Mendelian randomization, the use of common polymorphisms as surrogates for measuring exposure levels in epidemiologic studies, provides one method of assessing the causal nature of some environmental exposures. This can be illustrated by looking at the association between the ALDH2 polymorphism and esophageal cancer. Alcohol drinking is considered a risk factor for esophageal cancer, and exposure to high levels of acetaldehyde, the principal metabolite of alcohol, may be responsible for the increased cancer risk. The ability to metabolize acetaldehyde is encoded by the ALDH2 gene, which is polymorphic in some populations. The ALDH2*2 allele produces an inactive protein subunit, which is unable to metabolize acetaldehyde. An individual’s genotype at this locus may influence their esophageal cancer risk through two mechanisms, first through influencing alcohol intake and second through influencing acetaldehyde levels. We have carried out a meta-analysis of studies looking at the ALDH2 genotype and esophageal cancer and found that risk was reduced among *2/*2 homozygotes [odds ratio (OR), 0.36; 95% confidence interval (95% CI), 0.16-0.80] and increased among heterozygotes (OR, 3.19; 95% CI, 1.86-5.47) relative to *1/*1 homozygotes. This provides strong evidence that alcohol intake increases the risk of esophageal cancer and individuals whose genotype results in markedly lower intake, because they have an adverse reaction to alcohol are thus protected. This meta-analysis also provides evidence that acetaldehyde plays a carcinogenic role in esophageal cancer. The two different processes operating as a result of the ALDH2 genotype have implications for the interpretation of studies using the Mendelian randomization paradigm.

(Author: Sarah J. Lewis and George Davey Smith)

Received 3/18/05; revised 5/10/05; accepted 5/25/05.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1754 solely to indicate this fact.

Requests for reprints: Sarah J. Lewis, Department of Social Medicine, University of Bristol, Bristol, United Kingdom. Phone: 440-117-928-7266; Fax: 440-117-928-7225; E-mail: s.j.lewis@bristol.ac.uk

Copyright © 2005 American Association for Cancer Research.

doi:10.1158/1055-9565.EPI05-0196

Alcohol, ALDH2, and Esophageal Cancer: A Meta-analysis Which Illustrates the Potentials and Limitations of a Mendelian Randomization Approach

Sarah J. Lewis and George Davey Smith

Department of Social Medicine, University of Bristol, Bristol, United Kingdom

Abstract

Mendelian randomization, the use of common polymorphisms as surrogates for measuring exposure levels in epidemiologic studies, provides one method of assessing the causal nature of some environmental exposures. This can be illustrated by looking at the association between the ALDH2 polymorphism and esophageal cancer. Alcohol drinking is considered a risk factor for esophageal cancer, and exposure to high levels of acetaldehyde, the principal metabolite of alcohol, may be responsible for the increased cancer risk. The ability to metabolize acetaldehyde is encoded by the ALDH2 gene, which is polymorphic in some populations. The ALDH2*2 allele produces an inactive protein subunit, which is unable to metabolize acetaldehyde. An individual’s genotype at this locus may influence their esophageal cancer risk through two mechanisms, first through influencing alcohol intake and second through influencing acetaldehyde levels. We have carried out a meta-analysis of studies looking at the ALDH2 genotype and esophageal cancer and found that risk was reduced among *2/*2 homozygotes [odds ratio (OR), 0.36; 95% confidence interval (95% CI), 0.16-0.80] and increased among heterozygotes (OR, 3.19; 95% CI, 1.86-5.47) relative to *1/*1 homozygotes. This provides strong evidence that alcohol intake increases the risk of esophageal cancer and individuals whose genotype results in markedly lower intake, because they have an adverse reaction to alcohol are thus protected. This meta-analysis also provides evidence that acetaldehyde plays a carcinogenic role in esophageal cancer. The two different processes operating as a result of the ALDH2 genotype have implications for the interpretation of studies using the Mendelian randomization paradigm.
acetaldehyde, leading to an accumulation of acetaldehyde after alcohol intake (11). Individuals who are homozygous for the *ALDH2*2 allele have 18 times higher and heterozygotes have five times higher peak blood alcohol levels compared with *1*1 homozygotes (12). *ALDH2*2 homozygotes are characterized by a facial flushing response after consumption of alcohol coupled with nausea, drowsiness, headache, and other unpleasant symptoms that prevent them from heavy drinking (12). Heterozygotes have a limited ability to metabolize acetaldehyde but exhibit a less severe reaction than that seen among *ALDH2*2 homozygotes. We have carried out meta-analysis of studies looking at the *ALDH2* genotype and esophageal cancer to examine how the Mendelian randomization approach performs in this situation. A priori, we would expect *ALDH2*2 homozygotes to have a reduced esophageal cancer risk due to a lower alcohol intake among these individuals (Fig. 1).

### Materials and Methods

#### Search Strategy and Selection Criteria.

Papers published before the end of March 2004 were identified through a search of Medline (http://www.ncbi.nlm.nih.gov) and ISIS web of knowledge (http://wok.mimas.ac.uk), using the following search terms: “oesophageal” or “esophageal” and “ALDH2” or “aldehyde dehydrogenase.” A cited reference search of retrieved articles was carried out and publications were also identified by review of the bibliographies of retrieved articles. Articles reporting on *ALDH2* genotype in cases of esophageal cancer and controls were identified (13-19). Where authors had published more than one article using the same case series, we selected the most recent publication. Our a priori hypothesis was that *ALDH2*2/2 genotype protects against esophageal cancer because individuals with this genotype are unlikely to consume high quantities of alcohol. For this reason, we excluded two studies from the meta-analysis of *2*2 genotype versus *1*1 genotype, which specifically selected cases and/or controls on the basis of alcohol intake (14, 16). We also wanted to test the hypothesis that *ALDH2*1/2 genotype increases esophageal cancer risk relative to *1*1 genotype given a similar level of alcohol intake. We therefore carried out a stratified meta-analysis in which subjects were stratified according to alcohol intake; we excluded two studies that provided no information on alcohol intake by genotype (13, 14).

#### Statistical Analysis.

Alcohol intake was coded as 0 = nondrinkers, 1 = all individuals not classified as nondrinkers or heavy drinkers, and 2 = heavy drinkers. Unadjusted odds ratios (OR) were based on published genotype frequencies. Random effects models were used because the method of case/control ascertainment and the alcohol intake cutoffs both differed between the studies. We quantified the extent of heterogeneity using I², the percentage of total variation across studies that is attributable to heterogeneity rather than chance (20). The influence of alcohol intake on the relationship between *ALDH2* genotype and esophageal cancer risk was assessed using meta-regression analysis (21). A Monte Carlo permutation procedure was used to determine deviation from Hardy-Weinberg equilibrium among control populations using the HWSIM program provided on the following web site: http://krunch.med.yale.edu/hwsim. All other statistical analysis was carried out in Stata version 8 (Stata Corp., College Station, TX).

### Results

Seven studies, (13-19) with a total of 905 cases of esophageal cancer were identified and these are summarized in Table 1. The studies were carried out in Japan, Taiwan, and Thailand. Relative genotype frequencies in all control groups did not deviate from values predicted by Hardy-Weinberg equilibrium. *ALDH2*2 allele frequencies among control populations were lowest in a study in which the controls were alcoholics (16) and were also lower in the Thai study than in the Japanese studies (17). Our meta-analysis gave an overall OR of 0.36 [95% confidence interval (95% CI), 0.16-0.80] for the risk of esophageal cancer among *2*2 homozygotes compared with *1*1 homozygotes (Fig. 2) and an overall OR of 3.19 (95% CI, 1.86-5.47) for heterozygotes compared with *1*1 homozygotes (Fig. 3). Among nondrinkers, there was no strong evidence for an increase in risk among heterozygotes (OR, 1.31; 95% CI, 0.70-2.47) relative to *1*1 individuals. However, among heavy drinkers there was a 7-fold increase in risk (OR, 7.07; 95% CI, 3.67-13.6). Among all others with an intermediate alcohol intake the risk among heterozygotes versus *1*1 homozygotes was 2.49 (95% CI, 1.29-4.79). Meta-regression analysis showed evidence that alcohol intake influenced the effect of the *1*2 genotype on esophageal cancer risk (P = 0.008) and that the larger the amount of alcohol consumption (i.e., the greater the OR of *1*2 versus *1*1 genotypes).

There was no evidence of between study heterogeneity in the analysis of *2*2 versus *1*1 genotype (χ² = 2.14, P = 0.71, I² = 0.0%), but there was evidence of heterogeneity in the analysis of *1*2 versus *1*1 genotype (χ² = 53.5, P < 0.001).

**Figure 1.** Schematic diagram of the postulated relationships among *ALDH2*, alcohol consumption, acetaldehyde levels, and esophageal cancer.
This heterogeneity was slightly reduced within strata after stratification for alcohol intake and was confined to alcohol drinkers, both heavy drinkers ($\chi^2 = 9.01$, $P = 0.03$, $I^2 = 66.7\%$) and others ($\chi^2 = 11.6$, $P = 0.009$, $I^2 = 74.0\%$), and probably reflects differences in crude estimates of alcohol intake between studies, particularly as there was no between study heterogeneity among nondrinkers ($\chi^2 = 0.35$, $P = 0.84$, $I^2 = 0.0\%$). The overall test of heterogeneity among the three effect estimates for the pooled strata above gave the following results: $\chi^2 = 32.6$ and $P < 0.001$.

In the studies that provided alcohol use by genotype in the control group, heavy drinking was more common among individuals with the *1*1 genotype compared with the *1*2 genotype and there were no heavy drinkers among individuals with the *2*2 genotype (Table 2).

An Egger test (22) provided no evidence that effect estimates were related to study size ($P = 0.61$ for *2*2 analysis and $P = 0.11$ for *1*2 analysis), providing some reassurance that small study bias, such as publication bias has not distorted the findings.

### Discussion

This meta-analysis shows strong evidence that the ALDH2*2*2 genotype reduces esophageal cancer risk by ~3-fold, and this is likely to be due to markedly lower levels of alcohol consumption in *2*2 versus *1*1 homozygotes. ALDH2*2*2 homozygotes are intolerant to alcohol and can exhibit a severe reaction following intake of a small amount of alcohol; hence, this genotype protects against esophageal cancer because it influences propensity to drink alcohol (Table 2). This provides support to the existing evidence from epidemiologic studies that the association between alcohol intake and esophageal cancer is causal in nature. As ALDH2 genotype is determined at birth, the findings with respect to genotype are not subject to reverse causation - ill-health influencing the exposure measure - unlike the findings with respect to directly measured alcohol intake. Furthermore, it is often difficult to separate the effects of heavy alcohol intake and smoking in observational studies, because the two are highly correlated. In a large Japanese cohort study, smoking was shown to be strongly associated with heavy alcohol intake (ref. 23; Table 3), however, smoking was not associated with ALDH2 genotype (refs. 15, 24; Table 3). Hence measuring ALDH2 genotype as a surrogate for alcohol intake is not subject to confounding, and can be used to verify the findings of epidemiologic studies. Finally, a tendency to inaccurately report alcohol intake, most probably underreporting, may be related to other risk factors for disease, leading to bias in the association between reported alcohol intake and disease. This inaccurate reporting will not be an issue if genotype is used as a proxy for alcohol intake.

A meta-analysis of observational studies of alcohol intake and esophageal cancer risk found relative risks of 1.8, 2.38, and 4.36 for risk among light, moderate, and heavy drinkers, respectively, compared with nondrinkers (25). In the control population in the study by Yokoyama et al. (19), 9.4% of *1*1 individuals were nondrinkers (never plus ex drinkers) and 28.2%, 39.6%, and 22.9% were light, moderate, and heavy drinkers, respectively. Whilst the cutoff between light and moderate drinkers was different in the two studies (29 g/d in the study by Yokoyama et al. and 39.99 g/d in the meta-analysis of observational studies); other cutoff points were similar. We therefore used the relative risks associated with different levels of alcohol consumption in the meta-analysis of

### Table 1. Studies of the ALDH2 polymorphism and esophageal cancer risk

<table>
<thead>
<tr>
<th>Reference</th>
<th>Country</th>
<th>Population</th>
<th>No. cases</th>
<th>No. controls</th>
<th>*2 Allele frequency in controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hori et al. (13)</td>
<td>Japan</td>
<td>93</td>
<td>70 healthy individuals</td>
<td>0.33</td>
<td></td>
</tr>
<tr>
<td>Chao et al. (14)</td>
<td>Taiwan</td>
<td>88 (59 alcoholics)</td>
<td>105 nonalcoholics</td>
<td>0.29</td>
<td></td>
</tr>
<tr>
<td>Matsuo et al. (15)</td>
<td>Japan</td>
<td>102</td>
<td>241 hospital outpatients</td>
<td>0.28</td>
<td></td>
</tr>
<tr>
<td>Yokoyama et al. (16)</td>
<td>Japan</td>
<td>112 alcoholics</td>
<td>526 alcoholics</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>Boonphpiphat et al. (17)</td>
<td>Thailand</td>
<td>210</td>
<td>261 hospital inpatients</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>Itoya et al. (18)</td>
<td>Japan</td>
<td>74</td>
<td>241 healthy individuals</td>
<td>0.22</td>
<td></td>
</tr>
<tr>
<td>Yokoyama et al. (19)</td>
<td>Japan</td>
<td>234</td>
<td>634 population</td>
<td>0.27</td>
<td></td>
</tr>
</tbody>
</table>

*Figures for genotype frequencies by case-control status obtained by personal correspondence.

Figure 2. Risk of esophageal cancer in individuals with the ALDH2*2*2 versus ALDH2*1*1 genotype.
observational studies (25) to calculate an approximate overall probability of disease given alcohol intake as reported by individuals in the study by Yokoyama et al. (19). To calculate this probability, we used the following equation: RR = \frac{\text{RR}_i}{P_i}, where i denotes the drinking category (non, light, moderate, and heavy), RR is the relative risk in the ith drinking category estimated by a meta-analysis (25), and P is the assumed proportion of ith drinking category among controls (19). We estimated an overall relative risk for this group of around 2.54 compared with nondrinkers (equivalent to individuals as virtually all of these individuals are nondrinkers). The OR of esophageal cancer for homozygotes versus heterozygotes in this meta-analysis was 2.77 (95% CI, 1.25-6.12). Risk ratios and ORs are virtually equivalent for esophageal cancer as the disease is very rare. Hence the greater risk seen among homozygotes is what would be expected given their drinking behavior and findings from observational epidemiologic studies relating drinking behavior to esophageal cancer risk.

There are markedly increased acetaldehyde levels among heterozygotes who drink alcohol, and heterozygosity is associated with an 3-fold overall increase in risk for esophageal cancer (Fig. 3). This suggests that acetaldehyde may be the mechanism through which drinking alcohol increases the risk of esophageal cancer. This meta-analysis and the results of the meta-regression show that the

Table 2. Drinking status by ALDH2 genotype among controls

<table>
<thead>
<tr>
<th>Reference</th>
<th>Drinking categories</th>
<th><em>I</em>1, n (%)</th>
<th><em>I</em>2, n (%)</th>
<th><em>2</em>2, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matsuo et al. (15)</td>
<td>Others</td>
<td>104 (82.5)</td>
<td>92 (95.8)</td>
<td>19 (100)</td>
</tr>
<tr>
<td></td>
<td>Heavy drinkers</td>
<td>22 (17.5)</td>
<td>4 (4.2)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Yokoyama et al. (16)</td>
<td>Alcoholics</td>
<td>476 (100)</td>
<td>50 (100)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Boonyaphiphit et al. (17)</td>
<td>Nondrinkers</td>
<td>104 (48.4)</td>
<td>24 (60.0)</td>
<td>Excluded due to small numbers</td>
</tr>
<tr>
<td></td>
<td>≤60g/d</td>
<td>66 (30.7)</td>
<td>11 (27.5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;60g/d</td>
<td>45 (20.9)</td>
<td>5 (12.5)</td>
<td></td>
</tr>
<tr>
<td>Itoya et al. (18)</td>
<td>Nondrinkers</td>
<td>135 (94.4)</td>
<td>65 (73.9)</td>
<td>0 (0)</td>
</tr>
<tr>
<td></td>
<td>Never/rare</td>
<td>21 (6.2)</td>
<td>80 (26.1)</td>
<td>41 (59.1)</td>
</tr>
<tr>
<td></td>
<td>Light</td>
<td>96 (28.2)</td>
<td>103 (41.2)</td>
<td>2 (4.7)</td>
</tr>
<tr>
<td></td>
<td>Moderate</td>
<td>135 (39.6)</td>
<td>35 (14.0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td></td>
<td>Heavy</td>
<td>78 (22.9)</td>
<td>27 (10.8)</td>
<td>0 (0)</td>
</tr>
<tr>
<td></td>
<td>Ex drinker</td>
<td>11 (3.2)</td>
<td>5 (2.0)</td>
<td>0 (0)</td>
</tr>
</tbody>
</table>

Table 3. Smoking status by heavy alcohol intake and by ALDH2 genotype

<table>
<thead>
<tr>
<th>Smoking status</th>
<th>Alcohol intake (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moderate (1-149 g/wk)</td>
<td>Heavy (≥450 g/wk)</td>
</tr>
<tr>
<td>Never</td>
<td>47.1</td>
</tr>
<tr>
<td>Former</td>
<td>22.7</td>
</tr>
<tr>
<td>Current</td>
<td>15.5</td>
</tr>
<tr>
<td>ALDH2 genotype</td>
<td><em>2</em>2</td>
</tr>
<tr>
<td>Takagi et al. (23), (men, n = 1919)</td>
<td>35.9</td>
</tr>
<tr>
<td>Current smokers</td>
<td>35.9</td>
</tr>
<tr>
<td>Matsuo et al. (15), (controls, n = 241)</td>
<td><em>2</em>2 and <em>1</em>2</td>
</tr>
<tr>
<td>Never</td>
<td>54.0</td>
</tr>
<tr>
<td>Former</td>
<td>22.0</td>
</tr>
<tr>
<td>Current</td>
<td>23.8</td>
</tr>
</tbody>
</table>

Figure 3. Risk of esophageal cancer in individuals with the ALDH2*1*2 versus ALDH2*1*1 genotype.
association between ALDH2 heterozygote genotype and esophageal cancer is dependent on alcohol consumption. Among nondrinkers, there is no strong evidence of an increased risk. Whereas among heavy drinkers, a substantial elevated risk is seen. This suggests that possession of an ALDH2*2 allele does not increase risk of esophageal cancer unless alcohol is consumed.

Limitations. In this meta-analysis, we did not have access to individual level data and were not able to reclassify individuals by alcohol intake; instead, we were forced to use the cutoffs used by the different studies as approximate measures of nondrinking, heavy drinking, and other. This is likely to have been a source of heterogeneity in the meta-analysis. Similarly, we did not have access to individual data on smoking and thus were unable to assess whether alcohol intake, as determined by ALDH2 genotype, influences the number of cigarettes smoked. However, we suggest that cigarette smoking is unlikely to be a confounder given the apparent independence of ALDH2 genotype and smoking status (Table 3).

In this example, it is easy to observe the two different processes operating as a result of the ALDH2 genotype because this polymorphism in ALDH2 is well characterized with reliable differences in phenotype by genotype. However, whereas this polymorphism can be viewed as having two related effects with opposing influences on disease risk, there are very few examples of well-characterized polymorphisms where this can be understood. Finally, canalization, the developmental buffering against the effect of a polymorphism during fetal development, is also often a potential problem in studies that apply the Mendelian randomization concept (1). It is unlikely to be a factor in this example, however, as alcohol consumption is only adopted in adolescence or adulthood.

In summary, this polymorphism influences the propensity for exposure to alcohol and modifies exposure to acetaldehyde among alcohol drinkers and can therefore be used to characterize both alcohol intake and acetaldehyde as components of a causal chain increasing the risk of esophageal cancer. In using ALDH2*2 as a surrogate for measuring alcohol intake, this study shows that alcohol drinking is related to elevated risk of esophageal cancer thus illustrating the potential of the Mendelian randomization concept. However, this study also shows that ALDH2*1*2 is related to both lower alcohol consumption and elevated risk of esophageal cancer when the amount of alcohol consumed is identical; thus, analyzing the relationship between ALDH2*1*2 and risk of esophageal cancer without considering the amount of alcohol consumption is misleading. The result of ALDH2*1*2 illustrates a potential limitation of a Mendelian randomization approach in that in less well characterized situations, similar gene-environment interactions may be occurring which are not recognized, leading to spurious conclusions being drawn from looking at the main effects.

Acknowledgments
We thank Roger Harbord for his statistical advice and for commenting on an earlier version of this article.

References
Alcohol, ALDH2, and Esophageal Cancer: A Meta-analysis Which Illustrates the Potentials and Limitations of a Mendelian Randomization Approach

Sarah J. Lewis and George Davey Smith


Updated version
Access the most recent version of this article at:
http://cebp.aacrjournals.org/content/14/8/1967

Cited articles
This article cites 22 articles, 11 of which you can access for free at:
http://cebp.aacrjournals.org/content/14/8/1967.full.html#ref-list-1

Citing articles
This article has been cited by 34 HighWire-hosted articles. Access the articles at:
/content/14/8/1967.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.