Mendelian randomization for strengthening causal inference in observational studies: application to gene by environment interaction.

Forthcoming, <u>Perspectives on Psychological Science</u>

George Davey Smith [1]

[1] MRC Centre for Causal Analyses in Translational Epidemiology

Department of Social Medicine

University of Bristol

Oakfield House

Oakfield Grove

Bristol BS8 2BN

Introduction

In 1875 George Darwin, the second son and fifth child of Charles Darwin, reviewed evidence on the putative detrimental effects on offspring health of cousin marriages, something of personal interest to him as he was the product of such a union (Darwin 1875). He concluded by reviewing the most comprehensive studies of the issue and described what may be the first presentation of gene-by-environment interaction informed by at least some understanding of heredity. In 1864, George Darwin tells us, "Dr Mitchell had come to the conclusion that under favourable 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 This is in striking accordance with some unpublished experiments of my father, Mr.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 The unpublished findings of Charles Darwin were later published in his 1877 book "The Effects of Cross and Self Fertilisation in the Vegetable Kingdom" (Darwin 1876).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 sub-optimal environmental circumstances. There are clearly echoes here of the celebrated contemporary gene by environment interactions, such as those between genetic variation in the monoamine oxidase A gene, childhood maltreatment, and the outcome antisocial behaviour in later life (Caspi et al 2002). Unlike, like recent examples of gene by environment interaction in the molecular genetic age (Risch et al 2009), which have failed to stand up to rigorous attempts at replication. (Risch et al 2009) The interesting claims made by Dr Mitchell have not been formally followed up in this way.   In this review I will suggest that gene by 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 focusing on other aspects of the interactions themselves.   To illustrate this I will utilise examples from the alcohol and health area, one of the many contested fields where disparate claims based on observational data have been made. I will briefly outline, and carrying out randomized controlled trials would be difficult if not impossible, the use of genetic main effects in the basic Mendelian randomization approach for strengthening causal inference. I will then discuss how gene by environment interactions can also be utilised in this regard, will discuss the typology of gene

by environment interaction as in some of the framework I advance. And conclude correlate briefly outlining the limitations of this approach.

Mendelian randomization: What is it and how does it work?

The basic principle 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 and Ebrahim 2003; Davey Smith and Ebrahim 2004; Davey Smith and Ebrahim 2005; Davey Smith 2006a; Lawlor et al 2008; Ebrahim and Davey Smith 2008; Davey Smith 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 behavioural, 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 and 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, or of 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. 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, when the increasing benefits seen over the relatively short trial period are projected to the expectation for differences over a lifetime (Davey Smith and Ebrahim 2004).

In the Mendelian randomization framework the associations of genotype with outcomes are of interest because of the strengthened inference they allow about the action of the environmental modifiable risk factors that the genotypes proxy for, rather than what they say about genetic mechanisms *per se*. Mendelian randomization studies are aimed at informing strategies to reduce disease risk through influencing the non-genetic 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 Fisher said: "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 conventional observational epidemiological studies comes from

4

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 behavioural, socio-economic 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 SNPs of known phenotypic effects and nearly 100 socio-culturual, behavioural and biological risk factors for disease (Davey Smith et al 2008b).

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 et al 1967; Newcomer et al 1978; Lower et al 1979; Honkanen et al 1996), with the term Mendelian randomization being introduced by Richard Gray and Keith Wheatley in 1991 (Wheatley and 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 leukaemia. More recently the term has been widely used in discussions of observational epidemiological studies (Davey Smith and Ebrahim 2003; Youngman et al 2000; Fallon et al 2001; Clayton and McKeigue 2001; Keavney 2002)_. 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 and Ebrahim 2003; Davey Smith and Ebrahim 2004; Davey Smith 2006b; Ebrahim and Davey Smith 2008; Davey Smith et al 2008a) ; 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 (coronary heart disease, depression and dementia) for example, the possible protective effect of moderate alcohol consumption on coronary heart disease CHD risk remains highly controversial (Marmot 2001; Bovet and Paccaud 2001; Klatsky 2001). Non-drinkers may be at a higher risk of CHD because health problems (perhaps induced by previous alcohol abuse) dissuade them from drinking (Shaper 2001). As well as this form of reverse causation, confounding could play a role, with non-drinkers being more likely to display an adverse profile of socioeconomic or other behavioural risk factors for CHD (Hart et al 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 an RCT of differential levels of alcohol intake, able to test 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 oxidized by aldehyde dehydrogenases (ALDHs) to acetate. Half of Japanese people are heterozygotes or homozygotes for a null variant of *ALDH2* and peak blood acetaldehyde concentrations post alcohol challenge are 18 times and 5 times higher respectively among homozygous null variant and heterozygous individuals compared with homozygous wild type individuals (Enomoto et al 1991). This renders the consumption of alcohol unpleasant through inducing facial flushing, palpitations, drowsiness and other symptoms. As Figure 1 shows, there are very considerable differences in alcohol consumption according to genotype among men (Takagi et al 2002). The principles of Mendelian randomization are seen to apply – two factors that would be expected to be associated with alcohol consumption, age and cigarette smoking, which would 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 *ALDH2* genotype influences diseases known to be related to alcohol consumption and as proof of principle 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, including data from short-term randomized controlled trials, suggests that alcohol increases HDL cholesterol levels (Haskell et al 1984; Burr et al 1986) (which should protect against CHD). In line with this, *ALDH2* genotype is strongly associated with HDL cholesterol in the expected direction. (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 coronary heart disease discussed earlier, we could be looking at a confounded 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), with headlines like "Moderate drinking may help men with high blood pressure".

Evidence on the association of alcohol drinking and alcohol consumption blood pressure can come from studies of A*LDH2* genotype and blood pressure. A meta-analysis of such studies suggest there is indeed a substantial positive effect of alcohol on blood pressure (Chen et al 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 compared with others the shape of the association remains similar. Among women, however, who drink very little compared to men there is no evidence of association between drinking and genotype.

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

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 associated 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 behaviours, 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 non-alcohol substance use (Irons et al 2007; Kandel and Yamaguchi 1993; Kandel et al 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 who *ALDH2* status was associated with alcohol use and alcohol use was associated with tobacco, marijuana, and other illegal drug use. *ALDH2* variation was not robustly associated with non-alcohol substance use, however, which was taken to provide 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 for very high levels of intake, in the form of fetal alcohol syndrome (Gemma et al 2007). However, the influence outside of this extreme situation is less easy to assess, particularly as higher levels of alcohol intake will be related to a wide array of potential socio-cultural, behavioural 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, 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 outcomes.

Studies have generally utilized a variant in the alcohol dehydrogenase gene (*ADH1B*3* allele). Alcohol dehydrogenase metabolises 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, faster metabolism will clear blood alcohol more rapidly, so less high levels will be reached and these will more quickly return to low levels. 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 among these mothers the alcohol-metabolizing genotypes 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 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 and 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, rather than 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 analysed in a variety of ways. First, the mothers of offspring with a particular outcome can be compared to 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, maternal genotype, rather than offspring genotype, will be the appropriate exposure indicator. Clearly, maternal and offspring genotype are associated, but conditional on each 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, whether fathers or women who have offspring without the health outcome, is generally preferable.

Other examples of Mendelian randomization: a brief catalogue

Mendelian randomization has now been utilised 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 principle of this approach comes from situations where the answer is known. For example, several genetic variants that are associated with blood cholesterol levels are also associated with coronary heart disease risk, in line with the substantial amount of evidence, including that from RCTs, that higher blood cholesterol levels causally increase disease risk (Davey Smith et al 2008a).

These studies demonstrate another strength of the Mendelian randomization approach, in that the observed association of genotype with coronary heart disease is larger than that predicted from its effect on cholesterol levels and the magnitude of association of cholesterol levels with coronary heart disease risk identified in RCTs. Since RCTs only lower cholesterol for a few years, and atherosclerosis is a life-long process, this is to be expected, as the genetic variants indicate differences in cholesterol levels over many decades, as opposed to the relatively short-term changes produced in RCTs. Genotypic differences in intermediate phenotypes can provide evidence of lifelong, 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) is found in observational studies to be strongly predictive of type 2 diabetes and coronary heart disease 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 2009; 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 randomized controlled trials

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 analogy with RCTs is also useful with respect to one objection that has been raised in conjunction with Mendelian randomization studies. This is 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 and 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; neither 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, where 50% of participants receive the medication and 50% received a placebo. If the antihypertensive therapy reduced blood pressure by a quarter of a standard deviation (i.e. a 5mmg reduction in systolic blood pressure given systolic blood pressure has a standard deviation of 20mmHg in the population) then within the whole study group, treatment assignment (i.e. antihypertensive use versus placebo) will explain $5/20^2 = 1.25\%$ of the variance. In the example of *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. Both logic and quantification fail to support criticisms of the Mendelian randomization approach based on either the obvious fact that many factors influence most phenotypes of interest or that particular genetic variants only account for a small proportion of variance in the phenotype.

Mendelian randomization and instrumental variable approaches

As well as the analogy with RCTs, 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, where 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 et al 2005b). The parallels between Mendelian randomization and instrumental variable approaches are discussed in more detail elsewhere (Thomas and Conti 2004; Lawlor et al 2008). 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, oesophageal and head and neck cancer: Gene by environment interaction, cause and mechanismGene by environment interaction of a different kind to 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 oesophageal 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 oesophageal cancer risk (Lewis and Davey Smith 2005)found that people who are homozygous for the null variant, who therefore consume considerably less alcohol, have a greatly reduced risk of oesophageal 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 oesophageal cancer risk (Burd 2006). A similar picture is seen when head and neck cancer is the outcome (Boccia et al 2009).

Thus, with respect to the homozygous null variant versus 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, through allowing comparison of a group of low drinkers to a group who drink considerable amounts of alcohol, with no confounding factors differing between these groups. With respect to both oesophageal 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 and Stickel 2007). Thus, drinkers those 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 do drink. When the heterozygotes are compared with wild type homozygotes, an interesting picture emerges – the risk of oesophageal cancer is *higher* in the heterozygotes, although 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 oesophageal cancer risk between *ALDH2* heterozygotes and those homozygous for the wild type are displayed, stratified by drinking status. In non-drinkers there is no robust evidence of any association between genotype and oesophageal cancer outcomes, as would be expected if the underlying environmentally modifiable causal factor were alcohol intake and the mechanism was through acetaldehyde levels. In further support of the hypothesis, amongst people who were drinking alcohol there was increased risk amongst 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 never drinkers and a graded association according to alcohol intake level among alcohol drinkers (Boccia et al 2009).

Gene by environment interactions interpreted within a Mendelian randomization framework

The meaning of gene by environment interactions has a contested history within human genetics. As James Tabery (Tabery 2000; Tabery 2007) has discussed, two distinct concepts can be identified. First, there is a developmental concept, pioneered by Lancelot Hogben, which considered 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 by 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). What is noticeable from considering Haldane's typology is that many apparent gene by environment interactions discussed in the molecular genetics era will fall into his first category, where there is not a clear cross-over of effects of genotype

according to environment, rather there is some apparent quantitative difference, with a genotype having a larger influence on phenotype in one environment than another. Haldane considered interactions to be seen when one genotype was associated with a beneficial effect in one environment and 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 quantitative interactions are scale dependent (Thompson 1991) - there must be an interaction on one scale (e.g. additive) if there is no interaction on another scale (e.g. multiplicative), if any effect of genotype exists.

In an important series of papers Ruth Ottman has explicated a typology of five models of gene-environment interactions. (Ottman 1990; Ottman 1996; Ottman 2006). Here I consider how these models would be interpreted within a Mendelian randomization framework. In model A (Figure 9) the genotype increases the level of expression of the risk factor, which in turn influences the risk of disease. Under some definitions this would not be interaction, rather 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 randomisation similar to that discussed earlier with respect to maternal alcohol metabolising genotypes. The genotype has no effect if it is decoupled from the intermediate risk factor – for example, in a society where few people drink alcohol (or among women in societies where 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. Intermediate phenotype Mendelian randomisation studies – e.g. genetic variants influencing cholesterol levels and through this coronary heart disease – are also examples of model A.

In Model B (figure 9) the risk factor influences disease risk and the genotype modifies this, but on its own the genotype will not influence outcomes. In the absence of alcohol drinking the variant will not be related to alcohol-related morbidity, but in the presence of drinking the variant will modify the severity of outcome, 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 oesophageal cancer in the absence of alcohol consumption, whereas alcohol consumption does increase risk of oesophageal cancer risk even in the absence of *ALDH2* wild type, although to a lesser degree.

Another example of Model B (Figure 9) relates to the influence of smoking tobacco on bladder cancer risk. Observational studies suggest an association, but clearly confounding

and a variety of biases could generate such an association. The potential carcinogens in tobacco smoke of relevance 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 slow acetylators, 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) (Gu et al 2005). The influence of the *NAT2* slow acetylation genotype is only appreciable among those also exposed to heavy smoking. Since 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 et al 2005).

In Model C (figure 9) the genotype has a direct effect on disease risk while the risk factor does not have this effect when acting by itself. Examples here come from the field of pharmacogenetics, where an otherwise benign exposure has a detrimental influence if accompanied by a particular genotype which 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. Use of barbiturates in generally benign, but in the presence of porphyria genotype leads to very severe attacks of skin blistering. Model D (figure 9) is similar to C, but in the latter case both modifiable and genetic risk factor do not produce outcomes alone, only 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 (figure 9) refers to the situation where 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, 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 variation relate to disease outcomes can provide evidence of the causal action of vitamin D levels in this situation, since 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 and Ebrahim 2003; Ebrahim and Davey Smith 2008) and are therefore only some issues of particular relevance for to gene by environment interaction are briefly considered here.

Confounding of genotype — environmentally-modifiable risk factor —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 when interpreting the results, this

possibility needs to be considered. Firstly, 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 producing confounding). 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 that 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, where 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 where 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 estimated. Empirical studies to date suggest that common genetic variants are largely unrelated to the behavioural 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, 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 genotype-blood pressure association depends upon 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, which are not in linkage disequilibrium with each other, but which 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, since 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. Investigation of the effect of alcohol on risk of head and neck cancer though comparing risk among *ALDH2* homozygous wild type and *ALDH2* homozygous null variant men was discussed above; the same issue has been addressed by studying the interaction between alcohol intake, *ADH* variation and head and neck cancer risk (Figure 10) (Hashibe et al 2008) , where the influence of genotype among drinkers, but not among non-drinkers, 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 et al 2009).

Special issues with confounding in studies of gene by environment interactions

It must be recognized that gene by 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 *NAT2* / smoking / bladder cancer example 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 produce the qualitative interaction of essentially no effect of the genotype in one social class stratum and an 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 versus 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 post-natal development, and thus influencing 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 (Wilkins 1997; Rutherford 2000; Gibson and Wagner 2000; Hartman et al 2001; Debat and David 2001; Kitami and Nadeau 2002; Hornstein and Shomron 2006). Such buffering can be achieved either through genetic redundancy (more than one gene having the same or similar function) or through alternative metabolic routes, where 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 post-natal growth may influence the expression of a wide range of other genes, leading to changes that may compensate for the influence of 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. there is 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 where a variant influences an adulthood environmental exposure — e.g. *ALDH2* variation and alcohol intake — developmental compensation to genotype will not be an issue. In many cases of gene by environment interaction interpreted with respect to causality of the environmental factor, the same applies, since 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 in many cases, even if a locus is involved in a disease-related metabolic process, there may be no suitable marker or functional polymorphism to allow study of this process (Weiss and Terwilliger 2000). In an earlier paper on Mendelian randomization (Davey Smith and Ebrahim 2003) we discussed the example of vitamin C, since 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 coronary heart disease 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 et al 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 et al 2001). We therefore used this as an example of a situation where suitable polymorphisms for studying the modifiable risk factor of interest could not be located. However, since the earlier paper was written,

functional variation in *SLC23A1* has been identified that is related to circulating vitamin C levels (Timpson *et al.* personal communication). 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: Mendelian randomization, what it is and what it isn't

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 over-selling of "discoveries" in genetics that reiterate obvious truths so clearly (albeit somewhat repetitively) that there is no need to repeat them here (e.g. Berkowitz 1996; Baird 2000; 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 and 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 within environmental limits will pertain at a group level (Wolf 1995). This is analogous to an RCT of antihypertensive agents, where at a collective level the group randomized to active medication will have lower mean blood pressure than the group randomized to placebo, but at an individual level many participants randomized to active treatment will have higher blood pressure than many individuals randomized to placebo. It is group level differences are what create the analogy between Mendelian randomization and RCTs, outlined in Figure 13.

Finally, the associations that Mendelian randomization depend upon do need to pertain to a definable group at a particular time, but do not need to be immutable. Thus, *ALDH2* variation will not be related to alcohol consumption in a society where alcohol is not consumed; the association will vary by gender, by cultural group and may change over time (Higuchi et al 1994; Hasin et al 2002). Within the setting of a study of a well-defined group, however, the genotype will be associated with group-level differences in alcohol 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 in any case the influence of genetic factors is not reversible. 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' 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 and Weiss 2003). 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 Apo 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 7-fold increased risk of CHD, will only generate a population attributable risk of 0.5% (Tybjaerg-Hansin et al 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 this is associated with a relative risk of 2-fold, 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 since genotype cannot be modified, genotype–disease associations are not of public health importance (Terwilliger and 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". 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 the then popular genetic approach to disease, eugenics. 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.

Table 1: Association of NAT2 Slow Acetylation Genotype with Bladder Cancer in Never and Ever Smokers and Overall. Odds Ratio (95% confidence intervals)(Garcia-Closas et al 2005)

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

P for interaction on multiplicative scale <0.01

Table 2: Association of smoking status and NAT2 Slow Acetylation Genotype with Bladder Cancer

	NAT2 Rapid	NAT2 Slow
Never Smoker	1.0	0.9
Occasional	1.2	1.6
Former	2.4	4.1
Current	5.2	7.5

Garcia-Closas et al. Lancet 2005; 366: 649-659

Figure 1.   Relationship between alcohol intake and ALDH2 genotype

Data from Takagi, et al. 2002 (Takagi et al 2002)
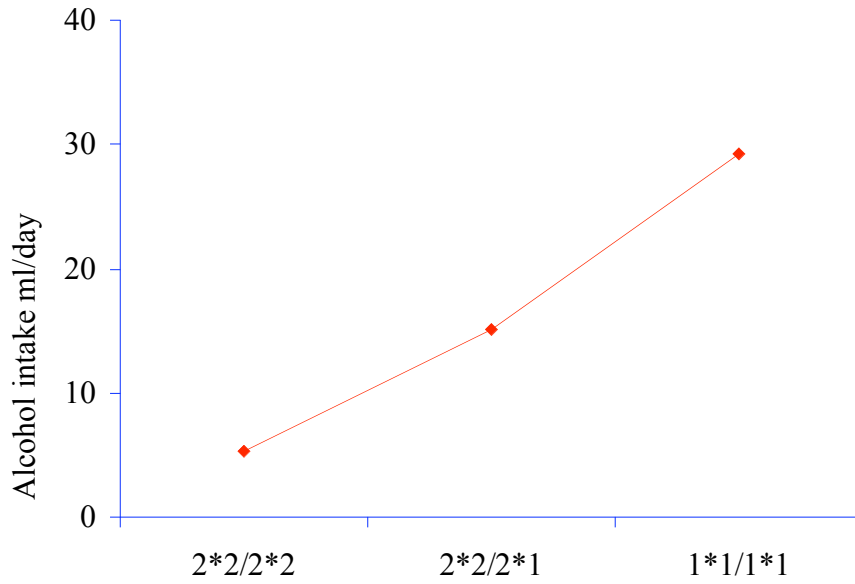
Figure 2. ALDH2 genotype by alcohol consumption, g/day: 5 studies, n=6815 (Chen et al 2008)
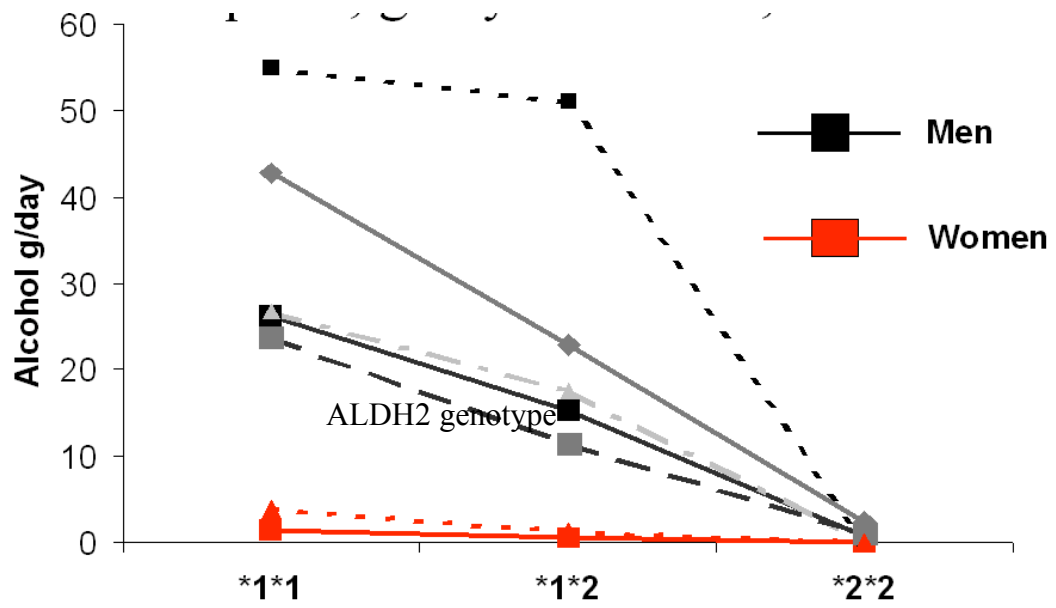
Figure 3. Forest plot of studies of *ALDH2* genotype and hypertension (Chen et al 2008)
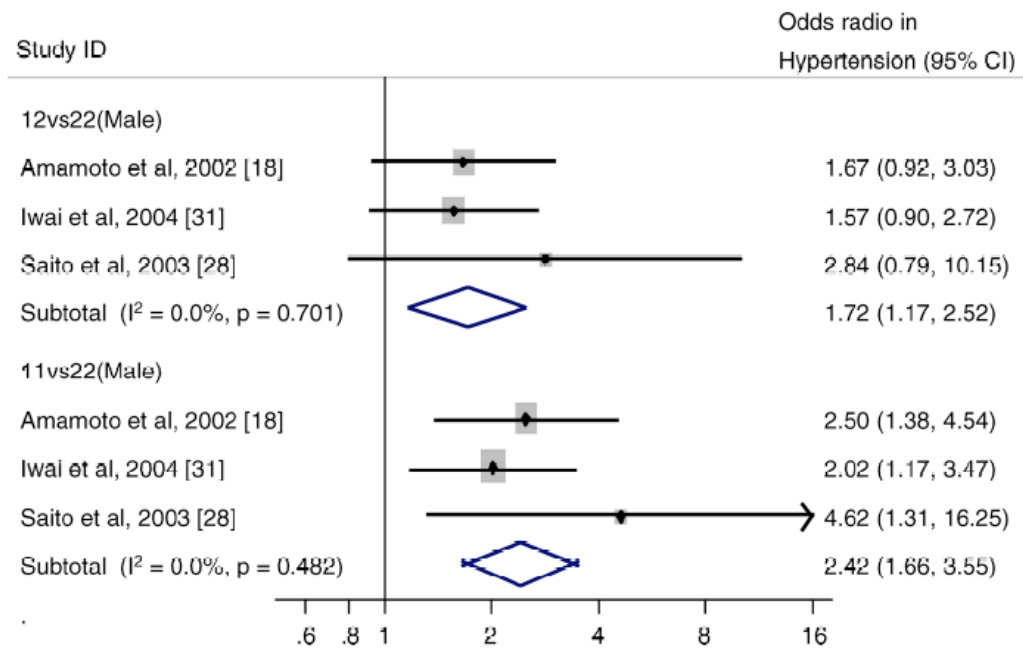
Figure 4. ALDH2 genotype and systolic blood pressure (Chen et al 2008)
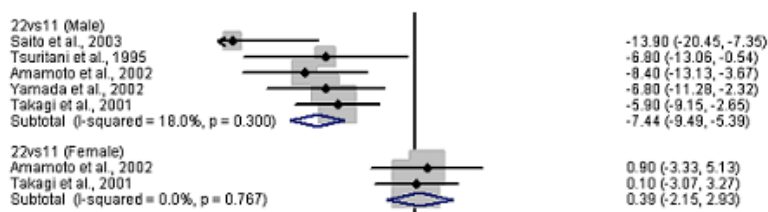
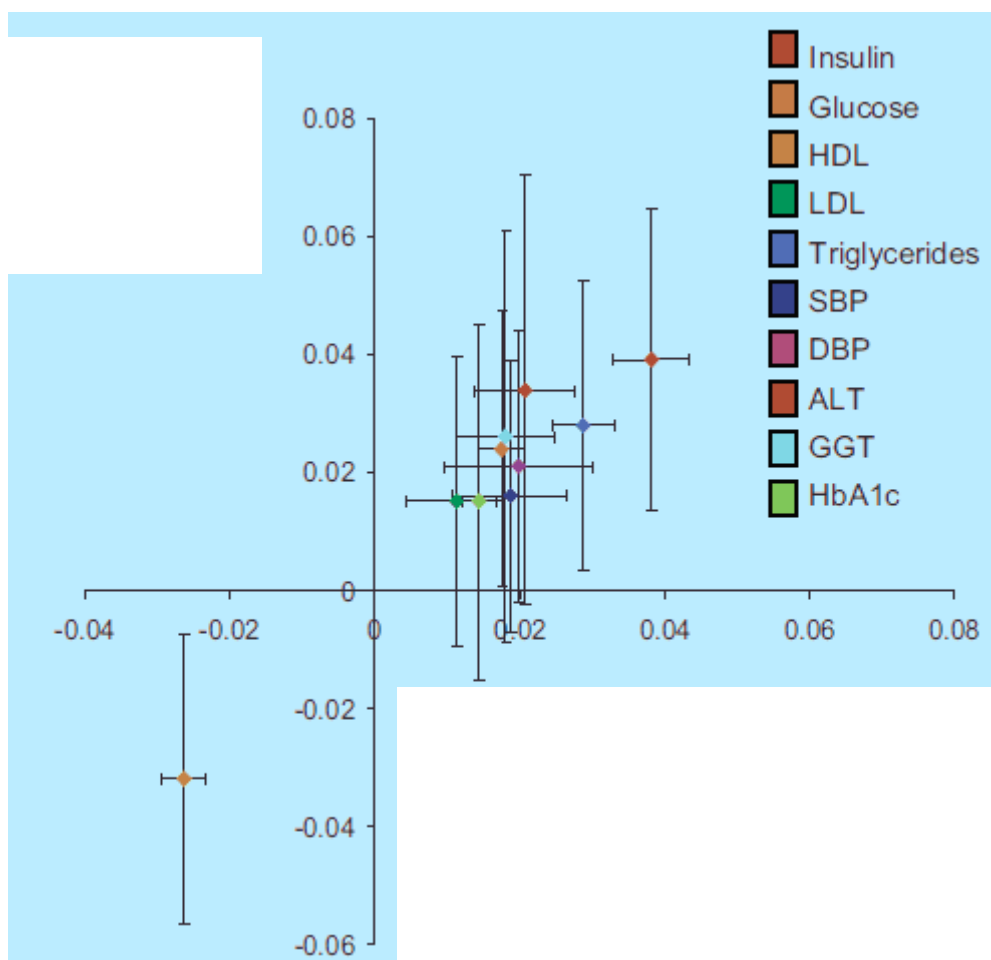**Figure 5 The observed effects of FTO variation on metabolic traits are exactly as expected**

Figure 6 Mendelian randomization and randomized controlled trial designs compared
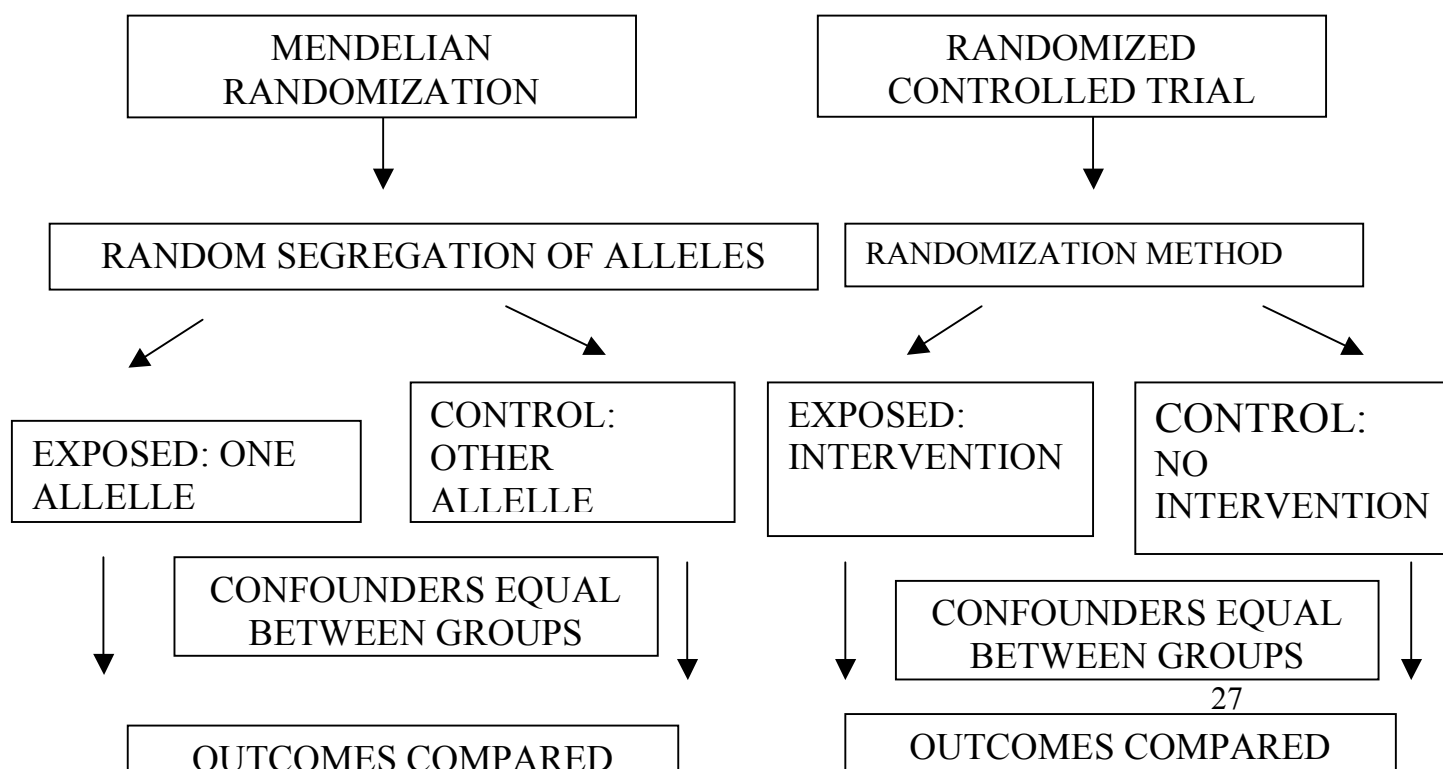
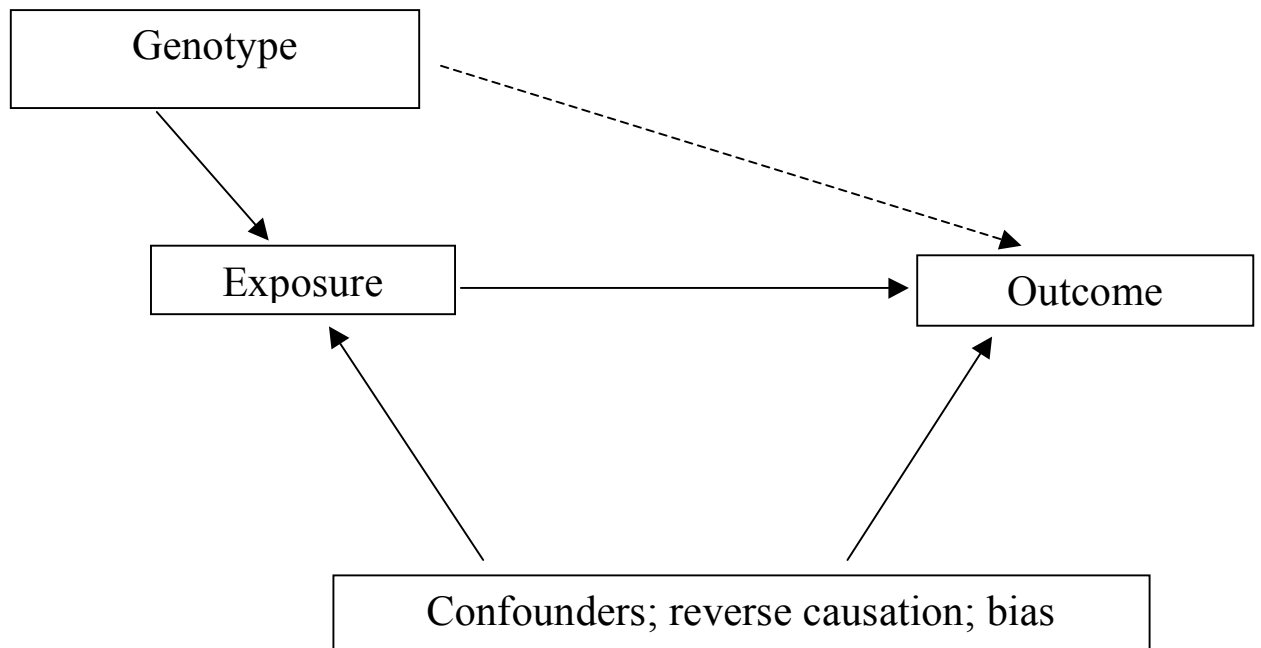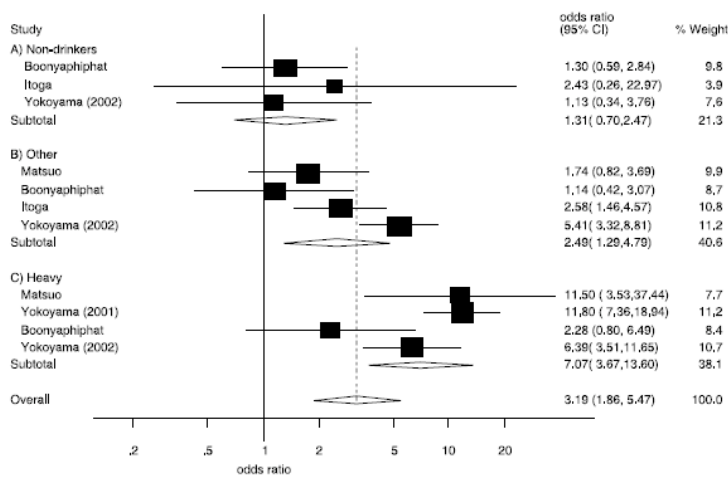**Figure 7. Mendelian randomization as an instrumental variables approach**


Figure 8

Non-drinkers = non-drinkers (Boonyaphiphat et al and Itoga et al), and never and ex-drinkers (Yokoyama et al, 2002).
Others = All those who were not heavy drinkers (Matsuo et al, 2001), drinkers ≤60g/day (Boonyaphiphat et al), habitual drinkers (Itoga et al, 2002) and 1-17.9 units/week; where 1 unit=22g ethanol (Yokoyama et al, 2002).
Heavy drinkers = 75ml ethanol per day ≥5 days/week (Matsuo et al, 2001), alcoholics (Yokoyama et al, 2001), >60g ethanol per day (Boonyaphiphat et al) and ≥18 units/week; where 1 unit=22g ethanol (Yokoyama et al, 2002),

Figure 3. Risk of esophageal cancer in individuals with the *ALDH2*1*2* versus *ALDH2*1*1* genotype.
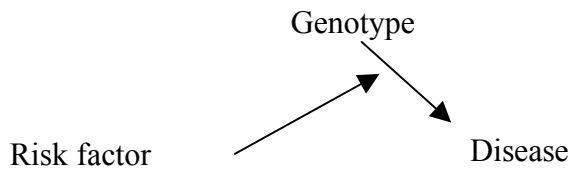
**Figure 9 Model A - The genotype increases expression of the risk factor**

Genotype

Risk factor $\longrightarrow$ Disease
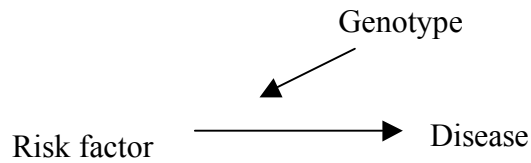
Model B: The genotype exacerbates the effect of the risk factor

Genotype

Risk factor $\longrightarrow$ Disease
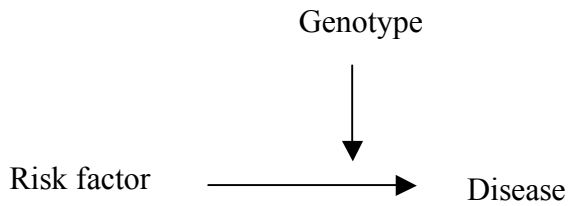
Model C: The risk factor exacerbates the effect of the genotype

Genotype

Risk factor $\longrightarrow$ Disease

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

Genotype

Risk factor $\longrightarrow$ Disease
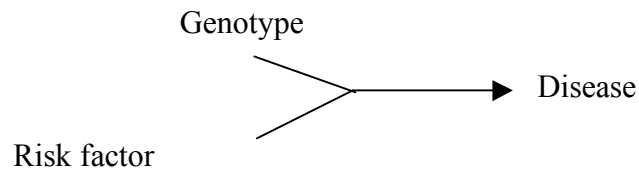
Model E: The genotype and risk factor each affect risk: combined effects can be additive or nonadditive

Genotype

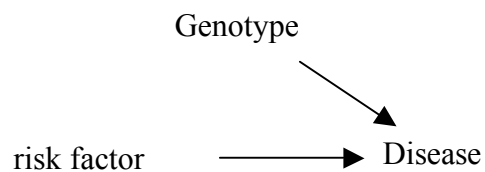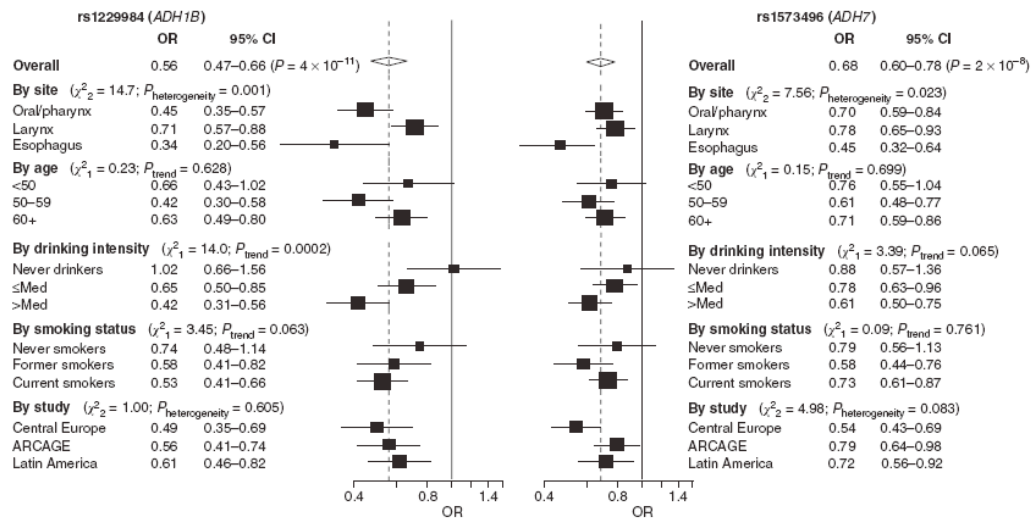risk factor $\longrightarrow$ Disease

Figure 10 (Hashibe et al 2008)

Box: JBS Haldane on Gene by Environment Interaction

In his polemical book *Heredity and Politics* Haldane presented the table below as exhausting the possibilities of gene-by-environment interaction. In the first situation genotype A is superior to genotype B in each environment, and environment X is more favourable 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 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 is interactions within this conceptual space that 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, with both yielding more milk on English pasture than on the Highland Scottish Moor, but the Jerseys performing better than the Highland cattle on pasture, whilst the Highland cattle perform better than the Jerseys on Highland Moore. 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 stabler temperament but less possibilities of intellectual achievement in a favourable environment". The third type of interaction involves genotype A performing better than genotype B independent of environment, but environment

X being better than Y for genotype A whereas environment Y is better than X for genotype B. Here, using the terminology of his day, he considered normal (A) and genetically mentally defective (B) children, where the first group performed better than the second in any type of school, but the second group do better in special schools than in standard schools, whereas the normal children do better in standard schools compared to 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-by-environment interaction and here Haldane considered length of life of English-origin populations doing better than long-term African origin groups when living in England, but in the African disease climate long-term African origin populations doing better than English migrants.

In Haldane's examples (also depicted in the figures) cross-overs of effect occur when outcomes are tabulated according to gene and environment combinations. He did not explicitly discuss examples of situations where a particular genotype has no influence on outcome in one environment but influences 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; the four examples which Haldane considered "exhaust[ed] the possibilities"

		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

References

Baird, P. (2000). Genetic technologies and achieving health for populations. International Journal of Health

Services, 30, 407-24.

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,, Gallı, 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;18,248–54.

Bovet, P., Paccaud, F. (2001) Alcohol, coronary heart disease and public health: which evidence-based policy? International Journal of Epidemiology, 30, 734-37.

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

Burr, M.L., Fehily, A.M., Butland, B.K., Bolton, C.H., Eastham, R.D. (1986) Alcohol and high-density-lipoprotein cholesterol: a randomiszed controlled trial. British Journal of Nutrition, 56, 81-86.

Caspi A, McClay J, Moffitt TE, Mill J, Martin J, Craig IW et al. (2002) Role of genotype in the cycle of violence in maltreated children. Science, 297:851-854.

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, 5, 461.

Clayton, D., McKeigue, P.M. (2001). Epidemiological methods for studying genes and environmental factors in complex diseases. Lancet, 358, 1356-60.

Darwin C. The Effects of Cross and Self Fertilisation in the Vegetable Kingdom. John Murray, London 1876.

Darwin GH. (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. James Lind Library (www.jameslindlibrary.org).

Davey Smith, G. (2006b). Cochrane Lecture: Randomised by (your) god: robust inference from an observational study design. Journal of Epidemiology and CommunityHhealth, 60, 382-388.

Davey Smith G, Ebrahim S. (2001) Epidemiology: is it time to call it a day? International Journal of Epidemiology, 30 :1–14.

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 2004; 33, 30-42.

Davey Smith, G., Ebrahim, S. (2005). What can Mendelian randomization tell us about modifiable behavioural and environmental exposures. BMJ , 330, 1076-9.

Davey Smith G, Ebrahim S. (2006) Folate supplementation and cardiovascular disease. Lancet, 366 :1679–81.

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, 1484-98.

Davey Smith, G., Harbord, R., Milton, J., Ebrahim, S., Sterne, J.A.C. (2005b) 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. (2008b). Clustered environments and randomized genes: a fundamental distinction between conventional and genetic epidemiology. PLoS Medicine, 4, 1985-1992

Davey Smith G, Phillips AN. (1996) Inflation in epidemiology: 'The proof and measurement of association between two things' revisited. Br Med J, 312:1659–61.

Davey Smith, G., Timpson, N., Ebrahim, S. (2008a). 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, 555-61.

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, 141-4.

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, 2623–27.

Fallon, U.B., Ben-Shlomo, Y., Elwood, P., Ubbink, J.B., Davey Smith, G. (2001). Homocysteine and coronary heart disease – Author's reply. Heart, published online March 14 (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, 221–9.

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, 659-61.

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

Gu, J., et al. (2005) Effects of N-acetyl transferase 1 and 2 polymorpshisms on bladder cancer risk in Caucasians. Mutation Research, 581, 97-104

Haldane, J.B.S. (1938). The Biology of Inequality. In: Heredity and Politics. George Allen and Unwin Ltd, London.

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. Br Med J, 318:1725–29.

Hartman, J.L., Garvik, B., Hartwell, L. (2001). Principles for the buffering of genetic variation. Science, 291, 1001-4.

Hashibe M, McKay JD, Curado MP, Oliveira JC, 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, 1432-4.

Haskell, W.L., Camargo, C., Williams, P.T., Vranizan, K.M., Krauss, R.M., Lindgren, F.T., et al. (1984) The effect of cessation and resumption of moderate alcohol intake on serum high-density-lipoprotein subfractions. New England Journal of Medicine, 310, 805-10.

Higuchi, S., Matsuushitam S., Imazeki, H., Kinoshita, T., Takagi, S., Kono, H. (1994). Aldehyde dehydrogenase genotypes in Japanese alcoholics. Lancet, 343, 741-2.

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, 30–7.

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, 730-36.

Kitami, T., Nadeau, J.H. (2002). Biochemical networking contributes more to genetic buffering in human and mouse metabolic pathways than does gene duplication. Nature Genet, 32, 191-4.

Klatsky, A.L. (2001) Commentary: Could abstinence from alcohol be hazardous to your health? International Journal of Epidemiology, 30, 739-42.

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, 606–10.

Lawlor, D.A., Harbord, R.M., Timpson, N.J., Lowe, G.D.O., Rumley, A., Gaunt, T.R et al. (2009). 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-acetyltransferase phenotype and risk in urinary bladder cancer: approaches in molecular epidemiology. Environmental Health Perspectives, 29, 71-9.

Marmot, M. (2001) Reflections on alcohol and coronary heart disease. International Journal of Epidemiology, 30, 729-34.

Memik, F. (2003) Alcohol and esophageal cancer, is there an exaggerated accusation? Hepatogastroenterology, 54, 1953-5

Morgan, T.H. (1935). The relation of genetics to physiology and medicine. The Scientific Monthly, 41, 5-18.results. Journal of the American Dietetic Association, 104: 942 – 950.

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: Genes, Behavior, and the Social Environment: Moving Beyond the Nature/Nurture Debate. The National Academies Press, Washington D.C.

Palmer, L., Cardon, L. (2005). Shaking tree: mapping complex disease genes with linkage disequilibrium. Lancet, 366, 1223-34.

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. JAMA,;301:2462-2471.

Rimm E. (2001) Commentary: Alcohol and coronary heart disease – laying the foundation for future work. International Journal of Epidemiology, 30, 738-39.

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 alcoholmediated carcinogenesis. Nature Reviews Cancer, 7, 599–612.

Shaper, A.G. (2001) Editorial: alcohol, the heart, and health. American Journal of Public Health 1993;83:799-801.

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, 112-45.

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

Tabery, J.G. (2000) Causation in the Nature-Nurture Debate: The Case of Genotype-Environment Interaction. (http://etd.library.pitt.edu/ETD/available/etd-07232007-121636/unrestricted/Tabery-Causation%5B1%5D.in.the.Nature-Nurture.Debate.pdf)

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 Medicine, 35, 532–44.

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., 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 NJ, Sayers A, Davey Smith G, Tobias JH. (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, 1577-1584.

Waddington, C.H. (1942). Canalization of development and the inheritance of acquired characteristics. Nature, 150, 563-5.

Warren, K.R., Li, T.K. (2005). Genetic polymorphisms: impact on the risk of fetal alcohol spectrum disorders. Birth Defects Research Clinical Molecular Teratololgy, 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.

Willett WC, Stampfer MJ. (2001) What vitamins should I be taking, doctor? New England Journal of Medicine, 345:1819–24.

Wolf, U. (1995). The genetic contribution to the phenotype. Human Genetics, 95, 127-48.

Youngman, L.D., Keavney, B.D., Palmer, A. 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 II), 31-32.