Perspectives on Psychological Science

http://pps.sagepub.com/

Mendelian Randomization for Strengthening Causal Inference in Observational Studies : Application to Gene × Environment Interactions

George Davey Smith

Perspectives on Psychological Science 2010 5: 527
DOI: 10.1177/1745691610383505

The online version of this article can be found at:
http://pps.sagepub.com/content/5/5/527

Published by:

http://www.sagepublications.com

On behalf of:

Association For Psychological Science

Additional services and information for *Perspectives on Psychological Science* can be found at:

Email Alerts: http://pps.sagepub.com/cgi/alerts

Subscriptions: http://pps.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Downloaded from pps.sagepub.com by guest on October 27, 2010

aps ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Perspectives on Psychological Science
5(5) 527–545
© The Author(s) 2010
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1745691610383505
http://pps.sagepub.com
SAGE

Mendelian Randomization for Strengthening Causal Inference in Observational Studies: Application to Gene × Environment Interactions

George Davey Smith
Department of Social Medicine, University of Bristol, Bristol, United Kingdom

Abstract

Identification of environmentally modifiable factors causally influencing disease risk is fundamental to public-health improvement strategies. Unfortunately, observational epidemiological studies are limited in their ability to reliably identify such causal associations, reflected in the many cases in which conventional epidemiological studies have apparently identified such associations that randomized controlled trials have failed to verify. The use of genetic variants as proxy measures of exposure—an application of the Mendelian randomization principle—can contribute to strengthening causal inference. Genetic variants are not subject to bias due to reverse causation (disease processes influencing exposure, rather than vice versa) or recall bias, and if simple precautions are applied, they are not influenced by confounding or attenuation by errors. The principles of Mendelian randomization are illustrated with specific reference to studies of the effects of alcohol intake on various health-related outcomes through the utilization of genetic variants related to alcohol metabolism (in *ALDH2* and *ADH1B*). Ways of incorporating Gene × Environment interactions into the Mendelian randomization framework are developed, and the strengths and limitations of the approach discussed.

Keywords

Gene × Envrionment interactions, Mendelian randomization, causal inference, epidemiology, confounding, alcohol

In 1875, George Darwin, the second son and fifth child of Charles Darwin, reviewed evidence on the putative detrimental effects of cousin marriages on offspring health, something of personal interest to him as he was the product of such a union (G.H. Darwin, 1875). He concluded by reviewing the most comprehensive studies of the issue and described what may be the first presentation of Gene × Environment interaction informed by at least some understanding of heredity. In 1864, George Darwin stated that, "Dr. Mitchell had come to the conclusion that under favorable conditions of life, the apparent ill effects were frequently almost nil, whilst if the children were ill-fed, badly housed and clothed, the evil might become very marked. This is in striking accordance with some unpublished experiments of my father, Mr. Charles Darwin, on the in-and-in breeding of plants; for he has found that in-bred plants; when allowed enough space and good soil, frequently show little or no deterioration, whilst when placed in competition with another plant, they frequently perish or are much stunted." The unpublished findings of Charles Darwin were later published in his 1876 book *The Effects of Cross and Self Fertilization in the Vegetable Kingdom* (C. Darwin, 1876).

The effects of cousin marriage, which would now be considered to reflect disorders generated by homozygosity for uncommon variants, were apparently mainly seen in suboptimal environmental circumstances. There are clearly echoes here of celebrated contemporary Gene × Environment interactions, such as that between genetic variation in the serotonin transporter gene (*5-HTTLPR*), stressful life events, and the risk of depression (Caspi et al., 2003). Unlike recent examples of Gene × Environment interaction in the molecular genetic age (Caspi et al., 2003, 2007), which have failed to stand up to rigorous attempts at replication (Risch et al., 2009; Steer, Davey Smith, Emmett, Hibbeln, & Golding, in press), the interesting claims made by Dr. Mitchell have not been formally followed up in this way. In this review, I suggest that

Corresponding Author:
George Davey Smith, MRC Centre for Causal Analyses in Translational Epidemiology, Department of Social Medicine, University of Bristol, Oakfield House, Oakfield Grove, Bristol BS8 2BN
E-mail: KZ.davey-smith@bristol.ac.uk

Downloaded from pps.sagepub.com by guest on October 27, 2010

Gene × Environment interactions can provide useful evidence as to the causal effect of the environmental exposure on disease, and that, in some circumstances, this could have more utility for strategies to improve population health than would focusing on other aspects of the interactions themselves. To illustrate this, I utilize examples from the alcohol and health area, one of the many contested fields in which disparate claims based on observational data have been made. I will briefly outline the use of genetic main effects in the basic Mendelian randomization approach for strengthening causal inference. I discuss how Gene × Environment interactions can also be utilized in this regard and discuss the typology of Gene × Environment interaction in some of the framework I advance. Finally, I conclude by briefly outlining the limitations of this approach.

Mendelian Randomization: What Is It and How Does It Work?

The basic reasoning utilized in the Mendelian randomization approach is that if genetic variants either alter the level of or mirror the biological effects of a modifiable environmental exposure that itself alters disease risk, then these genetic variants should be related to disease risk to the extent predicted by their influence on exposure to the risk factor. Common genetic polymorphisms that have a well-characterized biological function (or are markers for such variants) can therefore be utilized to study the effect of a suspected environmental exposure on disease risk (Davey Smith, 2006a; Davey Smith & Ebrahim, 2003; Davey Smith & Ebrahim, 2004; Davey Smith & Ebrahim, 2005; Davey Smith, Timpson, & Ebrahim, 2008; Ebrahim & Davey Smith, 2008; Lawlor et al., 2008). The variants should not have an association with the disease outcome except through their link to the modifiable risk process of interest.

It may seem counterintuitive to study genetic variants as proxies for environmental exposures rather than measure the exposures themselves. However, there are several crucial advantages of utilizing functional genetic variants (or their markers) in this manner that relate to the problems with observational studies outlined above. First, unlike environmental exposures, genetic variants are not generally associated with the wide range of behavioral, social, and physiological factors that can confound associations. This means that if a genetic variant is used as a proxy for an environmentally modifiable exposure, it is unlikely to be confounded in the way that direct measures of the exposure will be. Further, aside from the effects of population structure, (Palmer & Cardon, 2005), such variants will not be associated with other genetic variants, except through linkage disequilibrium (the association of alleles located close together on a chromosome).

Second, inferences drawn from observational studies may be subject to bias due to reverse causation. Disease processes may influence exposure levels, such as alcohol intake, or measures of intermediate phenotypes, such as cholesterol levels and C-reactive protein. However, germline genetic variants associated with average alcohol intake or circulating levels of intermediate phenotypes will not be influenced by the onset of disease. This will be equally true with respect to reporting bias generated by knowledge of disease status in case-control studies and to differential reporting bias in any study design.

Finally, a genetic variant will indicate long-term levels of exposure, and, if the variant is considered to be a proxy for such exposure, it will not suffer from the measurement error inherent in phenotypes that have high levels of variability. For example, differences between groups defined by cholesterol level-related genotype will, over a long period, reflect the cumulative differences in absolute cholesterol levels between the groups. For individuals, blood cholesterol is variable over time, and the use of single measures of cholesterol will underestimate the true strength of association between cholesterol and, for instance, coronary heart disease (CHD). Indeed, use of the Mendelian randomization approach predicts a strength of association that is in line with randomized controlled trial findings of effects of cholesterol lowering, in which the increasing benefits seen over the relatively short trial period are projected to the expectation for differences over a lifetime (Davey Smith & Ebrahim, 2004).

In the Mendelian randomization framework, the associations of genotype with outcomes are of interest because of the strengthened inference about the action of the environmental modifiable risk factors that the genotypes proxy for rather than what genotypes say about genetic mechanisms per se. Mendelian randomization studies are aimed at informing strategies to reduce disease risk by influencing the nongenetic component of modifiable risk processes.

Mendelian Randomization: Is the Principle Sound?

The principle of Mendelian randomization relies on the basic (but approximate) laws of Mendelian genetics. If the probability that a postmeiotic germ cell that has received any particular allele at segregation contributes to a viable conceptus is independent of environment (following from Mendel's first law), and if genetic variants sort independently (following on from Mendel's second law), then, at a population level, these variants will not be associated with the confounding factors that generally distort conventional observational studies. This particular strength of genetic studies was explicitly recognized by the pioneering statistician R.A. Fisher from the 1920s onwards, as illustrated in his 1951 Bateson memorial lecture

> Genetics is indeed in a peculiarly favored condition in that Providence has shielded the geneticist from many of the difficulties of a reliably controlled comparison. The different genotypes possible from the same mating have been beautifully randomized by the meiotic process . . . Generally speaking, the geneticist, even if he foolishly wanted to, could not introduce systematic errors into the comparison of genotypes, because for most of the relevant time he has not yet recognized them. (Fisher, 1952)

Empirical evidence that there is lack of confounding of genetic variants with factors that confound exposures in

Downloaded from pps.sagepub.com by guest on October 27, 2010

conventional observational epidemiological studies comes from several sources. For example, consider the virtually identical allele frequencies in the British 1958 birth cohort and British blood donors (Wellcome Trust Case Control Consortium, 2007). Blood donors are clearly a very selected sample of the population, whereas the 1958 birth cohort comprised all births born in 1 week in Britain with minimal selection bias. Blood donors and the general population sample would differ considerably with respect to the behavioral, socioeconomic, and physiological risk factors that are the confounding factors in observational epidemiological studies. However, they hardly differ in terms of allele frequencies. Similarly, we have demonstrated the lack of association between a range of single nucleotide polymorphisms of known phenotypic effects and nearly 100 sociocultural, behavioral, and biological risk factors for disease (Davey Smith, Lawlor, et al., 2008).

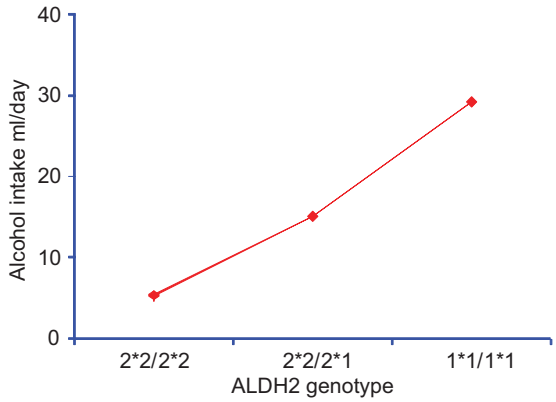


Fig. 1. Relationship between alcohol intake and ALDH2 genotype. Data from Takagi et al. (2002).

Mendelian Randomization in Practice

The principle of using genetic variation to proxy for a modifiable exposure was explicitly utilized in observational studies from the 1960s (Birge, Keutmann, Cuatrecasas, & Wheedon, 1967; Honkanen et al., 1996; Lower et al., 1979; Newcomer, Hodgson, Douglas, & Thomas, 1978), with the term *Mendelian randomization* being introduced by Richard Gray and Keith Wheatley in 1991 (Wheatley & Gray, 2004) in the context of an innovative genetically informed observational approach to assess the effects of bone marrow transplantation in the treatment of childhood acute myeloid leukemia. More recently, the term has been widely used in discussions of observational epidemiological studies (Clayton & McKeigue, 2001; Davey Smith & Ebrahim, 2003; Fallon, Ben-Shlomo, Elwood, Ubbink, & Davey Smith, 2001; Keavney, 2002; Youngman et al., 2000). Further discussion of the origin of this approach is given elsewhere (Davey Smith, 2006a).

There are several categories of inference that can be drawn from studies utilizing the Mendelian randomization approach. In the most direct forms, genetic variants can be related to the probability or level of exposure (''exposure propensity'') or to intermediate phenotypes believed to influence disease risk. Less direct evidence can come from genetic variant-disease associations that indicate that a particular biological pathway may be of importance, perhaps because the variants modify the effects of environmental exposures. Several examples from of these categories have been given elsewhere (Davey Smith 2006b; Davey Smith & Ebrahim, 2003; Davey Smith & Ebrahim, 2004; Davey Smith, Timpson, & Ebrahim, 2008; Ebrahim & Davey Smith, 2008); here, I will focus on studies of alcohol and various health and social outcomes that can be informed by this approach.

Alcohol Intake and Blood Pressure

The consequences of alcohol drinking for health range from the well-established (effects on liver cirrhosis and accidents) to the uncertain (CHD, depression, and dementia)—for example, the possible protective effect of moderate alcohol consumption on CHD risk remains highly controversial (Bovet & Paccaud, 2001; Klatsky, 2001; Marmot, 2001). Nondrinkers may be at a higher risk of CHD because health problems (perhaps induced by previous alcohol abuse) dissuade them from drinking (Shaper, 1993). As well as this form of reverse causation, confounding could play a role, with nondrinkers being more likely to display an adverse profile of socioeconomic or other behavioral risk factors for CHD (Hart, Davey Smith, Hole, & Hawthorne, 1999). Alternatively, alcohol may have a direct biological effect that lessens the risk of CHD—for example by increasing the levels of protective high density lipoprotein (HDL) cholesterol (Rimm, 2001). It is, however, unlikely that a randomized control trial (RCT) of differential levels of alcohol intake, which tests whether there is a protective effect of alcohol on CHD events, will ever be carried out.

Alcohol is oxidized to acetaldehyde, which in turn is used by aldehyde dehydrogenases (ALDHs) to oxidize to acetate. Half of Japanese people are heterozygotes or homozygotes for a null variant of *ALDH2*, and peak blood acetaldehyde concentrations after alcohol challenge are 18 times and 5 times higher among homozygous null variant and heterozygous individuals (respectively) in comparison with homozygous wild-type individuals (Enomoto, Takase, Yasuhara, & Takada, 1991). This renders the consumption of alcohol unpleasant through facial flushing, palpitations, drowsiness, and other symptoms. As Figure 1 shows, there are considerable differences in alcohol consumption according to genotype among men (Takagi et al., 2002). The principles of Mendelian randomization apply here: two factors that would be expected to be associated with alcohol consumption (age and cigarette smoking) and that confound conventional observational associations between alcohol and disease are not related to genotype despite the strong association of genotype with alcohol consumption.

It would be expected that the *ALDH2* genotype influences diseases known to be related to alcohol consumption, and it has been shown that *ALDH2* null variant homozygosity—associated with low alcohol consumption—is indeed related to a lower risk of liver cirrhosis (Chao et al., 1994). Considerable evidence,

Downloaded from pps.sagepub.com by guest on October 27, 2010

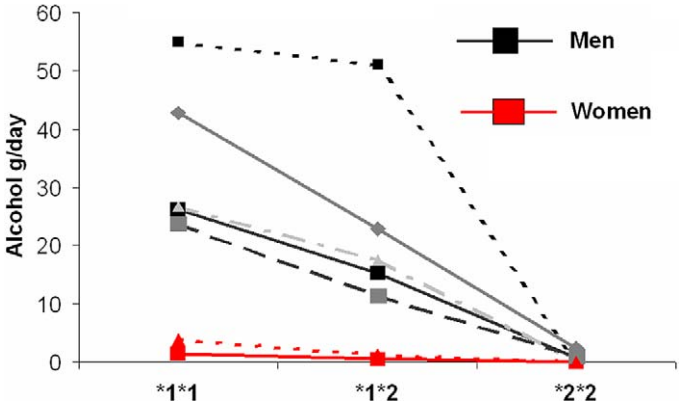


Fig. 2. ALDH2 genotype by alcohol consumption (g/day). (5 studies, $N = 6{,}815$; Chen et al., 2008.)

including data from short-term randomized controlled trials, suggests that alcohol increases HDL cholesterol levels (Burr, Fehily, Butland, Bolton, & Eastham, 1986; Haskell et al., 1984), which should protect against CHD. In line with this, the *ALDH2* genotype is strongly associated with HDL cholesterol in the expected direction, given the positive association between alcohol intake and HDL cholesterol levels in epidemiological studies (Takagi et al., 2002). With respect to blood pressure, observational evidence suggests that long-term alcohol intake produces an increased risk of hypertension and higher prevailing blood pressure levels—the results from different studies vary, and there is clearly a very large degree of potential confounding between alcohol and other exposures that would influence blood pressure. As in the case of vitamin E intake and CHD, we could be looking at a confounded association rather than a causal association. Indeed, evidence of controversy in this area is reflected by newspaper coverage of a recent study suggesting that moderate alcohol consumption has beneficial effects, even for hypertensive men (Beulens et al., 2007).

Evidence on the association of alcohol drinking and alcohol consumption blood pressure can come from studies of the *ALDH2* genotype and blood pressure. A meta-analysis of such studies suggests there is indeed a substantial positive effect of alcohol on blood pressure (Chen, Davey Smith, Harbord, & Lewis, 2008). As shown in Figure 2, alcohol consumption is strongly related to genotype among men, and despite higher levels of overall alcohol consumption in some studies, the shape of the association remains similar. However, in comparison with men, there is no evidence of association between drinking and genotype among women who drink very little.

Figure 3 demonstrates that men who are homozygous for the wild type have nearly two and a half times the risk of hypertension than men who are homozygous for the null variant. Heterozygous men who drink an intermediate amount of alcohol have a more modest elevated risk of hypertension than men who are homozygous for the null variant. Thus, a dose-response association of hypertension and genotype is seen, which is in line with the dose-response association between genotype and alcohol intake. Among men who are homozygous for the null variant and who drink considerably less alcohol than those who are homozygous for the wild type, systolic and diastolic blood pressures are considerably lower. By contrast, there is no association between genotype and blood pressure among women, for whom genotype is unrelated to alcohol intake (Fig. 4). The differential associations between genotype and blood pressure in men and women suggest that there is no other mechanism linking genotype and blood pressure than that relating to alcohol intake. If alternative pathways existed, then both men and women would be expected to have the same association between genotype and blood pressure.

In this example, the interaction is between a genetic variant and gender. Gender indicates substantial differences in alcohol consumption, which lead to the genotype being strongly associated with alcohol consumption in one group (males) but not in the other group (females) because of very low levels of alcohol consumption, irrespective of genotype, among the latter group. The power of this interaction is that it indicates that it is the association with alcohol intake and not any other aspects of the function of the genotype that is influencing blood pressure. If it were due to a pleiotropic effect of the genetic variation, then the association between genotype and blood pressure would be seen for women as well as men.

Alcohol and Illegal Substance Use: Testing the Gateway Hypothesis

In many contexts, people who drink alcohol manifest higher rates of illegal substance use. This could reflect common social and environmental factors that increase uptake of several behaviors or underlying genetic vulnerability factors. An alternative is the *gateway hypothesis* that postulates that alcohol use itself increases liability to initiate and maintain use of nonalcohol substance use (Irons, McGue, Iacono, & Oetting, 2007; Kandel & Yamaguchi, 1993; Kandel, Yamaguchi, & Chen, 1992). The Mendelian randomization approach has been applied in a study of East Asian Americans, all born in Korea but living in the United States from infancy, among whom *ALDH2* status was associated with alcohol use and alcohol use was associated with tobacco, marijuana, and other illegal drug use. However, *ALDH2* variation was not robustly associated with nonalcohol substance use, which provides evidence against the gateway hypothesis (Irons et al., 2007).

Maternal Drinking, the Intrauterine Environment, and Offspring Outcomes

The influence of high levels of alcohol intake by pregnant women on the health and development of their offspring is well recognized in the form of fetal alcohol syndrome (Gemma, Vichi, & Testai, 2007). However, outside of this extreme situation, the influence of alcohol is less easy to assess, particularly as higher levels of alcohol intake will be related to a wide array of potential sociocultural, behavioral, and environmental confounding factors. Furthermore, there may be systematic bias in how mothers report alcohol intake during pregnancy, which could distort associations with health outcomes. Therefore,

Downloaded from pps.sagepub.com by guest on October 27, 2010

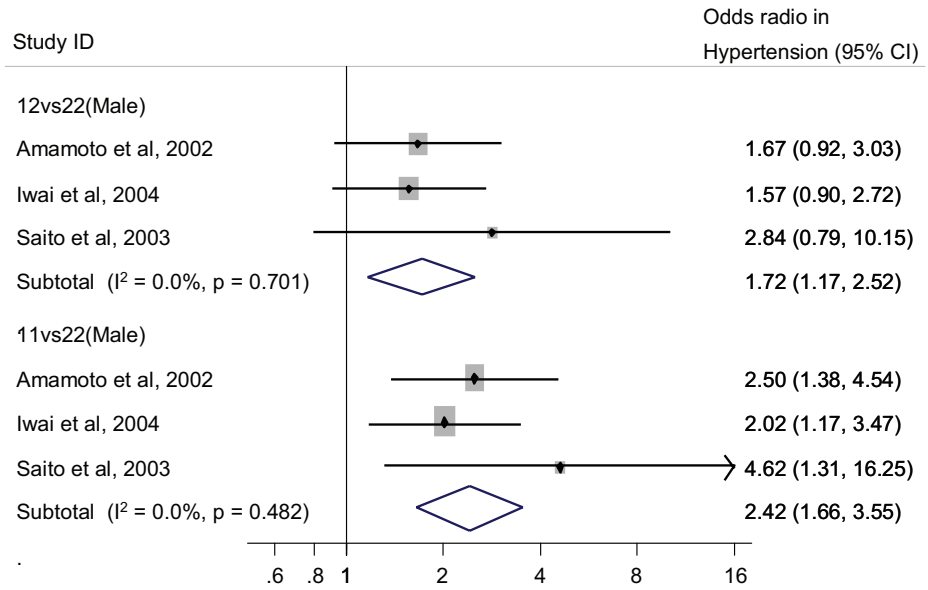


Fig. 3. Forest plot of studies of ALDH2 genotype and hypertension (Chen et al., 2008).

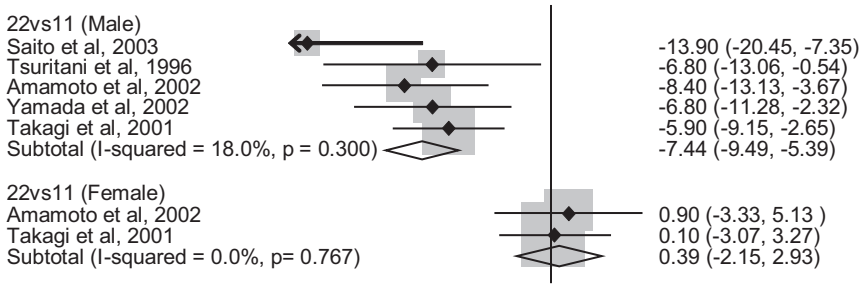


Fig. 4. ALDH2 genotype and systolic blood pressure (Chen et al., 2008).

outside of the case of very high alcohol intake by mothers, it is difficult to establish a causal link between maternal alcohol intake and offspring developmental characteristics. Some studies have approached this within the Mendelian randomization framework by investigating alcohol-metabolizing genotypes in mothers and offspring.

Studies have generally utilized a variant in the alcohol dehydrogenase gene (*ADH1B*3* allele). Alcohol dehydrogenase metabolizes alcohol to acetaldehyde and the *ADH1B* variant influences the rate of such metabolism. The *ADH1B*3* variant has a reasonable prevalence among African Americans and is related to faster alcohol metabolism. This can relate to a lower level of drinking, possibly because the faster metabolism leads to a more rapid spike in acetaldehyde with its aversive effects. At a given level of drinking, a faster metabolism will clear blood alcohol more rapidly, so the levels do not rise as high and they fall more quickly. Both of these processes, if occurring in the mother, would protect the fetus from the effects of alcohol. Some studies have selected mothers who have a universally high level of alcohol consumption and the alcohol-metabolizing genotypes among these mothers will relate to alcohol levels that could have a toxic effect on the developing fetus but not to their drinking, which is universally high. In this circumstance the genotypic differences will mimic the differences in level of alcohol intake with regard to the fetal exposure to maternal circulating alcohol. Although sample sizes have been low and the analysis strategies are not optimal, studies applying this approach provide some evidence to support the influence of maternal genotype, and thus of alcohol, on offspring outcomes (Gemma et al., 2007; Jacobson et al., 2006; Warren & Li, 2005). Studies that have been able to analyze both maternal genotype and fetal genotype find that it is the maternal genotype that is related to offspring outcomes, as anticipated if the crucial exposure related to maternal alcohol intake and alcohol levels.

As in other examples of Mendelian randomization, these studies are of relevance because they provide evidence of the influence of maternal alcohol levels on offspring development and not because they highlight a particular maternal genotype that is of importance. In the absence of alcohol drinking, the maternal genotype would presumably have no influence on offspring outcomes. Studies utilizing maternal genotype as a proxy for environmentally modifiable influences on the intrauterine environment can be analyzed in a variety of ways. First,

Downloaded from pps.sagepub.com by guest on October 27, 2010

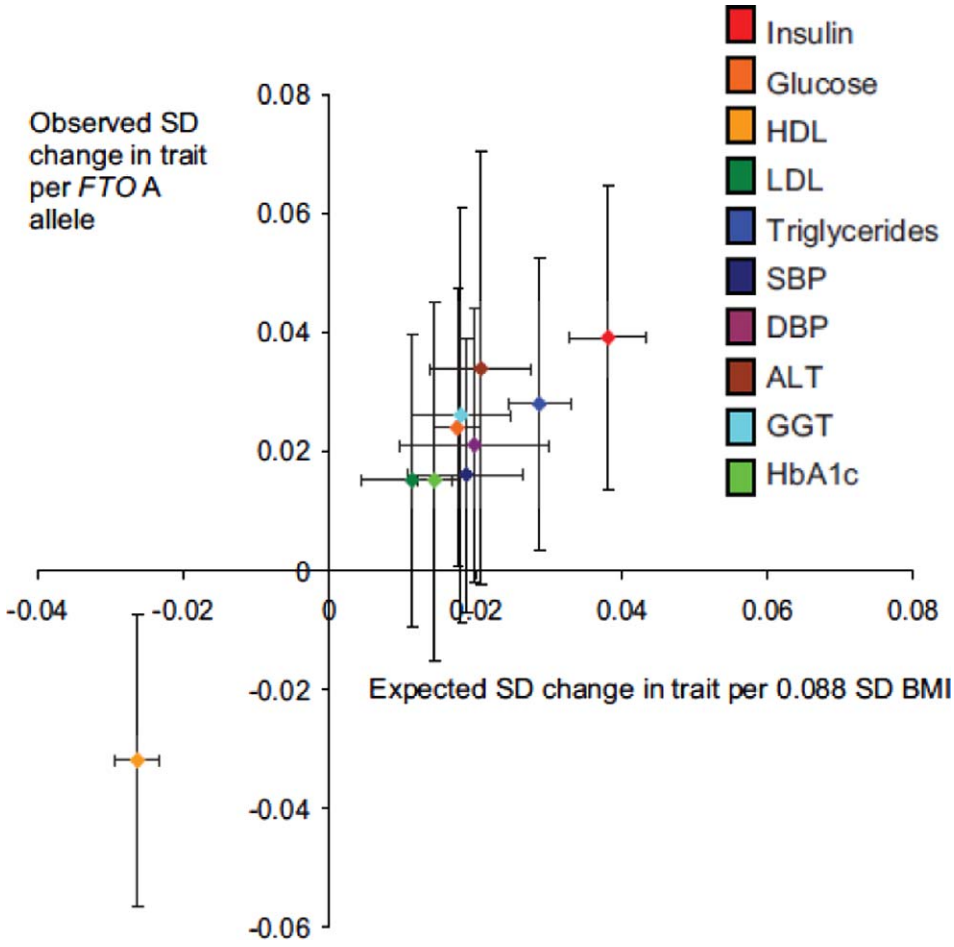


Fig. 5. The observed effects of *FTO* variation on metabolic traits. HDL = high density lipoproteins; LDL = low density lipoproteins; SBP = systolic blood pressure; DBP = diastolic blood pressure; ALT = alanine aminotransferase; GGT = gamma-glutamyltransferase. Adapted from Freathy et al. (2008).

the mothers of offspring with a particular outcome can be compared with a control group of mothers who have offspring without the outcome, in a conventional case-control design, but with the mother as the exposed individual (or control) rather than the offspring with the particular health outcome (or the control offspring). Fathers could serve as a control group when autosomal genetic variants are being studied. If the exposure is mediated by the mother, then maternal genotype, rather than offspring genotype, will be the appropriate exposure indicator. Clearly, maternal and offspring genotype are associated, but each are conditional based on the other—it should be the maternal genotype that shows the association with the health outcome among the offspring. Indeed, in theory it would be possible to simply compare genotype distributions of mothers and offspring, with a higher prevalence among mothers providing evidence that maternal genotype, through an intrauterine pathway, is of importance. However, the statistical power of such an approach is low, and an external control group, consisting of either fathers or women who have offspring without the health outcome, is generally preferable.

Other Examples of Mendelian Randomization: A Brief Catalog

Mendelian randomization has now been utilized in a wide variety of specific situations. Many of these relate to intermediate phenotypes; genotypic differences in such intermediate phenotypes can be related to genotypic influences on outcomes to investigate whether the intermediate phenotype causally influences disease outcome. Proof of this approach comes from situations in which the answer is known. For example, several genetic variants that are associated with blood cholesterol levels are also associated with CHD risk, in line with the substantial amount of evidence, including that from RCTs, that higher blood cholesterol levels causally increase disease risk (Davey Smith, Timpson, & Ebrahim, 2008).

These studies demonstrate another strength of the Mendelian randomization approach, in that the observed association of genotype with CHD is larger than that predicted from its effect on cholesterol levels and the magnitude of association of cholesterol levels with CHD risk identified in RCTs. This is

Downloaded from pps.sagepub.com by guest on October 27, 2010

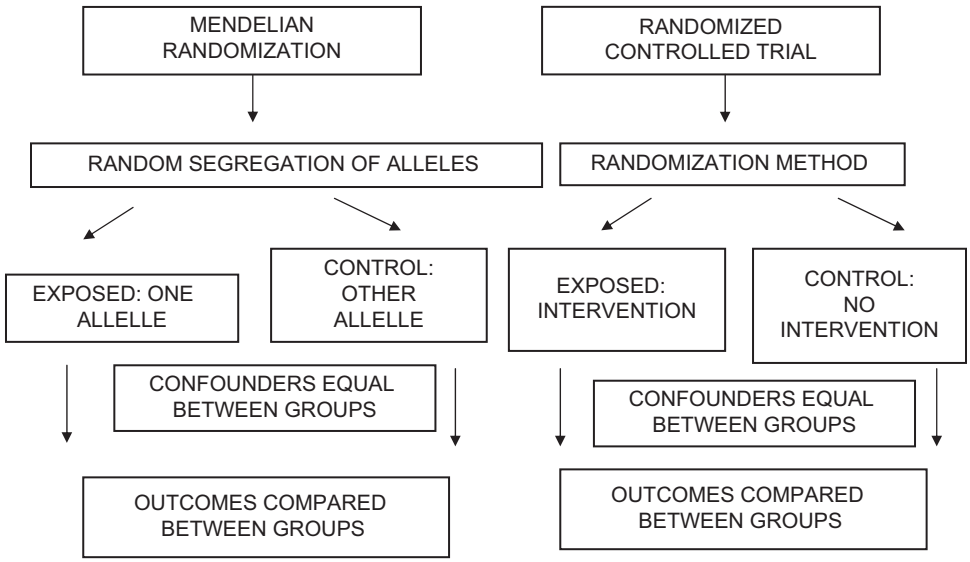


Fig. 6. Comparison of Mendelian randomization and randomized controlled trial designs.

to be expected, as RCTs only lower cholesterol for a few years, and atherosclerosis is a life-long process; in addition the genetic variants indicate differences in cholesterol levels over many decades, whereas RCTs produce relatively short-term changes. Genotypic differences in intermediate phenotypes can provide evidence of life-long, as opposed to short-term, influences of intermediate phenotypes on disease.

Another example of intermediate phenotype is seen in studies of the association of high body mass index (BMI) and a variety of cardiovascular risk factors. A variant in the *FTO* gene is robustly associated with differences in BMI, and, as shown in Figure 5, *FTO* variation predicts risk factor level to the degree expected, given its effect on BMI and a causal association between BMI and these risk factors (Freathy et al., 2008). Conversely, another intermediate phenotype, C-reactive protein (CRP), has been found to be strongly predictive of Type 2 diabetes and CHD risk. Genetic variants in the *CRP* gene that are related to differences in circulating CRP levels do not influence the risk of these diseases, suggesting that the observed associations are not causal (Lawlor et al., 2008; Timpson et al., 2005). This suggests that developing methods to pharmacotherapeutically lower CRP levels would not reduce disease risk, despite the strong observational associations.

Mendelian Randomization and RCTs

RCTs are clearly the definitive means of obtaining evidence on the effects of modifying disease risk processes. There are similarities in the logical structure of RCTs and Mendelian randomization as illustrated in Figure 6, which draws attention to the unconfounded nature of exposures for which genetic variants serve as proxies (analogous to the unconfounded nature of a randomized intervention), the impossibility of reverse causation as an influence on exposure-outcome associations in both Mendelian randomization and RCT settings, and the importance of intention to treat analyses (i.e., analysis by group defined by genetic variant, irrespective of associations between the genetic variant and the exposure for which this is a proxy within any particular individual).

The RCT analogy is also useful with respect to one objection that has been raised in conjunction with Mendelian randomization studies. Researchers have stated that the environmentally modifiable exposure for which genetic variants serve as proxies (such as alcohol intake) is influenced by many other factors in addition to the genetic variants (Jousilahti & Salomaa., 2004). This is of course true. However, consider an RCT of blood-pressure-lowering medication. Blood pressure is mainly influenced by factors other than taking blood-pressure-lowering medication—obesity, alcohol intake, salt consumption and other dietary factors, smoking, exercise, physical fitness, genetic factors and early-life developmental influences are all of importance. However, the randomization that occurs in trials ensures that these factors are balanced between the groups that receive the blood-pressure-lowering medication and those that do not. Thus, the fact that many other factors are related to the modifiable exposure does not compromise the power of RCTs, nor does it diminish the strength of Mendelian randomization designs. A related objection is that the genetic variants often explain only a trivial proportion of the variance in the environmentally modifiable risk factor for which the genetic variants are surrogate variables (Glynn, 2006). Again, consider an RCT of blood-pressure-lowering medication, in which 50% of participants receive the medication and 50% receive a placebo. If the antihypertensive therapy reduced blood pressure by a quarter of a standard deviation (i.e., a 5-mmHg reduction in systolic blood pressure, given that systolic blood pressure has a standard deviation of 20-mmHg in the population) then within the whole study group, treatment assignment (i.e., antihypertensive use versus placebo) will explain $5/20^2$ or 1.25% of the variance. In the example of

Downloaded from pps.sagepub.com by guest on October 27, 2010

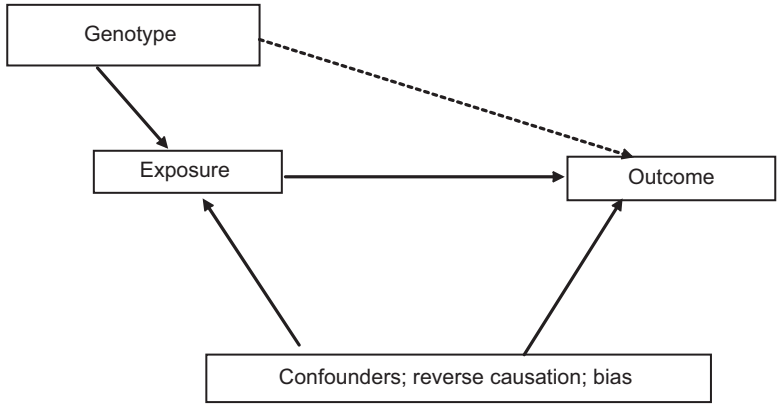


Fig. 7. Mendelian randomization as an instrumental variables approach.

ALDH2 variation and alcohol, the genetic variant explains about 2% of the variance in alcohol intake in the largest study available on this issue (Takagi et al., 2002). As can be seen, the quantitative association of genetic variants as instruments can be similar to that of randomized treatments with respect to biological processes that such treatments modify. Genetic variants are often equally strong—if not stronger—predictors of unconfounded differences in exposures as are the randomized treatments in RCTs.

Mendelian Randomization and Instrumental Variable Approaches

Mendelian randomization can also be likened to instrumental variable approaches that have been heavily utilized in econometrics and social science, although rather less so in epidemiology. In an instrumental variable approach, the instrument is a variable that is only related to the outcome through its association with the modifiable exposure of interest. The instrument is not related to confounding factors, nor is its assessment biased in a manner that would generate a spurious association with the outcome. Furthermore, the instrument will not be influenced by the development of the outcome (i.e., there will be no reverse causation). Figure 7 presents this basic schema, in which the dotted line between genotype and the outcome provides an unconfounded and unbiased estimate of the causal association between the exposure, for which the genotype is a proxy, and the outcome. The development of instrumental variable methods within econometrics, in particular, has led to a sophisticated suite of statistical methods for estimating causal effects, and these have now been applied within Mendelian randomization studies (Davey Smith, Harbord, Milton, Ebrahim, & Sterne, 2005b). The parallels between Mendelian randomization and instrumental variable approaches are discussed in more detail elsewhere (Lawlor et al., 2008; Thomas & Conti, 2004). The instrumental variable method allows for the estimation of the causal effect size of the modifiable environmental exposure of interest and the outcome together with estimates of the precision of the effect. Thus, in the example of alcohol intake (indexed by *ALDH2* genotype) and blood pressure, it is possible to utilize the joint associations of *ALDH2* genotype and alcohol intake and *ALDH2* genotype and blood pressure to estimate the causal influence of alcohol intake on blood pressure.

Alcohol, Esophageal, and Head and Neck Cancer: Gene × Environment Interaction, Cause, and Mechanism

A different form of Gene × Environment interaction than that discussed above in relation to gender-specific effects of *ALDH2* and blood pressure applies in the investigation of alcohol as a potential cause of esophageal and head and neck cancer. For these cancers, alcohol intake appears to increase the risk, although some have questioned the importance of its role (Memik, 2003). A meta-analysis of studies of *ALDH2* genotype and esophageal cancer risk (Lewis & Davey Smith, 2005) found that people who are homozygous for the null variant, and therefore consume considerably less alcohol, have a greatly reduced risk of esophageal cancer. The reduction in risk is close to that predicted from the size of effect of genotype on alcohol consumption and the dose response of alcohol on esophageal cancer risk (Burd, 2006). A similar picture is seen when examining the link to head and neck cancer risk (Boccia et al., 2009).

Thus, with respect to the homozygous null variant and homozygous wild type, the situation is similar to that of our blood pressure example: The genotypic association provides evidence of the effect of alcohol consumption, when researchers compare a group of infrequent drinkers to a group who drink considerable amounts of alcohol with no confounding factors differing between these groups. With respect to both esophageal and head and neck cancer, acetaldehyde (the metabolite that is increased in people carrying the null variant who do drink alcohol) is considered to be carcinogenic (Seitz & Stickel, 2007). Thus, drinkers who carry the null variant have higher levels of acetaldehyde than those who do not carry the variant. As shown above, people who are homozygous for the null variant drink very little alcohol, but heterozygous individuals

Downloaded from pps.sagepub.com by guest on October 27, 2010

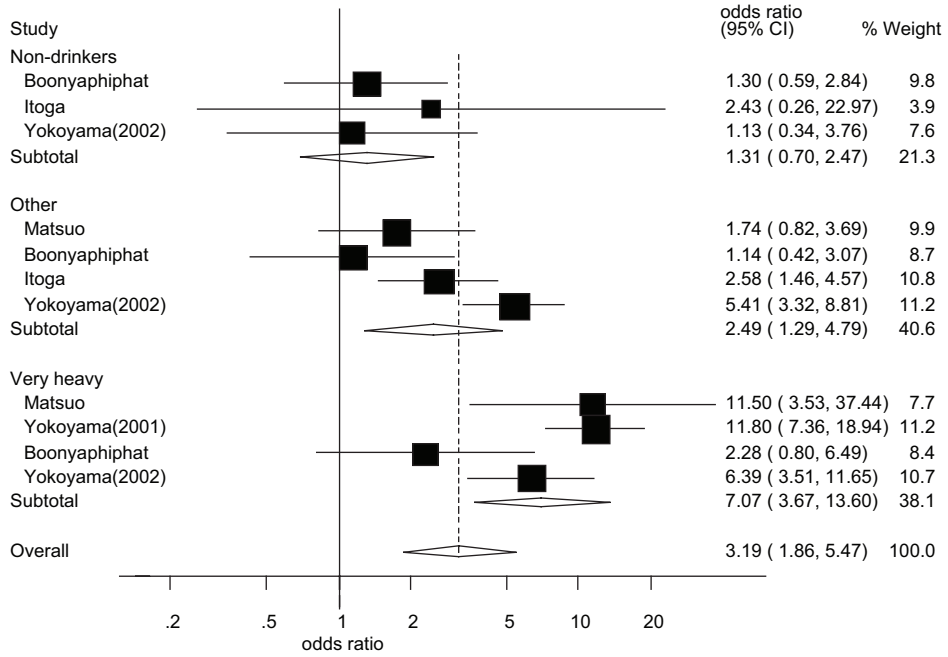


Fig. 8. Risk of esophageal cancer in individuals with the *ALDH2*1*2* genotype versus those with the *ALDH2*1*1* genotype. Non-drinkers = those who do not or never drink and exdrinkers; others = those who drink less than 60g of ethanol per day and were not habitual drinkers; heavy drinkers = 75 mL of ethanol per day for 5 or more days a week. Adapted from Lewis and Davey Smith (2005).

do drink. When the heterozygotes are compared with wild type homozygotes, an interesting picture emerges: The risk of esophageal cancer is higher in the heterozygotes, even though they drink less alcohol than the homozygotes. If alcohol itself acted directly as the causal factor, cancer risk would be intermediate in the heterozygotes compared with the other two groups. Acetaldehyde is the more likely causal factor, as heterozygotes drink some alcohol but metabolize it inefficiently, leading to accumulation of higher levels of acetaldehyde than would occur in homozygotes for the common variant, who metabolize alcohol efficiently, and homozygotes for the null variant, who drink insufficient alcohol to produce raised acetaldehyde levels. In Figure 8, the difference in esophageal cancer risk between *ALDH2* heterozygotes and those homozygous for the wild type are displayed, stratified by drinking status. In nondrinkers, there is no robust evidence of any association between genotype and esophageal cancer outcomes, as would be expected if the underlying environmentally modifiable causal factor were alcohol intake and the mechanism was acetaldehyde levels. In further support of the hypothesis, amongst people who were drinking alcohol there was increased risk in heterozygotes, who have higher acetaldehyde levels, and this was especially marked in heavy drinkers, who would have the greatest difference in acetaldehyde levels according to genotype. A similar analysis has been performed for head and neck cancer and again demonstrates no association of genotype and cancer risk in nondrinkers and a graded association according to alcohol intake level among alcohol drinkers (Boccia et al., 2009).

Gene × Environment Interactions Interpreted Within a Mendelian Randomization Framework

The meaning of Gene × Environment interactions has a contested history within human genetics. As James Tabery (Tabery, 2000, 2007) has discussed, two distinct concepts can be identified. First, there is a developmental concept, pioneered by Lancelot Hogben, which considers how gene–environment interplay influences particular developmental trajectories during ontogenesis. This notion can be contrasted with the biometric tradition, exemplified by R.A. Fisher, which considers interactions with respect to how much (if at all) they contribute to estimates of heritability. The clearest early statement of possible categories of Gene × Environment interaction came from one of the other founders of population genetics, J.B.S. Haldane, who tabulated the possible outcomes of gene–environment interplay as he saw them and stated that "the enumeration is so simple that no one has ever troubled to make it" (Haldane, 1938; see Box 1). What is noticeable from Haldane's typology is that many apparent Gene × Environment interactions discussed in the molecular genetics era will fall into his first category, in which there is no clear cross-over of effects of genotype according to environment but there is some apparent quantitative difference, with a genotype having a larger influence on phenotype in one environment than in another. Haldane considered a Gene × Environment interaction to be when one genotype was associated with a

Downloaded from pps.sagepub.com by guest on October 27, 2010

Box 1. J.B.S. Haldane on Gene × Environment Interaction

In his polemical book *Heredity and Politics*, Haldane presented the table below as an exhaustive list of the possibilities of Gene × Environment interaction. In the first situation, Genotype A is superior to Genotype B in each environment, and Environment X is more favorable than Environment Y independent of genotype. He considered mastiffs and dachshunds on a poor or good diet as an example of this—the mastiffs as a group would always be heavier than the dachshund, and those bred on a good diet would always be heavier than those on a poor diet. Within this basic arrangement, the exact quantitative way in which genotypic and environmental influences combined was not considered important by Haldane, but it the interactions within this conceptual space have received much attention in the current era of molecular genetic research.

In the second example, Genotype A performs better than Genotype B in Environment X, and Environment X provides better outcomes than Environment Y for both genotypes, but Genotype B performs better than Genotype A in Environment Y. Here Haldane considered Jersey cattle and Highland cattle, as both yield more milk on English pasture than on the Scottish highland moors, but the Jerseys perform better than the Highland cattle on pasture, and the Highland cattle perform better than the Jerseys on highland moors. He also used himself as an example of this type of interaction: "Had I been born in a Glasgow slum I should very probably have become a chronic drunkard, and if so I might by now be a good deal less intelligent than many men of a more stable temperament but less possibilities of intellectual achievement in a favorable environment." The third type of interaction involves Genotype A performing better than Genotype B independent of environment, but Environment X being better than Environment Y for Genotype A, whereas Environment Y is better than Environment X for Genotype B. Here, using the terminology of his day, he considered normal (A) and genetically mentally defective (B) children and found that although the first group performed better than the second in any type of school, the second group did better in special schools than in standard schools, and the normal children did better in standard schools than in special schools. Finally, in his fourth example, Genotype A performs better than Genotype B in Environment X but worse than Genotype B in Environment Y, and Environment Y produces superior outcomes among Genotype B but worse outcomes among Genotype A. This is clearly the most marked form of Gene × Environment interaction, and for his example Haldane considered how English-origin populations have a longer life span than long-term African-origin groups when living in England, and how long-term African-origin populations have a longer life span than English migrants in the African climate.

In Haldane's examples (also depicted in the figures), crossovers of effect occur when outcomes are tabulated according to gene and environment combinations. He did not explicitly discuss examples of situations in which a particular genotype has no influence on outcome in one environment but does influence outcome in another, although this could also be considered a form of qualitative interaction and has been a particular focus of some studies of Gene × Environment interaction.

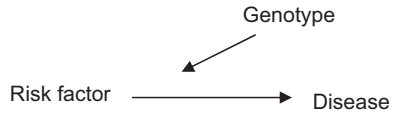
Table. Order of Achievement of Four Groups Designated by Genotypes A and B and Environments X and Y

		X	Y			X	Y
1.	A	1	2	or	A	1	3
	B	3	4		B	2	4
		X	Y				
2.	A	1	4				
	B	2	3				
		X	Y				
3.	A	1	2				
	B	4	3				
		X	Y			X	Y
4.	A	1	3		A	1	4
	B	4	2		B	3	2

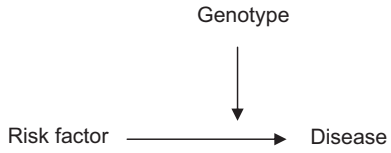
beneficial effect in one environment and with an adverse effect in another environment or vice versa. The latter can be referred to as *qualitative interactions*. A focus on qualitative interactions has clear advantages in that they are scale dependent (Thompson, 1991)—if any effect of genotype exists, then there must be an interaction on one scale (e.g., additive) if there is no interaction on another scale (e.g., multiplicative).

In an important series of papers, Ruth Ottman has explicated a typology of five models of Gene× Environment interactions. (Ottman, 1990, 1996, 2006). Here, I consider how these models

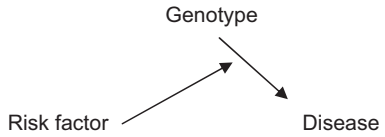
Downloaded from pps.sagepub.com by guest on October 27, 2010



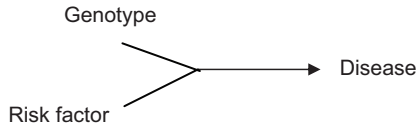
Model B: The genotype exacerbates the effect of the risk factor



Model C: The risk factor exacerbates the effect of the genotype



Model D: Both the genotype and the risk factor are required to raise risk



Model E: The genotype and risk factor each affect risk: combined effects can be additive or nonadditive

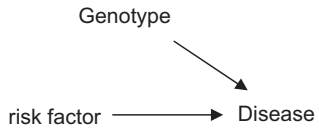


Fig. 9. Model A: Genotype increases expression of the risk factor. Model B: Genotype exacerbates the effect of the risk factor. Model C: Risk factor exacerbates the effect of the genotype. Model D: Both the genotype and the risk factor are required to raise risk. Model E: Genotype and risk factor each affect risk—combined effects can be additive or nonadditive.

would be interpreted within a Mendelian randomization framework. In Model A (Fig. 9), the genotype increases the level of expression of the risk factor, which in turn influences the risk of disease. In some definitions, this would not be interaction, but instead a causal chain of the kind that provide the essence of the Mendelian randomization approach. For example, the genotype could be the *ALDH2* null variant, which reduces alcohol intake and, through this, influences blood pressure in the manner discussed above. An example given by Ottman is of maternal phenylketonuria increasing the risk of mental retardation among the offspring due to the higher maternal blood levels of phenylalanine the fetus is exposed to, a form of intergenerational Mendelian randomization similar to that discussed earlier with respect to maternal alcohol metabolizing genotypes. The genotype has no effect if it is decoupled from the intermediate risk factor: for example, in a society in which few people drink alcohol (or among women in societies in which women drink little), a genotype will not be related to the disease outcomes, but it will be associated when it is coupled with the exposure (in this case, alcohol intake). Intermediate phenotype Mendelian randomization studies (e.g., genetic variants influencing cholesterol levels and through this CHD) are also examples of Model A.

Downloaded from pps.sagepub.com by guest on October 27, 2010

Table 1. Association of NAT2 Slow Acetylation Genotype With Bladder Cancer

Overall	Never Smokers	Ever Smokers
1.4 (1.2–1.7)	0.9 (0.6–1.3)	1.6 (1.3–1.9)

Note: Data are odds rations with 95% confidence intervals in parentheses. The baseline category is the odds ratio for bladder cancer with nonsmokers who are rapid NAT2 metabolizers. Adapted from Garcia Closas et al. (2005). $p < .01$.

Table 2. Association of Smoking Status and NAT2 Slow Acetylation Genotype With Bladder Cancer

Group	NAT2 Rapid	NAT2 Slow
Nonsmoker	1.0	0.9
Occasional smoker	1.2	1.6
Former smoker	2.4	4.1
Current smoker	5.2	7.5

Note: Adapted from Garcia-Closas et al. (2005).

In Model B (Fig. 9), the risk factor influences disease risk and the genotype modifies this, but the genotype will not influence outcomes on its own. In the absence of alcohol drinking, the variant will not be related to alcohol-related morbidity, but the variant will modify the severity of outcome in the presence of drinking, in the way that maternal *ADHD1B* is related to offspring outcomes among mothers who drink. Similarly carrying the wild-type *ALDH2* variant does not increase the risk of esophageal cancer in the absence of alcohol consumption, whereas alcohol consumption does increase risk of esophageal cancer risk even in the absence of *ALDH2* wild type, although to a lesser degree.

Model B (Fig. 9) also illustrates the influence of smoking tobacco on bladder cancer risk. Observational studies suggest an association, but it is clearly confounding and a variety of biases could generate such an association. The potential carcinogens in tobacco smoke relevant to bladder cancer risk include aromatic and heterocyclic amines, which are detoxified by N-acetyltransferase 2 (NAT2). Genetic variation in the *NAT2* gene leads to slower or faster acetylation states. If particular carcinogens in tobacco smoke do increase the risk of bladder cancer, then it would be expected that those with slow acetylate states, who have a reduced rate of detoxification of these carcinogens, would be at an increased risk of bladder cancer if they were smokers, whereas if they were not exposed to these carcinogens (the major exposure route for those outside of particular industries being through tobacco smoke) then an association of genotype with bladder cancer risk would not be anticipated (see Table 1; Garcia-Closas et al., 2005). The influence of the *NAT2* slow acetylation genotype is only appreciable among those also exposed to heavy smoking. As the genotype will be unrelated to confounders, it is difficult to reason why this situation should arise unless smoking is a causal factor with respect to bladder cancer. Thus, the presence of a sizable effect of genotype in the exposed group, but not in the unexposed group, provides evidence as to the causal nature of the environmentally modifiable risk factor—in this example, smoking. Table 2 illustrates that smoking has detrimental effects on bladder cancer risk in both genotype groups, and the somewhat lower risk amongst one group does not indicate that targeting prevention policies would be a useful strategy for public health (Davey Smith, Ebrahim, et al., 2005).

In Model C (Fig. 9), the genotype has a direct effect on disease risk whereas the risk factor does not have this effect when acting by itself. Examples of this are found in pharmacogenetics, where an otherwise benign exposure has a detrimental influence if accompanied by a particular genotype that increases the risk of adverse outcome even when the exposure is not present. Ottman discusses the autosomal dominant condition *porphyria variegate*, which increases risk of various skin conditions. Barbiturate use is generally benign, but in the presence of porphyria genotype it leads to very severe attacks of skin blistering. Model D (Fig. 9) is similar to Model C, but both modifiable and genetic risk factor do not produce outcomes alone in the latter case—they only do so in combination. For example, Stevens-Johnson syndrome can occur with carbamazepine use among individuals carrying the *HLA-B1502* allele. Models C and D do not allow for Mendelian randomization focused on the identification of environmentally modifiable risk factors that influence disease risk in the whole population, but they benefit from the Mendelian randomization principle in that randomization of the drug therapy is not required, given that the genotypes are essentially randomized with respect to use of the drug during periods before the interactions are detected and genetic testing allows for avoiding treating susceptible individuals. This is a specific example of how observational studies of unexpected adverse treatment consequences do not generally suffer from the same problems of confounding and bias that are experienced in conventional observational studies of risk factors for disease.

Model E (Fig. 9) refers to the situation in which genotype and risk factor both independently influence disease risk. The expected joint effect could be additive or multiplicative (given the scale dependence issue discussed above) and within model E the effect can either be of the expected order: either greater than anticipated (synergistic) or less than anticipated (antagonistic). If the genotype serves as a proxy for a modifiable cause of disease, then Model E is simply an expanded version of any Mendelian randomization study. The genotype would be expected to combine with other risk factors in the same way as would the modifiable risk factor it is a proxy for, with the advantage that the genotype provides more robust evidence of the causal effect of the modifiable risk factor. If a directly measured risk factor is studied, then confounding and bias can influence how the effect combines with other risk factors. For example, the joint effect of smoking and alcohol consumption on health outcomes could be investigated through study of *ALDH2* variation, smoking, and outcome. In some situations, genetic variation does not directly influence risk factor levels (as in Model A), but could proxy for such risk factor levels through influencing response to the risk factor. For example, genetic variation in the vitamin D receptor which does not influence vitamin D levels can proxy for such differences though being related to differential biological response to a given level of vitamin D. Studying how both levels and genetic

Downloaded from pps.sagepub.com by guest on October 27, 2010

variation relate to disease outcomes can provide evidence of the causal action of vitamin D levels in this situation, as concordance would support a direct biological (as opposed to biased or confounded) link between vitamin D and disease.

Problems and Limitations of Mendelian Randomization

The Mendelian randomization approach provides useful evidence on the influence of modifiable exposures on health outcomes. However, there are several limitations to this approach. These have been discussed at considerable length elsewhere (Davey Smith & Ebrahim, 2003; Ebrahim & Davey Smith, 2008) and therefore only some issues of particular relevance for Gene × Environment interaction are briefly considered here.

Confounding of Genotype, Environmentally Modifiable Risk Factors, and Disease Associations

The power of Mendelian randomization lies in its ability to avoid the often substantial confounding seen in conventional observational epidemiology. However, confounding can be reintroduced into Mendelian randomization studies, and this possibility needs to be considered when interpreting the results. First, it is possible that the locus under study is in linkage disequilibrium (i.e., is associated) with another polymorphic locus, with the former being confounded by the latter. It may seem unlikely, given the relatively short distances over which linkage disequilibrium is seen in the human genome, that a polymorphism influencing, for instance, CHD risk, would be associated with another polymorphism influencing CHD risk (and thus these being confounding between the two genetic variants). There are, nevertheless, examples of different genes influencing the same metabolic pathway being in physical proximity. For example, different polymorphisms influencing alcohol metabolism appear to be in linkage disequilibrium (Osier et al., 2002).

Second, Mendelian randomization is most useful when it can be used to relate a single intermediate phenotype to a disease outcome. However, polymorphisms may (and probably often will) influence more than one intermediate phenotype, and this may mean they proxy for more than one environmentally modifiable risk factor. This pleiotropy can be generated through multiple effects mediated by their RNA expression or protein coding; through alternative splicing, in which one polymorphic region contributes to alternative forms of more than one protein (Glebart, 1998); or through other mechanisms. The most robust interpretations will be possible when the functional polymorphism appears to directly influence the level of the intermediate phenotype of interest (as in the cholesterol example), but such examples are probably going to be less common in Mendelian randomization than in cases in which the polymorphism could in principle influence several systems, with different potential interpretations of how the effect on outcome is generated.

Linkage disequilibrium and pleiotropy can reintroduce confounding and thus reduce the potential value of the Mendelian randomization approach. Genomic knowledge may help in estimating the degree to which these are likely to be problems in any particular Mendelian randomization study, through, for instance, explication of genetic variants that may be in linkage disequilibrium with the variant under study or the function of a particular variant and its known pleiotropic effects. Furthermore, genetic variation can be related to measures of potential confounding factors in each study, and the magnitude of such confounding can be estimated. Empirical studies to date suggest that common genetic variants are largely unrelated to the behavioral and socioeconomic factors considered to be important confounders in conventional observational studies. However, relying on measurement of confounders does, of course, remove the central purpose of Mendelian randomization, which is to balance unmeasured as well as measured confounders.

In some circumstances, the genetic variant will be related to the environmentally modifiable exposure of interest in some population subgroups but not in others. The alcohol *ALDH2* genotype and blood pressure association affecting men (but not women), discussed earlier, is an example of this. If *ALDH2* genetic variation influenced blood pressure for reasons other than its influence on alcohol intake—for example, if it was in linkage disequilibrium with another genetic variant that influenced blood pressure through another pathway, or if there was a direct pleiotropic effect of the genetic variant on blood pressure—then the same genotype-blood pressure association should be seen among both men and women. If the genetic variant only influences blood pressure through its effect on alcohol intake, an effect should only be seen in men, which is what is observed. This further strengthens the evidence that the association between genotype and blood pressure depends on the genotype influencing alcohol intake and that the associations do indeed provide causal evidence of an influence of alcohol intake on blood pressure.

In some cases, it may be possible to identify two separate genetic variants that are not in linkage disequilibrium with each other but that both serve as proxies for the environmentally modifiable risk factor of interest. If both variants are related to the outcome of interest and point to the same underlying association, then it becomes much less plausible that reintroduced confounding explains the association, as it would have to be acting in the same way for these two unlinked variants. This can be likened to RCTs of different blood pressure lowering agents, which work through different mechanisms and have different potential side effects but lower blood pressure to the same degree. If the different agents produce the same reductions in cardiovascular disease risk, then it is unlikely that this is through agent-specific effects of the drugs—rather, it points to blood pressure lowering as being key. We previously discussed investigating the effect of alcohol on risk of head and neck cancer by comparing risk among *ALDH2* homozygous

Downloaded from pps.sagepub.com by guest on October 27, 2010

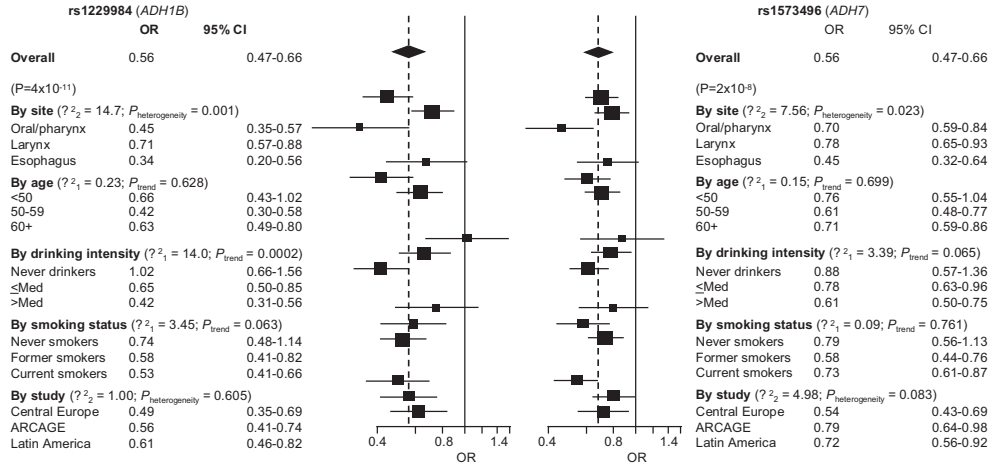


Fig. 10. The interaction between alcohol intake, *ADH* variation, and head and neck cancer risk. Adapted from Hashibe et al. (2008).

wild type and *ALDH2* homozygous null variant men; the same issue has been addressed by studying the interaction between alcohol intake, *ADH* variation, and head and neck cancer risk (Fig. 10; Hashibe et al., 2008), in which the influence of genotype among drinkers, but not among nondrinkers, provides evidence as to the causal role of alcohol. In another context, two distinct genetic variants acting as instruments for higher body fat content have been used to demonstrate that greater adiposity is related to higher bone mineral density (Timpson, Sayers, Davey Smith, & Tobias, 2009). With the large number of genetic variants that are being identified in genome wide association studies in relation to particular phenotypes (e.g., more than 50 independent variants that are related to height, more than 10 that are related to total cholesterol, and more than 20 related to fasting glucose) it is possible to generate many independent combinations of such variants and, from these, many independent instrumental variable estimates of the causal associations between an environmentally modifiable risk factor and a disease outcome. The independent estimates will not be plausibly influenced by any common pleiotropy or linkage-disequilibrium-induced confounding, and therefore, any consistency displayed provides strong evidence against any interpretation that reintroduced confounding is generating the associations.

Special Issues With Confounding in Studies of Gene × Environment Interactions

It must be recognized that Gene × Environment interactions interpreted within the Mendelian randomization framework as evidence regarding the causal nature of environmentally modifiable exposures are not protected from confounding to the same extent as main genetic effects. In the example regarding *NAT2*, smoking, and bladder cancer, any factor related to smoking—such as social class—will tend to show a greater association with bladder cancer within NAT2 slow acetylators than within NAT2 rapid acetylators. Because there is not a 1-to-1 association of social class with smoking, this will not result in an effect of the genotype in one social class stratum and no effect in the other social class stratum, as in the *NAT2*/smoking interaction, but rather a qualitative interaction of a greater effect of *NAT2* in the poorer social classes (amongst whom smoking is more prevalent) and a smaller (but still evident) effect in the better-off social classes, amongst whom smoking tends to be less prevalent. Thus, situations in which both the biological basis of an expected interaction is well understood and in which a qualitative (effect vs. no effect) interaction may be postulated are the ones that are most amenable to interpretations related to the general causal nature of the environmentally modifiable risk factor.

Canalization and Developmental Stability

Perhaps a greater potential problem for Mendelian randomization than reintroduced confounding arises from the developmental compensation that may occur through a polymorphic genotype being expressed during fetal or early postnatal development and thus influences development in such a way as to buffer against the effect of the polymorphism. Such compensatory processes have been discussed since C.H. Waddington introduced the notion of canalization in the 1940s (Waddington, 1942). Canalization refers to the buffering of the effects of either environmental or genetic forces attempting to perturb development and Waddington's ideas have been well developed both empirically and theoretically (Debat & David, 2001; Gibson & Wagner, 2000; Hartman, Garvik, & Hartwell, 2001; Hornstein & Shomron, 2006; Kitami & Nadeau, 2002; Rutherford, 2000; Wilkins, 1997). Such buffering can be achieved either through genetic redundancy (more than one gene having the same or similar function) or through alternative metabolic routes, in which the complexity of metabolic pathways allows recruitment of different pathways to reach the same phenotypic endpoint. In effect, a functional polymorphism expressed during fetal development or postnatal growth may influence the expression of a wide range of other genes, leading to changes that may compensate for the influence of

Downloaded from pps.sagepub.com by guest on October 27, 2010

the polymorphism. Put crudely, if a person has developed and grown from the intrauterine period onwards within an environment in which one factor is perturbed (e.g., elevated cholesterol levels due to genotype) then they may be rendered resistant to the influence of life-long elevated circulating cholesterol, through permanent changes in tissue structure and function that counterbalance its effects. In intervention trials—for example, RCTs of cholesterol-lowering drugs—the intervention is generally randomized to participants during their middle age; similarly, in observational studies of this issue, cholesterol levels are ascertained during adulthood. In Mendelian randomization, on the other hand, randomization occurs before birth. This leads to important caveats when attempting to relate the findings of conventional observational epidemiological studies to the findings of studies carried out within the Mendelian randomization paradigm.

In some Mendelian randomization designs, developmental compensation is not an issue. For example, when maternal genotype is utilized as an indicator of the intrauterine environment (e.g., maternal *ADH* variation discussed above), then the response of the fetus will not differ whether the effect is induced by maternal genotype or by environmental perturbation, and the effect on the fetus can be taken to indicate the effect of environmental influences during the intrauterine period. Also in cases in which a variant influences an adulthood environmental exposure (e.g., *ALDH2* variation and alcohol intake), developmental compensation to genotype will not be an issue. This also applies in Gene × Environment interactions that are interpreted with respect to causality of the environmental factor, as development will not have occurred in the presence of the modifiable risk factor of interest and thus developmental compensation will not have occurred.

Lack of Suitable Genetic Variants to Proxy for Exposure of Interest

An obvious limitation of Mendelian randomization is that it can only examine areas for which there are functional polymorphisms (or genetic markers linked to such functional polymorphisms) that are relevant to the modifiable exposure of interest. In the context of genetic association studies, it has been pointed out more generally that there may be no suitable marker or functional polymorphism to allow study of this process in many cases, even if a locus is involved in a disease-related metabolic process (Weiss & Terwilliger, 2000). In an earlier paper on Mendelian randomization (Davey Smith & Ebrahim, 2003), we discussed the example of vitamin C, as observational epidemiology appeared to have got the wrong answer related to associations between vitamin C levels and disease. We considered whether the association between vitamin C and CHD could have been studied utilizing the principles of Mendelian randomization. We stated that polymorphisms exist that are related to lower circulating vitamin C levels—for example, in the haptoglobin gene (Langlois, Delanghe, De Buyzere, Bernard, & Ouyang,1997)—but in this case, the effect on vitamin C is not direct and these other phenotypic differences could have an influence on CHD risk that would distort examination of the influence of vitamin C levels through relating genotype to disease. *SLC23A1*—a gene encoding for the vitamin C transporter SVCT1, which is involved in vitamin C transport by intestinal cells—would be an attractive candidate for Mendelian randomization studies. However, by 2003 (the date of our earlier paper), a search for variants had failed to find any common SNP that could be used in such a way (Erichsen, Eck, Levine, & Chanock, 2001). We therefore used this as an example of a situation in which suitable polymorphisms for studying the modifiable risk factor of interest could not be located. However, since the earlier paper was written, researchers have identified functional variation in *SLC23A1* that is related to circulating vitamin C levels (Timpson et al., 2010). We use this example not to suggest that the obstacle of locating relevant genetic variation for particular problems is observational—epidemiology will always be overcome—but to point out that rapidly developing knowledge of human genomics will identify more variants that can serve as instruments for Mendelian randomization studies.

Conclusions: What Mendelian Randomization Is and Is Not

Mendelian randomization is not predicated on the assumption that genetic variants are major determinants of health and disease within populations. There are many cogent critiques of genetic reductionism and the overselling of ''discoveries'' in genetics that reiterate obvious truths so clearly (albeit somewhat repetitively) that there is no need to repeat them here (e.g., Baird, 2000; Berkowitz, 1996; Holtzman, 2001; Strohman, 1993). Mendelian randomization does not depend upon there being genes for particular traits, and certainly not in the strict sense of a gene for a trait being one that is maintained by selection because of its causal association with that trait (Kaplan & Pigliucci, 2001). The association of genotype and the environmentally modifiable factor that it proxies for will be like most genotype–phenotype associations, one that is contingent and cannot be reduced to individual level prediction but will pertain at a group level within environmental limits (Wolf, 1995). This is analogous to an RCT of antihypertensive agents, in which the group randomized to active medication will have lower mean blood pressure than the group randomized to placebo at a collective level, but many participants randomized to active treatment will have higher blood pressure than many individuals randomized to placebo at an individual level. Group level differences are what create the analogy between Mendelian randomization and RCTs, as outlined in Figure 13.

Finally, the associations that Mendelian randomization depend on do need to pertain to a definable group at a particular time, but they do not need to be immutable. Thus, *ALDH2* variation will not be related to alcohol consumption in a society in which alcohol is not consumed; the association will vary by gender and by cultural group, and it may change over time (Hasin et al., 2002; Higuchi et al., 1994). Within the setting of a study of a well-defined group, however, the genotype will be associated with group-level differences in alcohol

Downloaded from pps.sagepub.com by guest on October 27, 2010

consumption, and group assignment will not be associated with confounding variables.

Critiques of contemporary genetic epidemiology often focus on two features of findings from genetic association studies: that the population attributable risk of the genetic variants is low and that the influence of genetic factors is not reversible in any case. Illustrating both of these criticisms, Terwilliger and Weiss suggest the following as reasons for considering that many of the current claims regarding genetic epidemiology are hype: (a) that alleles identified as increasing the risk of common diseases "tend to be involved in only a small subset of all cases of such diseases" (Terwilliger & Weiss, 2003, p. 532) and (b) that in any case "while the concept of attributable risk is an important one for evaluating the impact of removable environmental factors, for non-removable genetic risk factors, it is a moot point" (Terwilliger & Weiss, 2003, p. 540). These evaluations of the role of genetic epidemiology are not relevant when considering the potential contributions of Mendelian randomization. This approach is not concerned with the population attributable risk of any particular genetic variant, but the degree to which associations between the genetic variant and disease outcomes can demonstrate the importance of environmentally modifiable factors as causes of disease, for which the population attributable risk is of relevance to public health prioritization. Consider, for example, the case of familial hypercholesterolaemia or familial defective Apoliprotein B. The genetic mutations associated with these conditions will only account for a trivial percentage of cases of CHD within the population (i.e., the population attributable risk will be low). For example, in a Danish population, the frequency of familial defective Apo B is 0.08%, and, despite its sevenfold increased risk of CHD, it only generates a population attributable risk of 0.5% (Tybjaerg-Hansin, Steffensen, Meinertz, Schnohr, & Nordestgaard, 1998). However, by identifying blood cholesterol levels as a causal factor for CHD, the triangular association between genotype, blood cholesterol, and CHD risk identifies an environmentally modifiable factor with a very high population attributable risk: Assuming that 50% of the population have raised blood cholesterol above 6.0 mmol/l and that this is associated with a relative twofold risk of CHD, a population attributable risk of 33% is obtained. The same logic applies to the other examples discussed above—the attributable risk of the genotype is low, but the population attributable risk of the modifiable environmental factor identified as causal through the genotype–disease associations is large. The same reasoning applies when considering the suggestion that genotype–disease associations are not of public health importance as genotype cannot be modified (Terwilliger & Weiss, 2003). The point of Mendelian randomization approaches is not to attempt to modify genotype, but to utilize genotype–disease associations to strengthen inferences regarding modifiable environmental risks for disease and then reduce disease risk in the population through applying this knowledge.

Mendelian randomization differs from other contemporary approaches to genetic epidemiology in that its central concern is not with the magnitude of genetic variant influences on disease, but rather on what the genetic associations tell us about environmentally modifiable causes of disease. As David B. Abrams, former director of the Office of Behavioral and Social Sciences Research at the U.S. National Institutes of Health has said, "The more we learn about genes the more we see how important environment and lifestyle really are" (Abrams, 2002). Many years earlier, the pioneering geneticist Thomas Hunt Morgan articulated a similar sentiment in his Nobel Prize acceptance speech when he contrasted his views with eugenics, the then-popular genetic approach to disease. He thought that "through public hygiene and protective measures of various kinds we can more successfully cope with some of the evils that human flesh is heir to. Medical science will here take the lead—but I hope that genetics can at times offer a helping hand" (Morgan, 1935). More than seven decades later, it might now be time for genetic research to strengthen the knowledge base of public health directly.

Acknowledgments

Thank you to Tom Palmer, who estimated the variance explained by ALDH2 genotype for alcohol consumption, and to Caroline Relton, Ken Weiss, Debbie Lawlor, Ezra Susser, Maria Glymour, Marc Schuckit, and David Reiss for comments on an earlier draft of this article.

Declaration of Conflicting Interests

The author declared that he had no conflicts of interest with respect to his authorship or the publication of this article.

References

Abrams, D.B. (2002). President's message. *Quarterly Newsletter of the Society of Behavioral Medicine*, pp. 1–2, 7, 15. Retrieved from http://www.sbm.org/outlook/02_03_Winter.pdf

Baird, P. (2000). Genetic technologies and achieving health for populations. *International Journal of Health Services*, *30*, 407–424.

Berkowitz, A. (1996). Our genes, ourselves? *Bioscience*, *46*, 42–51.

Beulens, J.W.J., Rimm, E.B., Ascherio, A., Spiegelman, D., Hendriks, H.F.J., & Mukamal, K.J. (2007). Alcohol consumption and risk for coronary heart disease among men with hypertension. *Annals of Internal Medicine*, *146*, 10–19.

Birge, S.J., Keutmann, H.T., Cuatrecasas, P., & Wheedon, G.D. (1967). Osteoporosis, intestinal lactase deficiency and low dietary calcium intake. *New England Journal of Medicine*, *276*, 445–448.

Boccia, S., Hashibe, M., Galli, P., De Feo, E., Asakage, T., Hashimoto, T., et al. (2009). Aldehyde dehydrogenase 2 and head and neck cancer: A meta-analysis implementing a Mendelian randomization approach. *Cancer Epidemiology Biomarkers and Prevention*, *18*, 248–254.

Bovet, P., & Paccaud, F. (2001). Alcohol, coronary heart disease and public health: Which evidence-based policy? *International Journal of Epidemiology*, *30*, 734–737.

Burd, L.J. (2006). Interventions in FASD: we must do better. *Child Care Health Development*, *33*, 398–400.

Burr, M.L., Fehily, A.M., Butland, B.K., Bolton, C.H., & Eastham, R.D. (1986). Alcohol and high-density-lipoprotein

Downloaded from pps.sagepub.com by guest on October 27, 2010

cholesterol: A randomized controlled trial. *British Journal of Nutrition*, *56*, 81–86.

Caspi, A., Sugden, K., Moffitt, T.E., Taylor, A., Craig, I.W., Harrington, H., et al. (2003). Influence of life stress on depression: Moderation by a polymorphism in the 5-HTT gene. *Science*, *301*, 386–389.

Caspi, A., Williams, B., Kim-Cohen, J., Craig, I.W., Milne, B.J., Poulton, R., et al. (2007). Moderation of breastfeeding effects on the IQ by genetic variation in fatty acid metabolism. *Proceedings of the National Academy of Sciences, USA*, *104*, 18860–18865.

Chao, Y.-C., Liou, S.-R., Chung, Y.-Y., Tang, H.-S., Hsu, C.-T., Li, T.-K., et al. (1994). Polymorphism of alcohol and aldehyde dehydrogenase genes and alcoholic cirrhosis in Chinese patients. *Hepatology*, *19*, 360–366.

Chen, L., Davey Smith, G., Harbord, R., & Lewis, S. (2008). Alcohol intake and blood pressure: A systematic review implementing Mendelian randomization approach. *PLoS Medicine*, *461*, 5.

Clayton, D., & McKeigue, P.M. (2001). Epidemiological methods for studying genes and environmental factors in complex diseases. *Lancet*, *358*, 1356–1360.

Darwin, C. (1876). *The effects of cross and self fertilisation in the vegetable kingdom*. London: John Murray.

Darwin, G.H. (1875). Marriages between first cousins and their effects. *Journal of the Statistical Society of London*, *38*, 153–184.

Davey Smith, G. (2006a). *Capitalising on Mendelian randomization to assess the effects of treatments*. Retrieved from http://www.jameslindlibrary.org/trial_records/20th_Century/1990s/gray/gray-commentary.html

Davey Smith, G. (2006b). Cochrane lecture: Randomised by (your) god—robust inference from an observational study design. *Journal of Epidemiology and Community Health*, *60*, 382–388.

Davey Smith, G., & Ebrahim, S. (2003). Mendelian randomization: Can genetic epidemiology contribute to understanding environmental determinants of disease? *International Journal of Epidemiology*, *32*, 1–22.

Davey Smith, G., & Ebrahim, S. (2004). Mendelian randomization: Prospects, potentials, and limitations. *International Journal of Epidemiology*, *33*, 30–42.

Davey Smith, G., & Ebrahim, S. (2005). What can Mendelian randomization tell us about modifiable behavioural and environmental exposures. *British Medical Journal*, *330*, 1076–1079.

Davey Smith, G., Ebrahim, S., Lewis, S., Hansell, A.L., Palmer, L.J., & Burton, P.J. (2005). Genetic epidemiology and public health: Hope, hype, and future prospects. *Lancet*, *366*, 1484–1498.

Davey Smith, G., Harbord, R., Milton, J., Ebrahim, S., & Sterne, J.A.C. (2005). Does elevated plasma fibrinogen increase the risk of coronary heart disease? Evidence from a meta-analysis of genetic association studies. *Arteriosclerosis Thrombosis and Vascular Biology*, *25*, 2228–2233.

Davey Smith, G., Lawlor, D., Harbord, R., Timpson, N., Day, I., & Ebrahim, S. (2008). Clustered environments and randomized genes: A fundamental distinction between conventional and genetic epidemiology. *PLoS Medicine*, *4*, 1985–1992.

Davey Smith, G., Timpson, N., & Ebrahim, S. (2008). Strengthening causal inference in cardiovascular epidemiology through Mendelian randomization. *Annals of Medicine*, *12*, 524–541.

Debat, V., & David, P. (2001). Mapping phenotypes: Canalization, plasticity and developmental stability. *Trends in Ecology and Evolution*, *16*, 555–561.

Ebrahim, S., & Davey Smith, G. (2008). Mendelian randomization: Can genetic epidemiology help redress the failures of observational epidemiology? *Human Genetics*, *123*, 15–33.

Enomoto, N., Takase, S., Yasuhara, M., & Takada, A. (1991). Acetaldehyde metabolism in different aldehyde dehydrogenase-2 genotypes. *Alcoholism Clinical and Experimental Research*, *15*, 141–144.

Erichsen, H.C., Eck, P., Levine, M., & Chanock, S. (2001). Characterization of the genomic structure of the human vitamin C transporter SVCT1 (SLC23A2). *Journal of Nutrition*, *131*, 2623–2627.

Fallon, U.B., Ben-Shlomo, Y., Elwood, P., Ubbink, J.B., & Davey Smith, G. (2001). *Homocysteine and coronary heart disease: Author's reply*. Retrieved from http://heart.bmjjournals.com/cgi/eletters/85/2/153

Fisher, R.A. (1952). Statistical methods in genetics. *Heredity*, *6*, 1–2.

Freathy, R.M., Timpson, N.J., Lawlor, D.A., Pouta, A., Ben-Shlomo, Y., Ruokonen, A., et al. (2008). Common variation in the FTO gene alters diabetes-related metabolic traits to extent expected, given its effect on BMI. *Diabetes*, *57*, 1419–1426.

Garcia-Closas, M., Malats, N., Silverman, D., Dosemeci, M., Kogevinas, M., Hein, D.W., et al. (2005). NAT2 slow acetylation, GSTM1null genotype, and risk of bladder cancer: Results from the Spanish Bladder Cancer Study and meta-analyses. *Lancet*, *366*, 649–659.

Gemma, S., Vichi, S., & Testai, E. (2007). Metabolic and genetic factors contributing to alcohol induced effects and fetal alcohol syndrome. *Neuroscience & Biobehavioral Reviews*, *31*, 221–229.

Gibson, G., & Wagner, G. (2000). Canalization in evolutionary genetics: A stabilizing theory? *BioEssays*, *22*, 372–380.

Glebart, W.M. (1998). Databases in genomic research. *Science*, *282*, 659–661.

Glynn, R.K. (2006). Commentary: Genes as instruments for evaluation of markers and causes. *International Journal of Epidemiology*, *35*, 932–934.

Haldane, J.B.S. (1938). *Heredity and politics*. London: George Allen and Unwin.

Hart, C., Davey Smith, G., Hole, D., & Hawthorne, V. (1999). Alcohol consumption and mortality from all causes, coronary heart disease, and stroke: Results from a prospective cohort study of Scottish men with 21 years of follow up. *British Medical Journal*, *318*, 1725–1729.

Hartman, J.L., Garvik, B., & Hartwell, L. (2001). Principles for the buffering of genetic variation. *Science*, *291*, 1001–1004.

Hashibe, M., McKay, J.D., Curado, M.P., Oliveira, J.C., Koifman, S., Koifman, R., et al. (2008). Multiple *ADH* genes are associated with upper aerodigestive cancers. *Nature Genetics*, *40*, 707–709.

Hasin, D., Aharonovich, E., Liu, X., Mamman, Z., Matseoane, K., Carr, L., et al. (2002). Alcohol and ADH2 in Israel: Ashkenazis, Sephardics, and recent Russian immigrants. *American Journal of Psychiatry*, *159*, 1432–1434.

Haskell, W.L., Camargo, C., Williams, P.T., Vranizan, K.M., Krauss, R.M., Lindgren, F.T., et al. (1984). The effect of cessation

Downloaded from pps.sagepub.com by guest on October 27, 2010

and resumption of moderate alcohol intake on serum high-density-lipoprotein subfractions. *New England Journal of Medicine*, *310*, 805–810.

Higuchi, S., Matsuushitam, S., Imazeki, H., Kinoshita, T., Takagi, S., & Kono, H. (1994). Aldehyde dehydrogenase genotypes in Japanese alcoholics. *Lancet*, *343*, 741–742.

Holtzman, N.A. (2001). Putting the search for genes in perspective. *International Journal of Health Services*, *31*, 445.

Honkanen, R., Pulkkinen, P., Järvinen, R., Kröger, H., Lindstedt, K., Tuppurainen, M., et al. (1996). Does lactose intolerance predispose to low bone density? A population-based study of perimenopausal Finnish women. *Bone*, *19*, 23–28.

Hornstein, E., & Shomron, N. (2006). Canalization of development by microRNAs. *Nature Genetics*, *38*, S20–S24.

Irons, D.E., McGue, M., Iacono, W.G., & Oetting, W.S. (2007). Mendelian randomization: A novel test of the gateway hypothesis and models of gene–environment interplay. *Development and Psychopathology*, *19*, 1181–1195.

Jacobson, S.W., Carr, L.G., Croxford, J., Sokol, R.J., Li, T.K., & Jacobson, J.L. (2006). Protective effects of the alcohol dehydrogenase-ADH1B allele in children exposed to alcohol during pregnancy. *Journal of Pediatrics*, *148*, 30–37.

Jousilahti, P., & Salomaa, V. (2004). Fibrinogen, social position, and Mendelian randomisation. *Journal of Epidemiology and Community Health*, *58*, 883.

Kandel, D., & Yamaguchi, K. (1993). From beer to crack: Developmental patterns of drug involvement. *American Journal of Public Health*, *83*, 851–855.

Kandel, D.B., Yamaguchi, K., & Chen, K. (1992). Stages of progression in drug involvement from adolescence to adulthood: Further evidence for the gateway theory. *Journal of Studies on Alcohol*, *53*, 447–457.

Kaplan, J.M., & Pigliucci, M. (2001). Genes "for" phenotypes: A modern history view. *Biology and Philosophy*, *16*, 189–213.

Keavney, B. (2002). Genetic epidemiological studies of coronary heart disease. *International Journal of Epidemiology*, *31*, 730–736.

Kitami, T., & Nadeau, J.H. (2002). Biochemical networking contributes more to genetic buffering in human and mouse metabolic pathways than does gene duplication. *Nature Genetics*, *32*, 191–194.

Klatsky, A.L. (2001). Commentary: Could abstinence from alcohol be hazardous to your health? *International Journal of Epidemiology*, *30*, 739–742.

Langlois, M.R., Delanghe, J.R., De Buyzere, M.L., Bernard, D.R., & Ouyang, J. (1997). Effect of haptoglobin on the metabolism of vitamin C. *American Journal of Clinical Nutrition*, *66*, 606–610.

Lawlor, D.A., Harbord, R.M., Timpson, N.J., Lowe, G.D.O., Rumley, A., Gaunt, T.R., et al. (2008). The association of C-reactive protein and CRP genotype with coronary heart disease: Findings from five studies with 4,610 cases amongst 18,637 participants. *PLoS One*, *8*, e3011.

Lewis, S., & Davey Smith, G. (2005). Alcohol, ALDH2 and esophageal cancer: A meta-analysis which illustrates the potentials and limitations of a Mendelian randomization approach. *Cancer Epidemiology Biomarkers and Prevention*, *14*, 1967–1971.

Lower, G.M., Nilsson, T., Nelson, C.E., Wolf, H., Gamsky, T.E., & Bryan, G.T. (1979). N-acetylransferase phenotype and risk in urinary bladder cancer: approaches in molecular epidemiology. *Environmental Health Perspectives*, *29*, 71–79.

Marmot, M. (2001). Reflections on alcohol and coronary heart disease. *International Journal of Epidemiology*, *30*, 729–734.

Memik, F. (2003). Alcohol and esophageal cancer, is there an exaggerated accusation? *Hepatogastroenterology*, *54*, 1953–1955.

Morgan, T.H. (1935). The relation of genetics to physiology and medicine. *Scientific Monthly*, *41*, 5–18.

Newcomer, A.D., Hodgson, S.F., Douglas, M.D., & Thomas, P.J. (1978). Lactase deficiency: Prevalence in osteoporosis. *Annals of Internal Medicine*, *89*, 218–220.

Osier, M.V., Pakstis, A.J., Soodyall, H., Comas, D., Goldman, D., Odunsi, A., et al. (2002). A global perspective on genetic variation at the ADH genes reveals unusual patterns of linkage disequilibrium and diversity. *American Journal of Human Genetics*, *71*, 84–99.

Ottman, R. (1990). An epidemiologic approach to gene–environment interaction. *Genetic Epidemiology*, *7*, 177–185.

Ottman, R. (1996). Theoretical epidemiology: Gene–environmental interaction—Definitions and study designs. *Preventive Medicine*, *25*, 764–770.

Ottman, R. (2006). Study design and analysis for assessment of interactions. In L. M. Hernandez and D. G. Blazer (Eds.), *Genes, behavior, and the social environment: Moving beyond the nature/nurture debate* (pp. 161–180). Washington, DC: National Academies Press.

Palmer, L., & Cardon, L. (2005). Shaking the tree: Mapping complex disease genes with linkage disequilibrium. *Lancet*, *366*, 1223–1234.

Rimm, E. (2001). Commentary: Alcohol and coronary heart disease: Laying the foundation for future work. *International Journal of Epidemiology*, *30*, 738–739.

Risch, N., Herrell, R., Lehner, T., Liang, K.-Y., Eaves, L., Hoh, J., et al. (2009). Interaction between the serotonin transporter gene (*5-HTTLPR*), stressful life events, and risk of depression. *Journal of the American Medical Association*, *301*, 2462–2471.

Rutherford, S.L. (2000). From genotype to phenotype: Buffering mechanisms and the storage of genetic information. *BioEssays*, *22*, 1095–1105.

Seitz, H.K., & Stickel, F. (2007). Molecular mechanisms of alcohol-mediated carcinogenesis. *Nature Reviews Cancer*, *7*, 599–612.

Shaper, A.G. (1993). Editorial: Alcohol, the heart, and health. *American Journal of Public Health*, *83*, 799–801.

Steer, C.D., Davey Smith, G., Emmett, P.M., Hibbeln, J.R., & Golding, J. (in press). FADS2 polymorphisms modify the effect of breastfeeding on child IQ. *PLoS 1*.

Strohman, R.C. (1993). Ancient genomes, wise bodies, unhealthy people: The limits of a genetic paradigm in biology and medicine. *Perspectives in Biology and Medicine*, *37*, 112–145.

Tabery, J.G. (2000). Causation in the nature–nurture debate: The case of genotype–environment interaction. Retrieved from http://etd.library.pitt.edu/ETD/available/etd-07232007-121636/unrestricted/Tabery-Causation%5B1%5D.in.the.Nature-Nurture.Debate.pdf

Downloaded from pps.sagepub.com by guest on October 27, 2010

Tabery, J. (2007). Biometric and developmental gene–environment interactions: Looking back, moving forward. *Development and Psychopathology*, *19*, 961–976.

Takagi, S., Iwai, N., Yamauchi, R., Kojima, S., Yasuno, S., Baba, T., et al. (2002). Aldehyde dehydrogenase 2 gene is a risk factor for myocardial infarction in Japanese men. *Hypertension Research*, *25*, 677–681.

Terwilliger, J.D., & Weiss, W.M. (2003). Confounding, ascertainment bias, and the blind quest for a genetic "fountain of youth." *Annals of Medicine*, *35*, 532–544.

Thomas, D.C., & Conti, D.V. (2004). Commentary on the concept of "Mendelian randomization." *International Journal of Epidemiology*, *33*, 17–21.

Thompson, W.D. (1991). Effect modification and the limits of biological inference from epidemiological data. *Journal of Clinical Epidemiology*, *44*, 221–232.

Timpson, N.J., Forouhi, N.G., Brion, M.J., Harbord, R.M., Cook, D.G., Johnson, P., et al. (2010). Genetic variation at the *SCL23A1* locus is associated with circulating concentrations of L-ascorbic acid (Vitamin C): Evidence from 5 independent studies with >15,000 participants. *American Journal of Clinical Nutrition*, *92*, 375–382.

Timpson, N.J., Lawlor, D.A., Harbord, R.M., Gaunt, T.R., Day, I.N.M., Palmer, L.J., et al. (2005). C-reactive protein and its role in metabolic syndrome: Mendelian randomisation study. *Lancet*, *366*, 1954–1959.

Timpson, N.J., Sayers, A., Davey Smith, G., & Tobias, J.H. (2009). How does body fat influence bone mass in childhood? A Mendelian randomisation approach. *Journal of Bone and Mineral Research*, *24*, 522–533.

Tybjaerg-Hansen, A., Steffensen, R., Meinertz, H., Schnohr, P., & Nordestgaard, B.G. (1998). Association of mutations in the apolipoprotein B gene with hypercholesterolemia and the risk of ischemic heart disease. *New England Journal of Medicine*, *338*, 1577–1584.

Waddington, C.H. (1942). Canalization of development and the inheritance of acquired characteristics. *Nature*, *150*, 563–565.

Warren, K.R., & Li, T.K. (2005). Genetic polymorphisms: Impact on the risk of fetal alcohol spectrum disorders. *Birth Defects Research Part A: Clinical Molecular Teratology*, *73*, 195–203.

Weiss, K., & Terwilliger, J. (2000). How many diseases does it take to map a gene with SNPs? *Nature Genetics*, *26*, 151–157.

Wellcome Trust Case Control Consortium. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*, *447*, 661–678.

Wheatley, K., & Gray, R. (2004). Commentary: Mendelian randomization—An update on its use to evaluate allogeneic stem cell transplantation in leukaemia. *International Journal of Epidemiology*, *33*, 15–17.

Wilkins, A.S. (1997). Canalization: A molecular genetic perspective. *BioEssays*, *19*, 257–262.

Wolf, U. (1995). The genetic contribution to the phenotype. *Human Genetics*, *95*, 127–148.

Youngman, L.D., Keavney, B.D., Palmer, A., Parish, S., Clark, S., Danesh, J., et al. (2000). Plasma fibrinogen and fibrinogen genotypes in 4685 cases of myocardial infarction and in 6002 controls: test of causality by 'Mendelian randomization'. *Circulation*, *102*(Suppl. 2), 31–32.

Downloaded from pps.sagepub.com by guest on October 27, 2010