerebral ischemia, occlusion of precerebral arteries, and death from these conditions. Peripheral events included bypass surgery of the peripheral arteries, aneurysm, and death from these conditions. Occurrences of CVD events were obtained from PRISMANT, the Dutch National Registry of hospital discharge diagnoses.29 The secondary end point was all-cause mortality, which was ascertained by data linkage with the Dutch Central Bureau of Statistics. Data were coded according to the International Classification of Diseases, Tenth Revision (ICD-10). Mortality was categorized into CVD, cancer, or “other causes” by ICD-10 coding.

Covariates
During the baseline visit, participants were asked to complete questionnaires about lifestyle factors, family history for CVD, medical history, and medication use. Education level was self-reported on the basis of highest ascertainment and stratified according to 3 categories: low (primary education or intermediate vocational education),
middle (higher secondary education), and high (higher vocational education and university). Smoking status was categorized into the following categories: (1) “never smoking,” (2) “former smoking,” (3) “<6 cigarettes/day,” (4) “≥6 to 20 cigarettes/day,” and (5) “≥20 cigarettes/day.” Physical activity was measured as self-reported frequency of exercise and was categorized into 3 categories: (1) “no/hardly,” (2) “less than once a week,” and (3) “twice or more times a week.” Body mass index (BMI) was calculated as measured weight in kilograms divided by the square of height in meters and was categorized into 5 categories: (1) “BMI <20 kg/m²,” (2) “BMI 20 to 22.9 kg/m²,” (3) “BMI 23 to 24.9 kg/m²,” (4) “BMI 25 to 29.9 kg/m²,” and (5) “BMI >30 kg/m².”

We defined type 2 diabetes mellitus (T2DM) as self-reported T2DM, use of antidiabetic medication, or fasting blood glucose at baseline ≥7.0 mmol/L. Hypertension at baseline was defined as self-reported hypertension, use of antihypertensive medication, or a blood pressure at baseline of >140 mm Hg systolic or >90 mm Hg diastolic. Hypercholesterolemia was defined as self-reported hypercholesterolemia, use of cholesterol-lowering drugs, or a total cholesterol level at baseline of ≥6.5 mmol/L. As a measure of kidney function, estimated glomerular filtration rate (eGFR) was calculated using the combined creatinine–cystatin C–based Chronic Kidney Disease Epidemiology Collaboration equation.

Statistical Analysis

All statistical analyses were performed using IBM SPSS 25.0 for Windows and R studio version 3.4.1. Descriptive statistics were used to assess the distribution of the data. Because eGFR was the only variable with considerable missingness (N=288 [5.0%]), we imputed missing values with the mean (93.3 mL/min per 1.73 m²). We excluded small numbers of missing values (<1.1%) for T2DM, hypertension, smoking, and physical activity. We compared self-reported alcohol consumption and EtG, high-density lipoprotein cholesterol, and CDT using Spearman correlation coefficients.

We fitted 3 adjusted Cox proportional hazards models to study the associations of EtG and self-reported alcohol consumption with cardiovascular outcomes, all-cause mortality, and cause-specific mortality. We additionally restricted the analyses to cardiac outcomes. Urinary EtG was assessed as both a continuous variable on a natural logarithmic scale, excluding undetectable EtG, and a categorical variable, which was divided into undetectable EtG concentrations (category 1) and quintiles of detectable EtG concentrations. The second category was considered the reference category to take light drinkers as the referent. As sensitivity analyses, we additionally assessed total excretion of EtG (ie, EtG concentration×urine volume) and EtG/urinary creatinine ratio to correct for urine dilution. Self-reported alcohol consumption was analyzed as a categorical variable, using the 5 consumption categories: abstinence, 1 to 4 units/month, 2 to 7 units/week, 1 to 3 units/day, and ≥4 units/day. The category 1 to 4 units/month was considered the reference category. Model 1 was adjusted for age and sex. Model 2 was adjusted for model 1 and smoking, BMI, physical activity, education level, and family history of CVD. Model 3 contained the same covariates as model 2, but additionally adjusted for T2DM, hypertension, hypercholesterolemia, and kidney function, as these factors were also considered potential mediators in the causal pathway. Age, sex, and eGFR are potential effect modifiers for the association between EtG concentrations and CVD and all-cause mortality. Therefore, these variables and their interaction terms with EtG were separately entered in the model. When suggestive interaction terms (P<0.10) were identified, analyses were stratified accordingly.

We plotted Martingale residuals against age and kidney function to test which functional form of these covariates best fitted the model. We used scaled Schoenfeld residuals to test the proportional hazards assumption. Results are presented as hazard ratios with 95% CIs. We tested for trend by adding the linear term in the model. To assess the presence of nonlinear relationships, we entered the quadratic and cubic terms of EtG with the linear term. If nonlinear relations were found (P<0.05 for quadratic/cubic term), splines were applied to fit different polynomials.

Sensitivity Analyses

As we had multiple measures of self-reported alcohol consumption, we performed a sensitivity analysis excluding participants who reported inconsistent alcohol consumption over time, defined as a shift of >1 category (N=212). As a second sensitivity analysis, we additionally fitted the models with simple time-varying alcohol consumption, using the self-reported alcohol consumption of the baseline visit and the 2 follow-up rounds.

To address misclassification by self-report, we combined information on EtG and CDT concentrations and self-reported alcohol consumption to exclude participants with misreported alcohol consumption. We performed a sensitivity analysis excluding participants with discrepant values for EtG and self-reported consumption, and for CDT and self-reported consumption. To do so, we regressed EtG concentration on self-reported alcohol consumption and excluded participants with the highest and lowest 2.5% residuals (N=146). In addition, we
excluded those participants who reported ≥1 glass of alcohol a day, but had a discrepant EtG concentration <100 ng/mL (N=126). Likewise, self-reported abstainers with EtG concentrations >100 ng/mL were excluded (N=102). Finally, the 5% highest residuals from the regression of CDT on self-report were excluded (N=234). Because heavy drinkers are most prone to underreport their alcohol consumption,24 we additionally excluded participants with CDT values >2.35%, which is the cutoff value for heavy alcohol consumption (N=59)25 (Figure 1).

RESULTS

Participant Characteristics
Among the 5676 eligible participants, mean age at baseline was 52.9 (SD, 11.8) years, and 51.2% were men. Urinary EtG was detected in 52.2% of the samples, consistent with intentional recent alcohol consumption. Urinary EtG concentrations ranged from 0 to 531 900 ng/mL. Abstention from alcohol was reported by 24% of the participants. In general, participants without detectable EtG concentrations were more often women, were slightly older, and reported lower levels of education and more comorbidities, particularly T2DM (Table 1). We observed a similar pattern when self-reported alcohol consumption categories were used (Table S1). Self-reported consumption categories were significantly correlated with EtG ($r_s=0.68; P<0.001$), high-density lipoprotein cholesterol ($r_s=0.11; P<0.001$), and CDT ($r_s=0.25; P<0.001$).

Median follow-up time from baseline until January 2011 was 8.3 years (25–75 percentile, 7.8–8.9 years). In this period, 385 (6.8%) cardiovascular events occurred. Most events were myocardial infarctions (N=102 [1.8%]) or ischemic heart disease (N=77 [1.4%]). Follow-up time for all-cause mortality was available from baseline until January 2017, with a median follow-up time of 14.1 years (25–75 percentile, 11.6–14.7 years). A total of 724 (12.8%) deaths occurred, of which 156 (2.7%) were cardiovascular deaths, 354 (6.2%) were cancer related, 212 (3.7%) were otherwise specified, and 2 (0.03%) were unknown.

Figure 1. Scatterplot of alcohol consumption categories and ethyl glucuronide (EtG) concentrations for 5676 PREVEND (Prevention of Renal and Vascular End-Stage Disease) study participants.
Exclusion of participants with misreported consumption (N=667), on the basis of discrepancies between self-reported consumption and concentrations of biomarkers EtG and carbohydrate-deficient transferrin (CDT). The lowest and highest 2.5% residuals of the regression between EtG and self-reported consumption and the highest 5% residuals of the regression between CDT and self-report were excluded. Moreover, participants who reported abstention, but with EtG concentrations >100 ng/mL, and vice versa were excluded. In addition, heavy drinkers were excluded, on the basis of CDT values. Alcohol consumption categories: 0, abstention; 1, 1 to 4 units/month; 2, 2 to 7 units/week; 3, 1 to 3 units/day; and 4, ≥4 units/day. One standard unit contains 10 g of alcohol.
van de Luitgaarden et al A New Marker for Alcohol in Epidemiologic Research

Alcohol and CVD

The association between self-reported alcohol consumption and CVD appeared to be nonlinear, with a higher CVD risk for the abstention category compared with the reference category of 1 to 4 units/month and a trend toward a higher risk in the heavier alcohol consumption categories. Adjustment for confounders slightly attenuated the associations (Table 2). A similar trend was found when EtG was used as the exposure measure: the lowest and highest categories appeared to be associated with a higher CVD risk compared with the other categories (Table 3). We observed a nonlinear association when ln EtG was tested continuously ($P$ for cubic term=0.04) (Figure 2). There was no effect modification by sex, age, or eGFR. Restricting the analyses to exclusively cardiac events (N=289) yielded similar results (data not shown). The shape of the association remained similar for both total EtG excretion and EtG/creatinine ratio but did not reach statistical significance (data not shown).

Alcohol and All-Cause Mortality

No significant associations were found between self-reported alcohol consumption and all-cause mortality or between EtG and all-cause mortality (Tables 2 and 3). Stratification by cause of death did not alter these results (data not shown). No effect modification by age, sex, or eGFR was found. Similar results were found when EtG was assessed as total EtG excretion and EtG/creatinine ratio.

Sensitivity Analyses

Exclusion of participants who reported unstable alcohol consumption over time, defined as a shift in >1 alcohol consumption category (N=212), did not lead to different associations (Table 4 and Table S2). Inclusion of a time-varying term for alcohol consumption demonstrated a lower mortality risk in the 1 to 3 units/day group compared with the 1 to 4 units/month group, which appeared to be driven by mortality other than cardiovascular or

Table 1. Baseline Characteristics of 5676 PREVEND Study Participants, by EtG Category

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Undetectable EtG (&lt;100 ng/mL)</th>
<th>Detectable EtG (≥100 ng/mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N (%)</td>
<td>Category 1</td>
<td>Category 2</td>
</tr>
<tr>
<td>Informations on EtG</td>
<td>2716(47.9)</td>
<td>592(10.4)</td>
</tr>
<tr>
<td>EtG level</td>
<td>0.0±0</td>
<td>350±186</td>
</tr>
<tr>
<td>Eto level</td>
<td>0 (0; 0)</td>
<td>320 (185:489)</td>
</tr>
<tr>
<td>Range of EtG level</td>
<td>0</td>
<td>100 to 635</td>
</tr>
<tr>
<td>Men</td>
<td>1206 (44.4)</td>
<td>305 (51.5)</td>
</tr>
<tr>
<td>Age, y</td>
<td>53.4±12.4</td>
<td>52.2±12.3</td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td>26.6 (24.0; 29.6)</td>
<td>25.9 (23.3; 28.7)</td>
</tr>
<tr>
<td>Smoking</td>
<td>Never smokers</td>
<td>953 (35.1)</td>
</tr>
<tr>
<td>Educational level</td>
<td>Low</td>
<td>1377 (50.7)</td>
</tr>
<tr>
<td>Physical activity</td>
<td>No exercise</td>
<td>481 (17.7)</td>
</tr>
<tr>
<td>Comorbidities</td>
<td>Diabetes mellitus</td>
<td>191 (7.0)</td>
</tr>
<tr>
<td>Hypertension</td>
<td>911 (33.5)</td>
<td>166 (28.0)</td>
</tr>
<tr>
<td>Hypercholesterolemia</td>
<td>917 (33.8)</td>
<td>212 (35.8)</td>
</tr>
<tr>
<td>Family history of CVD</td>
<td>990 (36.5)</td>
<td>212 (35.8)</td>
</tr>
<tr>
<td>Measurements at baseline</td>
<td>CDT, % of total transferrin</td>
<td>1.4 (1.3; 1.7)</td>
</tr>
<tr>
<td>HDL-C, mg/dL</td>
<td>48.7±11.5</td>
<td>48.5±12.3</td>
</tr>
<tr>
<td>eGFR, mL/min per 1.73 m²</td>
<td>91.5±16.7</td>
<td>94.4±16.3</td>
</tr>
</tbody>
</table>

Values represent numbers (percentages), means±SDs, or medians (25th–75th percentiles). Unit for EtG is ng/mL. BMI indicates body mass index; CDT, carbohydrate-deficient transferrin; CVD, cardiovascular disease; eGFR, estimated glomerular filtration rate; EtG, ethyl glucuronide; HDL-C, high density lipoprotein cholesterol; and PREVEND, Prevention of Renal and Vascular End-Stage Disease.
cancer-related mortality (Table S3). Finally, combining self-report and the biomarkers EtG and CDT into one measure and excluding the heavy drinkers did not significantly alter the associations between alcohol consumption and CVD and all-cause mortality (Table 4).

DISCUSSION
In this prospective cohort study, the association between alcohol consumption and CVD tended to be similar when EtG concentrations and self-report were used as measures of alcohol consumption. Although not statistically significant, the association between alcohol and CVD appeared to be nonlinear, with a lower risk for light-to-moderate drinkers compared with abstainers and heavy drinkers. We observed no associations between alcohol consumption measured by either EtG or self-report and all-cause mortality. Exclusion of participants who reported unstable alcohol consumption over time or had discrepant values for EtG/CDT and self-report did not alter our findings. Overall, our results support the reliability of self-reported consumption as a measure of habitual alcohol consumption.

Strengths and Limitations
To our knowledge, ours is the first long-term prospective cohort study to include EtG or any other direct markers of alcohol consumption in relation to CVD and all-cause mortality. Our findings suggest that self-reported consumption is a reliable measure of habitual alcohol consumption, and the use of EtG as a biomarker for alcohol consumption may provide additional insights into the health effects of alcohol.

Table 2. Associations of Self-Reported Alcohol Consumption With CVD Events and All-Cause Mortality in 5676 PREVEND Study Participants

<table>
<thead>
<tr>
<th>Variable</th>
<th>Alcohol Consumption Category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Abstention (N= 1366) 1 to 4/mo (N=960) 2 to 7/wk (N=1830) 1 to 3/d (N=1269) ≥4/d (N=251) P for Trend</td>
</tr>
<tr>
<td>CVD events, N (%)</td>
<td>115 (8) 52 (5) 105 (6) 90 (7) 24 (10)</td>
</tr>
<tr>
<td>Model 1</td>
<td>1.56 (1.12–2.17)* Reference 1.15 (0.82–1.61) 1.20 (0.85–1.69) 1.44 (0.88–2.34) 0.23</td>
</tr>
<tr>
<td>Model 2</td>
<td>1.43 (1.03–1.99)* Reference 1.12 (0.80–1.57) 1.22 (0.96–1.72) 1.28 (0.78–2.11) 0.42</td>
</tr>
<tr>
<td>Model 3</td>
<td>1.42 (1.02–1.98)* Reference 1.09 (0.78–1.52) 1.11 (0.79–1.58) 1.11 (0.68–1.84) 0.16</td>
</tr>
<tr>
<td>All-cause mortality, N (%)</td>
<td>204 (15) 118 (12) 198 (11) 165 (13) 39 (16)</td>
</tr>
<tr>
<td>Model 1</td>
<td>1.18 (0.94–1.48) Reference 1.14 (0.91–1.44) 1.12 (0.88–1.42) 1.27 (0.88–1.84) 0.88</td>
</tr>
<tr>
<td>Model 2</td>
<td>1.10 (0.87–1.38) Reference 1.08 (0.86–1.36) 1.04 (0.81–1.32) 1.02 (0.70–1.48) 0.66</td>
</tr>
<tr>
<td>Model 3</td>
<td>1.06 (0.84–1.34) Reference 1.06 (0.84–1.34) 1.01 (0.79–1.29) 0.97 (0.67–1.41) 0.67</td>
</tr>
</tbody>
</table>

Data are given as hazard ratios (95% CIs) for alcohol consumption categories vs the reference category with CVD events and all-cause mortality. Model 1, adjusted for age (years) and sex. Model 2, adjusted for model 1, smoking, education, physical activity, body mass index (categories), and parental history of CVD. Model 3, adjusted for model 2, hypertension, hypercholesterolemia, diabetes mellitus, and renal function (estimated glomerular filtration rate). Alcohol consumption categories are displayed in standard units per time period; 1 standard unit contains 10 g of alcohol. CVD indicates cardiovascular disease; and PREVEND, Prevention of Renal and Vascular End-Stage Disease.

*P<0.05.

Table 3. Associations of EtG Categories With CVD Events and All-Cause Mortality in 5676 PREVEND Study Participants

<table>
<thead>
<tr>
<th>Variable</th>
<th>Undetectable EtG (&lt;100 ng/mL)</th>
<th>Quintiles of Detectable EtG (≥100 ng/mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Quintile 1 (N=2716)</td>
<td>Quintile 2 (N=592)</td>
</tr>
<tr>
<td>CVD events, N (%)</td>
<td>205 (8)</td>
<td>37 (6)</td>
</tr>
<tr>
<td>Model 1</td>
<td>1.18 (0.83–1.68) Reference</td>
<td>0.58 (0.34–0.98)*</td>
</tr>
<tr>
<td>Model 2</td>
<td>1.14 (0.80–1.62) Reference</td>
<td>0.58 (0.34–0.99)*</td>
</tr>
<tr>
<td>Model 3</td>
<td>1.16 (0.82–1.65) Reference</td>
<td>0.60 (0.35–1.02)</td>
</tr>
<tr>
<td>All-cause mortality, N (%)</td>
<td>358 (13)</td>
<td>79 (13)</td>
</tr>
<tr>
<td>Model 1</td>
<td>0.91 (0.72–1.17) Reference</td>
<td>0.76 (0.54–1.07)</td>
</tr>
<tr>
<td>Model 2</td>
<td>0.91 (0.71–1.16) Reference</td>
<td>0.75 (0.53–1.05)</td>
</tr>
<tr>
<td>Model 3</td>
<td>0.89 (0.70–1.14) Reference</td>
<td>0.75 (0.53–1.06)</td>
</tr>
</tbody>
</table>

Data are given as hazard ratios (95% CIs) for EtG categories vs the reference category with CVD events and all-cause mortality. Model 1, adjusted for age (years) and sex. Model 2, adjusted for model 1, smoking, education, physical activity, body mass index (categories), and parental history of CVD. Model 3, adjusted for model 2, hypertension, hypercholesterolemia, diabetes mellitus, and renal function (estimated glomerular filtration rate). CVD indicates cardiovascular disease; EtG, ethyl glucuronide; and PREVEND, Prevention of Renal and Vascular End-Stage Disease.

*P<0.05.
biomarker of alcohol as a measure of habitual alcohol consumption in causative research. This enabled us to assess the impact of probable misclassification in self-reported consumption on its associations with CVD and mortality, robust to several sensitivity analyses.

Table 4. Sensitivity Analyses for the Associations of Self-Reported Alcohol Consumption With CVD Events and All-Cause Mortality

<table>
<thead>
<tr>
<th>Variable</th>
<th>Alcohol Consumption Category</th>
<th>P for Trend</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Abstention</td>
<td>1–4/mo</td>
</tr>
<tr>
<td>CVD events</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Main analysis model 3</td>
<td>1.42 (1.02–1.98)*</td>
<td>Reference</td>
</tr>
<tr>
<td>Exclusion of unstable consumption (n=5464)</td>
<td>1.47 (1.05–2.06)*</td>
<td>Reference</td>
</tr>
<tr>
<td>Time-varying consumption (n=5676)</td>
<td>1.02 (0.75–1.38)</td>
<td>Reference</td>
</tr>
<tr>
<td>Exclusion of misreported consumption (n=5068)</td>
<td>1.52 (1.08–2.14)*</td>
<td>Reference</td>
</tr>
<tr>
<td>Exclusion of misreported consumption+heavy drinkers (n=5009)</td>
<td>1.52 (1.08–2.14)*</td>
<td>Reference</td>
</tr>
<tr>
<td>All-cause mortality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Main analysis model 3</td>
<td>1.06 (0.84–1.34)</td>
<td>Reference</td>
</tr>
<tr>
<td>Exclusion of unstable consumption (n=5464)</td>
<td>1.05 (0.83–1.33)</td>
<td>Reference</td>
</tr>
<tr>
<td>Time-varying consumption (n=5676)</td>
<td>0.86 (0.69–1.07)</td>
<td>Reference</td>
</tr>
<tr>
<td>Exclusion of misreported consumption (n=5068)</td>
<td>1.10 (0.86–1.39)</td>
<td>Reference</td>
</tr>
<tr>
<td>Exclusion of misreported consumption+heavy drinkers (n=5009)</td>
<td>1.10 (0.86–1.39)</td>
<td>Reference</td>
</tr>
</tbody>
</table>

Data are given as hazard ratios (95% CIs) for alcohol consumption categories vs the reference category with CVD events and all-cause mortality. Models are adjusted for age (years), sex, smoking, education, physical activity, body mass index (categories), and parental history of cardiovascular disease, hypertension, hypercholesterolemia, diabetes mellitus, and renal function (estimated glomerular filtration rate). Alcohol consumption categories are displayed in standard units per time period; 1 standard unit contains 10 g of alcohol. CVD indicates cardiovascular disease.

*P<0.05.
One limitation of this study was that the event rate was relatively low. As a result, the precision of our estimates was insufficient to exclude plausibly sized effects on mortality. Furthermore, information on self-reported alcohol consumption was derived from a self-administered questionnaire that yielded limited information on the pattern of alcohol intake. Because the associations of alcohol consumption with CVD and mortality are both markedly affected by the pattern of drinking, we may have missed important associations of the quantity or frequency of alcohol intake with these outcomes.

Finally, urinary EtG as a marker also has its limitations: although it has a much longer detection window than most other direct biomarkers, EtG still is a short-term biomarker, covering only the 72 hours after consumption. Therefore, light drinkers in particular are easily misclassified and discriminating between abstainers and light drinkers can be problematic. Repeated sampling would decrease the misclassification of light drinkers, but is unlikely to be readily feasible in large-scale population research. EtG measured in hair might provide a better alternative, as this represents a more long-term measure of consumption, lasting several months. However, EtG in hair provides useful qualitative but not necessarily quantitative information.

**Previous Research**

To date, few studies have included indirect alcohol biomarkers in examining the associations between alcohol and disease. Jousilahti et al compared CDT, γ-glutamyl transferase (GGT), and self-report with coronary heart disease and found an inverse association for CDT, but a positive association for GGT with coronary heart disease. Self-reported consumption showed a borderline inverse association, which was attenuated after adjustment for confounders. The authors pointed out that self-reported levels of alcohol consumption in this study were low. Zatu et al studied CDT, GGT, and self-reported alcohol consumption with mortality. Only GGT was significantly positively associated with all-cause and cardiovascular mortality. Both studies emphasize that other factors than alcohol may influence these indirect markers. Indeed, other studies examined the association between GGT and coronary heart disease and confirmed that there is an independent mechanism linking serum GGT to coronary heart disease, which is also present in abstainers. By contrast, direct markers, such as EtG, are metabolites of the alcohol molecule and therefore specific for alcohol consumption.

In our study, we identified trends toward a J-shaped association with CVD, but could not definitively replicate previous observational studies that found a nonlinear relation between alcohol and CVD. Moreover, neither EtG nor self-reported consumption was associated with all-cause mortality, in contrast to previous studies that did find associations between alcohol and all-cause mortality and cause-specific mortality. This could have been because of the limited power of our study. In addition, the contribution of cardiovascular deaths to overall mortality was relatively small, and the association between alcohol and all-cause mortality is generally driven by cardiovascular mortality. Nevertheless, we observed similar results with EtG and self-report, as well as with self-report corrected for misclassification.

The consistency of our results across several measurement methods implies that findings from previous studies using self-report exclusively as a measure for alcohol consumption are unlikely to be heavily distorted by the subjectivity of self-report. At the same time, our results demonstrate the feasibility of incorporating urinary EtG in studies of populations in which self-report may be less reliable than the PREVEND study. Measurement of urinary EtG is inexpensive, easy, and noninvasive for the participant and therefore may be feasibly incorporated into even large-scale research.

In conclusion, self-reported alcohol consumption shows a similar association between alcohol consumption and CVD when compared with an objective measure of alcohol consumption. Moreover, these findings are consistent when the measures are combined to minimize misclassification. This argues for the validity of self-report; however, objective biomarkers can serve as effective supportive tools to complement self-report in the assessment of habitual alcohol consumption.


Table S1. Baseline characteristics of 5,676 PREVEND participants, by self-reported alcohol consumption category.

<table>
<thead>
<tr>
<th>Alcohol consumption category (by self-report)</th>
<th>No, almost never</th>
<th>1-4 units/month</th>
<th>2-7 units/week</th>
<th>1-3 units/day</th>
<th>≥ 4 units/day</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total N of participants</td>
<td>1366 (24.1%)</td>
<td>960 (16.9%)</td>
<td>1830 (32.2%)</td>
<td>1269 (22.4%)</td>
<td>251 (4.4%)</td>
</tr>
<tr>
<td>Sex:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>497 (36.4%)</td>
<td>420 (43.8%)</td>
<td>1011 (55.2%)</td>
<td>779 (61.4%)</td>
<td>199 (79.3%)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>54.8 ± 12.5</td>
<td>53.0 ± 12.6</td>
<td>50.9 ± 11.4</td>
<td>53.5 ± 10.8</td>
<td>54.2 ± 9.8</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>26.7 [24.1; 30.0]</td>
<td>26.3 [23.7; 29.3]</td>
<td>25.8 [23.4; 28.5]</td>
<td>25.6 [23.5; 28.4]</td>
<td>26.3 [23.5; 29.0]</td>
</tr>
<tr>
<td>Smoking:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never smokers</td>
<td>500 (36.6%)</td>
<td>355 (37.0%)</td>
<td>558 (30.5%)</td>
<td>234 (18.4%)</td>
<td>29 (11.6%)</td>
</tr>
<tr>
<td>Educational level:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>819 (60.0%)</td>
<td>421 (43.9%)</td>
<td>622 (34.0%)</td>
<td>428 (33.7%)</td>
<td>106 (42.2%)</td>
</tr>
<tr>
<td>Physical activity:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No exercise</td>
<td>287 (21.0%)</td>
<td>139 (14.5%)</td>
<td>226 (12.3%)</td>
<td>144 (11.3%)</td>
<td>62 (24.7%)</td>
</tr>
<tr>
<td>Comorbidities:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diabetes</td>
<td>123 (9.0%)</td>
<td>48 (5.0%)</td>
<td>76 (4.2%)</td>
<td>62 (4.9%)</td>
<td>13 (5.2%)</td>
</tr>
<tr>
<td>Hypertension</td>
<td>505 (37.0%)</td>
<td>291 (30.3%)</td>
<td>467 (25.5%)</td>
<td>394 (31.0%)</td>
<td>96 (38.2%)</td>
</tr>
<tr>
<td>Hypercholesterolemia</td>
<td>484 (35.4%)</td>
<td>327 (34.1%)</td>
<td>578 (31.6%)</td>
<td>481 (37.9%)</td>
<td>127 (50.6%)</td>
</tr>
<tr>
<td>Family history CVD</td>
<td>503 (36.8%)</td>
<td>370 (38.5%)</td>
<td>594 (32.5%)</td>
<td>428 (33.7%)</td>
<td>90 (35.9%)</td>
</tr>
<tr>
<td>Measurements at baseline</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDT (% of total transferrin)</td>
<td>1.4 [1.2; 1.6]</td>
<td>1.4 [1.2; 1.6]</td>
<td>1.5 [1.3; 1.8]</td>
<td>1.6 [1.4; 1.9]</td>
<td>1.9 [1.6; 2.4]</td>
</tr>
<tr>
<td>HDL-C (mg/dL)</td>
<td>46.5 ± 11.5</td>
<td>47.3 ± 11.4</td>
<td>48.4 ± 12.1</td>
<td>50.5 ± 12.9</td>
<td>50.6 ± 14.3</td>
</tr>
<tr>
<td>eGFR (mL/min/1.73m²)</td>
<td>90.1 ± 17.5</td>
<td>92.3 ± 16.1</td>
<td>95.2 ± 15.3</td>
<td>94.3 ± 15.2</td>
<td>94.4 ± 14.9</td>
</tr>
</tbody>
</table>

Values represent numbers (percentages); means ± standard deviations; medians [interquartile ranges]
EtG = ethyl glucuronide, BMI = body mass index, CVD = cardiovascular disease, CDT = carbohydrate-deficient transferrin, HDL-C = high density cholesterol, eGFR = estimated glomerular filtration rate and PREVEND = Prevention of Renal and Vascular End-Stage Disease.

Alcohol consumption categories are displayed in standard units per time period, one standard unit contains 10 grams of alcohol.
Table S2. Sensitivity analyses for the associations of EtG with CVD events and all-cause mortality.

<table>
<thead>
<tr>
<th>EtG categories</th>
<th>Undetectable EtG (&lt; 100 ng/mL)</th>
<th>Quintiles of detectable EtG (≥ 100 ng/mL)</th>
<th>P for trend</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Q1</td>
<td>Q2</td>
<td>Q3</td>
</tr>
<tr>
<td>CVD events</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 3</td>
<td>1.16 (0.82 – 1.65)</td>
<td>Ref</td>
<td>0.60 (0.35 – 1.02)</td>
</tr>
<tr>
<td>Exclusion of unstable consumption (n = 5464)</td>
<td>1.19 (0.83 – 1.71)</td>
<td>Ref</td>
<td>0.65 (0.38 – 1.11)</td>
</tr>
<tr>
<td>All-cause mortality</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 3</td>
<td>0.89 (0.70 – 1.14)</td>
<td>Ref</td>
<td>0.75 (0.53 – 1.06)</td>
</tr>
<tr>
<td>Exclusion of unstable consumption (n = 5464)</td>
<td>0.92 (0.72 – 1.19)</td>
<td>Ref</td>
<td>0.73 (0.51 – 1.04)</td>
</tr>
</tbody>
</table>

HRs and 95% confidence intervals for ethyl glucuronide versus the reference category with CVD events and all-cause mortality.

Models are adjusted for age (years), sex, smoking, education, physical activity, BMI (categories) and parental history of CVD, hypertension, hypercholesterolemia, diabetes and renal function (eGFR).

EtG = ethyl glucuronide, BMI = body mass index, CVD = cardiovascular disease, eGFR = estimated glomerular filtration rate, HR = hazard ratio, PREVEND = Prevention of Renal and Vascular End-Stage Disease.
Table S3. Associations of time varying self-reported alcohol consumption with all-cause and cause-specific mortality in 5,676 PREVEND participants.

<table>
<thead>
<tr>
<th>Alcohol consumption category</th>
<th>Abstention</th>
<th>1-4/month</th>
<th>2-7/week</th>
<th>1-3/day</th>
<th>≥ 4/day</th>
<th>P for trend</th>
</tr>
</thead>
<tbody>
<tr>
<td>All-cause mortality</td>
<td>N = 1366</td>
<td>N = 960</td>
<td>N = 1830</td>
<td>N = 1269</td>
<td>N = 251</td>
<td></td>
</tr>
<tr>
<td>Model 3</td>
<td>0.86 (0.69 – 1.07)</td>
<td>Ref</td>
<td>0.86 (0.68 – 1.07)</td>
<td>0.77 (0.61 – 0.98)*</td>
<td>0.80 (0.54 – 1.20)</td>
<td>0.06</td>
</tr>
<tr>
<td>CVD mortality</td>
<td>N = 43 (3%)</td>
<td>N = 28 (3%)</td>
<td>N = 48 (3%)</td>
<td>N = 25 (3%)</td>
<td>N = 5 (2%)</td>
<td></td>
</tr>
<tr>
<td>Model 3</td>
<td>0.71 (0.46 – 1.11)</td>
<td>Ref</td>
<td>0.66 (0.41 – 1.07)</td>
<td>0.63 (0.38 – 1.04)</td>
<td>0.54 (0.20 – 1.41)</td>
<td>0.06</td>
</tr>
<tr>
<td>Cancer mortality</td>
<td>N = 89 (7%)</td>
<td>N = 54 (6%)</td>
<td>N = 98 (5%)</td>
<td>N = 91 (7%)</td>
<td>N = 22 (9%)</td>
<td></td>
</tr>
<tr>
<td>Model 3</td>
<td>0.98 (0.70 – 1.39)</td>
<td>Ref</td>
<td>1.24 (0.89 – 1.74)</td>
<td>1.00 (0.70 – 1.43)</td>
<td>1.02 (0.58 – 1.80)</td>
<td>0.75</td>
</tr>
<tr>
<td>Mortality from other causes</td>
<td>N = 72 (5%)</td>
<td>N = 36 (4%)</td>
<td>N = 52 (3%)</td>
<td>N = 42 (3%)</td>
<td>N = 12 (5%)</td>
<td></td>
</tr>
<tr>
<td>Model 3</td>
<td>0.82 (0.57 – 1.20)</td>
<td>Ref</td>
<td>0.56 (0.36 – 0.85)*</td>
<td>0.62 (0.40 – 0.96)*</td>
<td>0.77 (0.38 – 1.57)</td>
<td>0.03</td>
</tr>
</tbody>
</table>

HRs and 95% confidence intervals for ethyl glucuronide versus the reference category with CVD events and all-cause mortality.

Models are adjusted for age (years), sex, smoking, education, physical activity, BMI (categories) and parental history of CVD, hypertension, hypercholesterolemia, diabetes and renal function (eGFR).

EtG = ethyl glucuronide, BMI = body mass index, CVD = cardiovascular disease, eGFR = estimated glomerular filtration rate, HR = hazard ratio, PREVEND = Prevention of Renal and Vascular End-Stage Disease.

Alcohol consumption categories are displayed in standard units per time period, one standard unit contains 10 grams of alcohol.

*P < 0.05.