Hum Genet (2008) 123:15–33
DOI 10.1007/s00439-007-0448-6

REVIEW

Mendelian randomization: can genetic epidemiology help redress the failures of observational epidemiology?

Shah Ebrahim · George Davey Smith

Received: 3 August 2007 / Accepted: 9 November 2007 / Published online: 23 November 2007

Abstract Establishing causal relationships between environmental exposures and common diseases is beset with problems of unresolved confounding, reverse causation and selection bias that may result in spurious inferences. Mendelian randomization, in which a functional genetic variant acts as a proxy for an environmental exposure, provides a means of overcoming these problems as the inheritance of genetic variants is independent of—that is randomized with respect to—the inheritance of other traits, according to Mendel's law of independent assortment. Examples drawn from exposures and outcomes as diverse as milk and osteoporosis, alcohol and coronary heart disease, sheep dip and farm workers' compensation neurosis, folate and neural tube defects are used to illustrate the applications of Mendelian randomization approaches in assessing potential environmental causes of disease. As with all genetic epidemiology studies there are problems associated with the need for large sample sizes, the non-replication of findings, and the lack of relevant functional genetic variants. In addition to these problems, Mendelian randomization findings may be confounded by other genetic variants in linkage disequilibrium with the variant under study, or by population stratification. Furthermore, pleiotropy of effect of a genetic variant may result in null associations, as may canalisation of genetic effects. If correctly conducted and carefully interpreted, Mendelian randomization studies can provide useful evidence to support or reject causal hypotheses linking environmental exposures to common diseases.

S. Ebrahim (✉)
Noncommunicable Diseases Epidemiology Unit,
Department of Epidemiology and Population Health,
London School of Hygiene and Tropical Medicine,
Keppel Street, WC1E 7HT London, UK
e-mail: shah.ebrahim@lshtm.ac.uk

G. Davey Smith
MRC Centre for Causal Analyses in Translational Epidemiology,
Department of Social Medicine, University of Bristol,
Canynge Hall, Whiteladies Road, BS8 2PR Bristol, UK

Introduction

The causes of common multi-factorial diseases

Genetic epidemiology has been tested over the past decade with frequent failures to find robust replicable associations between genetic variants and common diseases (Davey Smith et al. 2005a). The recent publication of the first large-scale genome-wide association studies is, therefore, something of a relief for those who have pioneered the common variant—common disease hypothesis. We now have strong evidence of associations derived from the Wellcome Trust Case Control Collaboration (WTCCC) which has required large-scale collaboration, development in statistical methods and exploitation of large genotyping chips (Wellcome Trust Case Control Consortium 2007). The headline result of 24 robust strong associations ($P < 5 \times 10^{-7}$and a further 58 less strong associations (P: 10^{-5} to 5×10^{-7}) will provide genetics laboratories with a considerable amount of work to identify the underlying functional genetic variants that have been detected in this study. But perhaps more importantly, the study has crystallized a growing realisation that genetic associations with common multi-factorial diseases are not strong, and by implication, are unlikely to be useful for clinical prediction in the way in which Francis Collins originally speculated would be the case a decade ago (Collins 1999).

Of course, tracking down the genetic variants that regulate metabolic pathways of relevance to common diseases

will, hopefully, provide better understanding of molecular pathophysiology through gene expression, metabolomic and proteomic research. Furthermore, it is hoped that therapeutic targets will be identified to aid effective drug discovery. The likely existence of far more gene–environment interactions than we already have knowledge of provides justification for a combined genetic and environmental epidemiological approach to understanding causation (Khoury et al. 2005). But this newly generated knowledge may also hold the key to discovering modifiable environmental causes of common diseases that could contribute far more to improving public health than is allowed by our conventional understanding of the public health implications of contemporary genetics (Khoury et al. 2004). The traditional toolkit for environmental epidemiology has relied on collection of difficult to measure exposures, such as components of diet, and linked these exposures to disease risk. Here, we explore the limitations of such conventional methods and define ways in which genetic epidemiology can provide new tools for understanding environmental determinants of disease.

Limits of observational epidemiology

Over the past two decades there has been a growing concern about failures of replication of observational epidemiology findings in randomized trials testing the same hypotheses (Davey Smith and Ebrahim 2002). These concerns are particularly acute in areas of nutritional exposures and common chronic diseases, particularly cardiovascular diseases. It might seem fairly straightforward to measure dietary intakes and convert reports into macro- and micronutrients, and then compare risks of cardiovascular diseases in those with different levels of nutrient intake—and indeed, it is precisely this approach that has been widely adopted. However, it is seldom fully realised how two major problems—confounding and reverse causation—can lead to completely misleading causal inferences. A good example relates to the observation that homocysteine levels are associated with increased risk of cardiovascular disease. This was initially shown in case–control studies in which the possibility that increased levels were the result of metabolic consequences of having disease could not be discounted. When similar, albeit somewhat smaller, effects were demonstrated in prospective studies a stronger case for a causal association was possible (Wald et al. 2002). Homocysteine levels are determined, in part, by folic acid intake, which is modifiable and would make a useful preventive strategy. Trials of folate supplementation showed that homocysteine levels were reduced by supplementation and led to the establishment of large-scale secondary prevention trials in which people with established cardiovascular disease (or at high risk of developing disease) were randomized to folate supplementation or placebo. These trials started reporting in the past 5 years and have been uniformly and disappointingly negative (see Fig. 1) (Davey Smith and Ebrahim 2005a; Bazzano et al. 2006). This disappointment reflects millions of dollars of research investment and raises the justifiable questions of "were we misled" and if so "where did we go wrong?".

In this case, the answer may be that we were misled by both confounding and reverse causality. These factors operate in the following ways: homocysteine levels are affected by a wide range of environmental variables—in particular smoking, blood pressure, and socio-economic status—and these factors are not all that easy to measure well in epidemiological studies. Consequently, these factors may confound any association observed between homocysteine and cardiovascular diseases as shown in Fig. 1. Even in those studies that made relevant adjustments for potential confounders, because generally only simple and single measures of smoking, blood pressure and socio-economic position are made, the possibility of residual confounding exists. Furthermore, developing cardiovascular disease may lead to an increased homocysteine level, perhaps through an influence on renal function, and this reverse causality will generate an apparent prospective association between homocysteine level and end-stage cardiovascular disease (Zoccali et al. 2006).

To illustrate the problems that beset observational epidemiologists, two studies conducted by the same research group of the associations of dietary folic acid intake and stroke published in the same year generated diametrically opposite findings. The first, published in February 2004, of participants in the physicians' health study demonstrated a protective effect of high folate levels (He et al. 2004). The second published later that year using data from the nurses'

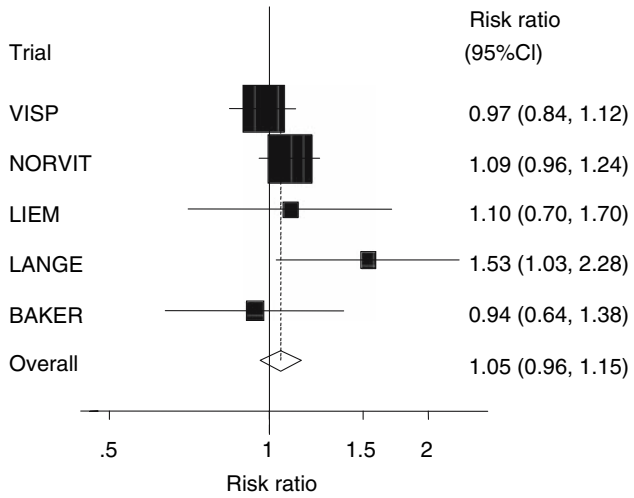

Fig. 1 Randomized controlled trials of folate/B vitamin supplementation and cardiovascular disease

health study demonstrated no important effect of folate intake on stroke risk (Al-Delaimy et al. 2004). The findings are shown in Table 1. One contribution to the apparent discrepancy may be that the "correct" answer for the association of dietary folate and stroke derived from a randomized controlled trial (RCT) was published between the two observational studies (Toole et al. 2004). The trial was actually cited in the second observational study in support of the negative finding. It is interesting that both studies showed similar effects in unadjusted analyses, but in the second study the effects attenuated on adjustment for potential confounders. A major difference between the two studies was the approach taken to adjusting for confounders. The second study adjusted for more confounders and also adjusted for vitamin E levels. In the British Women's Heart & Health Study, vitamin E levels are highly correlated with socio-economic position across the life-course (Lawlor et al. 2004). The significance of vitamin E adjustment is that as it is strongly associated with socio-economic position it thereby acts as an excellent adjustment for socio-economic confounding. Of course, in some circumstances randomized trials are not testing the same exposures as the observational studies—in particular, trials can only compare short-term exposure differences when life-long exposures may be more relevant in some cases.

Other processes in addition to confounding can generate robust, but non-causal, associations in observational studies. As mentioned above, reverse causation—where the disease influences the apparent exposure, rather than vice versa, may generate strong and replicable associations (see Fig. 2). For example, many studies have found that people with low circulating cholesterol levels are at increased risk of several cancers, including colon cancer. If causal, this is an important association as it might mean that lowering cholesterol levels to prevent coronary heart disease (CHD) would increase the risk of cancer. However, it is possible that the early stages of cancer may, many years before diagnosis or death, lead to a lowering in cholesterol levels, rather than low cholesterol levels increasing the risk of cancer. Similarly in studies of inflammatory markers such as

Table 1 Different effects of folate intake on risk of stroke by same research group in same year but using different analytic strategies

Quintiles of folate intake					
	1st	2nd	3rd	4th	5th
Folate does reduce the risk of stroke, men					
Age-adjusted relative risk	1	0.97	0.72	0.94	0.68
Adjusted relative risk (11 variables)	1	1.00	0.75	0.96	0.66
Folate does not reduce risk of stroke, women					
Age-adjusted relative risk	1	0.87	0.89	0.79	0.83
Adjusted relative risk (23 variables)	1	1.03	1.14	1.04	1.01

Data from He et al. (2004) and Al-Delaimy et al. (2004)

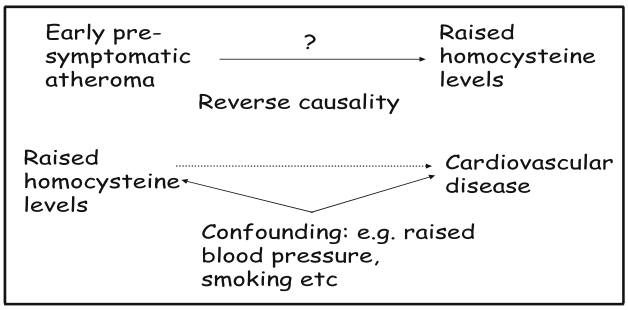

Fig. 2 Confounding and reverse causality may explain associations between homocysteine and cardiovascular disease

C-reactive protein (CRP) and cardiovascular disease risk it is possible that early stages of atherosclerosis—which is an inflammatory process starting early in life (McNamara et al. 1971)—lead to elevation in circulating inflammatory markers, and since people with atherosclerosis are more likely to experience cardiovascular events, a robust, but non-causal, association between levels of inflammatory markers and incident cardiovascular disease is generated (Davey Smith et al. 2005b). Reverse causation can also occur through behavioural processes—for example, people with early stages and symptoms of cardiovascular disease may reduce their consumption of alcohol, which would generate a situation in which alcohol intake appears to protect against cardiovascular disease. A form of reverse causation can also occur through reporting bias, with the presence of disease influencing reporting disposition. In case–control studies people with the disease under investigation may report on their prior exposure history in a different way than do controls—perhaps because the former will think harder about potential reasons that account for why they have developed the disease. Similarly the under-reporting of alcohol intake may be greater among those with symptoms of cardiovascular disease, since such people may have been advised to cut down on their drinking.

In observational studies associations between an exposure and disease will generally be biased if there is selection according to an exposure–disease combination in case–control studies, or according to an exposure–disease risk combination in prospective studies. Such selection may arise through differential participation in research studies, conducting studies in settings such as hospitals where cases and controls are not representative of the general population, or study of unusual populations (e.g. vegetarians). If, for example, those people experiencing an exposure but at low risk of disease for other reasons were differentially excluded from a study the exposure would appear to be positively related to disease outcome, even if there were no such association in the underlying population. This is a form of 'Berkson's bias', well known to epidemiologists (Berkson 1946). A possible example of such associative

selection bias relates to the finding in the large American Cancer Society volunteer cohort that high alcohol consumption was associated with a reduced risk of stroke (Thun et al. 1997). This is somewhat counter-intuitive as the outcome category included haemorrhagic stroke (for which there is no obvious mechanism through which alcohol would reduce risk) and because alcohol is known to increase blood pressure—a major causal factor for stroke. Population-based studies have found that heavy alcohol consumption tends to increase stroke risk, particularly haemorrhagic stroke (Hart et al. 1999; Reynolds et al. 2003). Heavy drinkers who volunteer for a study known to be about the health effects of their lifestyle are likely to be very unrepresentative of all heavy drinkers in the population, in ways that render them to be at low risk of stroke. Moderate and non-drinkers who volunteer may be more representative of moderate and non-drinkers in the underlying population. Thus the low risk of stroke in the heavy drinkers who volunteer for the study could erroneously make it appear that alcohol reduces the risk of stroke.

A further cause of concern in observational studies is that exposures are seldom measured repeatedly to properly characterize exposures that often show considerable within-person variability that leads to random measurement error. The strength of associations between truly causal risk factors and disease in observational studies is then underestimated due to this random measurement imprecision in characterizing the exposure. A century ago Charles Spearman demonstrated mathematically how such measurement imprecision would lead to what he termed the 'attenuation by errors' of associations (Spearman 1904; Davey Smith and Phillips 1996), which has, more recently, been renamed 'regression dilution bias' (Peto 1976).

For these reasons, observational studies can and do produce findings that either spuriously enhance or downgrade estimates of causal associations between modifiable exposures and disease. This has serious consequences for finding interventions to reduce disease risk in populations. It also undermines the scientific credibility of epidemiology as a discipline. For these reasons alternative approaches are needed; one of these is the Mendelian randomization framework that we review here.

Mendelian randomization

The basic principle of Mendelian randomization is that genetic variants which mirror the biological effects of a modifiable environmental exposure and alters disease risk should be associated with 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 2006; Davey Smith and Ebrahim 2003, 2004, 2005b, 2007).

Why "Mendelian randomization"?

Gregor Mendel (1822–1884) concluded from his hybridisation studies with pea plants that "the behaviour of each pair of differentiating characteristics [such as the shape and colour of seeds] in hybrid union is independent of the other differences between the two original plants" (Mendel 1866). This formulation was actually the only regularity that Mendel referred to as a "law", and in Carl Correns' 1900 paper (one of a trio appearing that year that are considered to represent the rediscovery of Mendel) he refers to this as Mendel's Law (Correns 1900; Olby 1966). Morgan discusses independent assortment and refers to this process as being realized "whenever two pairs of characters freely Mendelize" (Morgan 1913). Morgan's use of Mendel's surname as a verb did not catch on, but Morgan later christened this as Mendel's second law (Morgan 1918) and it has been known as this, or as "The Law of Independent assortment" since this time. The law suggests that inheritance of one trait is independent of—that is, randomized with respect to—the inheritance of other traits. The analogy with an RCT will clearly be most applicable to parent–offspring designs investigating the frequency with which one of two alleles from a heterozygous parent is transmitted to offspring with a particular disease. However, at a population level, traits influenced by genetic variants are generally not associated with the social, behavioural and environmental factors that confound relationships observed in conventional epidemiological studies. Thus while the 'randomization' is approximate and not absolute in genetic association studies, empirical observations suggest that it applies in most circumstances (Davey Smith et al. 2005b; Bhatti et al. 2005).

The term "Mendelian randomization" itself was first introduced in a somewhat different context, in which the random assortment of genetic variants at conception is utilized to provide an unconfounded study design for estimating treatment effects for childhood malignancies (Gray and Wheatley 1991; Wheatley and Gray 2004; Davey Smith 2007). Briefly, Gray and Wheatley wanted to obtain unbiased effect sizes for bone marrow transplantation for acute myeloid leukaemia where a randomized trial was not feasible, and direct comparisons of those transplanted and not transplanted would be uninterpretable due to selection by stage of disease, among other factors linked with prognosis. They reasoned that making comparisons of survival between leukaemic children on the basis of whether they had or did not have a genetically compatible sib (who could, in principle, provide a source of bone marrow), regardless of whether they actually had a transplant or not, was analogous to inten-

tion to treat analysis of randomised controlled trials. This approach has been used in further studies (Davey Smith 2007) but the term "Mendelian randomization" has recently become widely used with the meaning we ascribe to it here.

The notion that genetic variants can serve as an indicator of the action of environmentally modifiable exposures has been expressed in many contexts. For example, since the mid-1960s various investigators have pointed out that the autosomal dominant condition of lactase persistence is associated with milk drinking. Associations of lactase persistence with osteoporosis, bone mineral density or fracture risk thus provide evidence that milk drinking protects against these conditions (Birge et al. 1967; Newcomer et al. 1978; Honkanen et al. 1997; Corazza et al. 1995). In a related vein, it was proposed in 1979 that as N-acetyltransferase pathways are involved in the detoxification of arylamine, a potential bladder carcinogen, the observation of increased bladder cancer risk among people with genetically determined slow acetylator phenotype provided evidence that arylamines are involved in the aetiology of the disease (Lower et al. 1979). Since that time various commentators have pointed out that the associations of genetic variants of known function with disease outcomes provide evidence about aetiological factors (McGrath 1999; Ames 1999; Rothman et al. 2001; Brennan 2002; Kelada et al. 2003). However, these commentators have not emphasized the key strengths of Mendelian randomization—the avoidance of confounding, bias due to reverse causation or reporting tendency, and the underestimation of risk associations due to variability in behaviours and phenotypes (Davey Smith and Ebrahim 2004).

These key concepts were present in Martijn Katan's 1986 Lancet letter in which he suggested that genetic variants related to cholesterol level could be used to investigate whether the observed association between low cholesterol and increased cancer risk was real (Katan 1986) as shown in Fig. 3. Honkanen and colleagues also used these concepts in understanding how lactase persistence could better characterize the difficut-to-measure environmental influence of calcium intake than that could direct dietary reports (Honkanen et al. 1996).

Genetic variant	*ApoE* genotypes e3/e3	*ApoE* genotypes e2/e4	
Intermediate phenotype	Low cholesterol	High cholesterol	
Finding	↑cancer	↓cancer	Causal
Finding	= cancer	= cancer	Non-causal

Fig. 3 Katan's design for understanding the association between low blood cholesterol and cancers using a genetic variant

Since 2000 there have been several reports using the term 'Mendelian randomization' in the way it is used here (Davey Smith and Ebrahim 2003; Youngman et al. 2000; Fallon et al. 2001; Clayton and McKeigue 2001; Keavney 2002), and its use is becoming widespread (18,600 Google hits on 30 July 2007).

Phenocopy and genocopy

The exploitation of situations in which genotypic differences produce effects similar to environmental factors (and vice versa) clearly resonates with the concepts of phenocopy and genocopy in developmental genetics.

The term phenocopy is attributed to Goldschmidt (1938) and is used to describe the situation where an environmental effect could produce the same effect as was produced by a genetic mutation. As Goldschmidt wrote *"different causes produce the same end effect, presumably by changing the same developmental processes in an identical way"* (Goldschmidt 1938). In human genetics the term has generally been applied to refer to an environmentally produced disease state that is similar to a clear genetic syndrome. For example the niacin-deficiency disease pellagra is clinically similar to the autosomal recessive condition Hartnup disease (Baron et al. 1956), and pellagra has been referred to as a phenocopy of the genetic disorder (Snyder 1959; Guy 1993) Hartnup disease is due to reduced neutral amino acid absorption from the intestine and reabsorption from the kidney, leading to low levels of blood tryptophan which in turn leads to a biochemical anomaly which is similar to that seen when the diet is deficient in niacin (Kraut and Sachs 2005; Broer et al. 2004). Genocopy is a less widely utilized term, attributed to Schmalhausen (cited by Gause 1942) and has generally been considered to be the reverse of phenocopy—i.e. when genetic variation generates an outcome that could be produced by an environmental stimulus (Jablonka-Tavory 1982). It is clear that, even when the term genocopy is used polemically (Rose 1995) the two concepts are mirror-images reflecting differently motivated accounts of how both genetic and environmental factors influence physical state. For example Hartnup disease can be called a genocopy of pellagra, while pellagra can be considered a phenocopy of Hartnup disease. Mendelian randomization can, therefore, be viewed as an appreciation of the phenocopy–genocopy nexus that allows causation to be separated from association. However, the Mendelian randomization approach is distinct from the geneticist's search for causal mechanisms through appreciation of the biological interaction of genetic variants and environmental exposures. It seeks to define the causal, or non-causal, nature of environmental associations with common diseases, using genetic variants as proxies.

Phenocopies of major genetic disorders are generally rarely encountered in clinical medicine, but as Lenz comments, *"they are, however, most important as models which might help to elucidate the pathways of gene action"* (Lenz 1973). Mendelian randomization is concerned with less major (and thus common) disturbances and reverses the direction of phenocopy --> genocopy, to utilize genocopies, of known genetic mechanism, to inform us better about pathways through which the environment influences health.

The scope of phenocopy–genocopy has been discussed by Zuckerkandl and Villet (1988), who describe mechanisms through which there can be equivalence between environmental and genotypic influences. Indeed they state that *"no doubt all environmental effects can be mimicked by one or several mutations"*. The notion that genetic and environmental influences can be both equivalent and interchangeable has received considerable attention in developmental biology (West-Eberhard 2003; Leimar et al. 2006). Furthermore, population genetic analyses of correlations between different traits suggest there are common pathways of genetic and environmental influences, with Cheverud concluding that *"most environmentally caused phenotypic variants should have genetic counterparts and vice versa"* (Cheverud 1988).

Advantages of using genetic variants as proxies for environmental exposures

Given a general understanding that common diseases are neither "genetically" nor "environmentally" determined, a growing focus on examining the effects of a genetic variant in people who vary in their exposure to an environmental factor of interest is not surprising, and underpins much contemporary genetic epidemiology (Khoury et al. 1993). Presumably, the motivation behind this work is to identify individuals through genotyping in whom exposure to harmful environmental factors can be avoided—mirroring the value of genetic screening for some heritable conditions such as phenylketonuria. This has led some commentators to assume that Mendelian randomization can be understood as a form of gene–environment interaction study when, in fact, in Mendelian randomization, it is usually the *comparison* of main effects (genetic and environmental) that is of interest and not the interaction effect between them (Brennan 2004).

If main effects are of interest then why not simply measure the environmental exposures themselves rather than concern ourselves with both genetic and environmental effects? There are several crucial advantages of utilising functional genetic variants (or their markers) in this manner, which 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, for example, confound the association between homocysteine levels and CHD. This means that if a genetic variant is used to 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, excepting those with which they are in linkage disequilibrium. This powerful aspect of Mendelian randomization is illustrated in Tables 2 and 3 which shows the strong associations between key confounders and blood CRP levels, but no association of the same factors with genetic variants in the *CRP* gene.

Second, we have seen how 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 CRP. 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.

Table 2 Means or proportions of blood pressure, pulse pressure, hypertension and potential confounders by quarters of C-reactive protein (CRP) blood levels and by CRP genetic variants

	Means or proportions by quarters of C-reactive protein (range mg/l)				*P* trend across categories
	1(0.16–0.85)	2 (0.86–1.71)	3 (1.72–3.88)	4 (3.89–112.0)	
Hypertension (%)	45.8	49.7	57.5	60.0	<0.001
BMI (kg/m^2)	25.2	27.0	28.5	29.7	<0.001
Lifecourse socioeconomic position score	4.08	4.37	4.46	4.75	<0.001
Current smoker (%)	7.9	9.6	10.9	15.4	<0.001
Physically inactive (%)	11.3	14.9	20.1	29.6	<0.001
Moderate alcohol consumption (%)	22.2	19.6	18.8	14.0	<0.001

N = 3,529 (from Davey Smith et al. 2005b)

Table 3 Means or proportions of CRP systolic blood pressure, hypertension and potential confounders by 1059G/C genotype

	Means or proportions by genotype		*P*
	GG	GC or CC	
C-reactive protein (mg/l log scale)[a]	1.81	1.39	<0.001
Hypertension (%)	53.3	53.1	0.95
BMI (kg/m^2)	27.5	27.8	0.29
Lifecourse socioeconomic position score	4.35	4.42	0.53
Current smoker (%)	11.2	9.3	0.24
Physically inactive (%)	18.9	18.9	1.0
Moderate alcohol consumption (%)	18.6	19.8	0.56

[a] Geometric means

Third, associative selection bias in which selection into a study is related to both exposure level and disease risk and can generate spurious associations (as illustrated above with respect to alcohol and haemorrhagic stroke) are unlikely to occur with respect to genetic variants. For example empirical evidence supports a lack of association between a wide range of genetic variants and participation rates in three separate case–control studies: breast cancer, non-Hodgkins lymphoma, and lung cancer (Bhatti et al. 2005). Comparisons of genetic variants concerned with DNA repair, growth factors, immune responses, and oxidative stress (more than 100 SNPs and 15 tandem repeats) were compared in participants who had responded early or with minimal effort and participants who required incentives or increased time and contact to respond. Odds ratios for differences in prevalence of genetic variants between those willing and less willing to participate were generally null, showing no strong evidence to support any associations between genotype and willingness to participate in research (Bhatti et al. 2005). As these investigators noted, it is important that researchers test this assumption in their own data, as it is possible that other genotypes than those tested here, particularly those associated with health relevant behaviours (e.g. alcohol consumption), may show associations. Strong empirical evidence that selection bias is not an important concern for 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. In Fig. 4 (top panel) the results of statistical significance of differences in allele frequencies of 500,568 SNPs assayed using the Affimetrix 500k chip between subjects from the 1958 birth cohort and the UK blood donors, stratified by 12 broad regions of Britain. The bottom panel shows good agreement with the null distribution. The fact that very few robust differences between these two groups were found despite the difference in sampling and age indicates absence of selection bias effects with respect to a very wide range of SNP variants. Recently we have demonstrated the lack of association between a range of SNPs of known phenotypic effects and nearly 100 socio-cultrual, behavioural and biological risk factors for disease (Davey Smith et al. 2007).

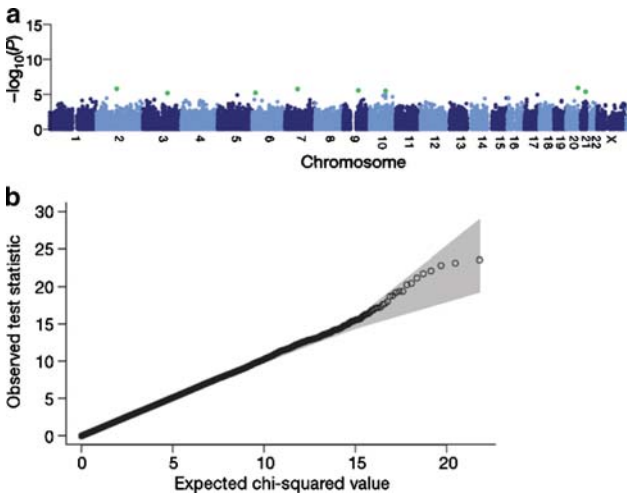

Fig. 4 Comparisons of 500,568 SNP variants in the 1958 birth cohort and UK blood donors. Data from Wellcome Trust Case Control Consortium, Nature (2007). **a** Statistical significance of differences in allele frequencies; **b** agreement with null distribution of differences. Reprinted by permission from Macmillan Publishers Ltd: Nature (doi:10.1038/nature)

Finally, a genetic variant will indicate long-term levels of exposure and if the variant is taken as a proxy for such exposure it will not suffer from the measurement error and thus attenuation by errors inherent in phenotypes that have high levels of variability. For example, groups defined by cholesterol-level related genotype will, over a long period, have experienced the cholesterol difference seen between the groups.

Mendelian randomization in animal studies

The approach to causal inference underlying Mendelian randomization has also been explicitly utilized in animal studies. For instance in investigations of the structural neuroanatomical factors underlying behavioural traits in rodents genetic crosses that lead to different on-average structural features have been carried out (Roderic et al. 1976; Weimer 1973; Lipp et al. 1989). Lipp et al. refer to this as "meiotic randomization" and consider that the advantages of this method are that the brain morphology differences that are due to genetic difference occur before any of the behavioural traits develop, and therefore the

brain morphology differences cannot be a feedback function of behaviour (which is equivalent to the avoidance of reverse causality in human Mendelian randomization studies) and that other difference between the animals are randomized with respect to the brain morphology differences of interest (equivalent to the avoidance of confounding in human Mendelian randomization studies). Li and colleagues (2006) apply this method to the dissection of adiposity and body composition in mice and point out that in experimental crosses *"meiosis serves as a randomization mechanism that distributes naturally occurring genetic variation in a combinatorial fashion among a set of cross progeny. Genetically randomized populations share the properties of statically designed experiments that provide a basis for causal inference. This is consistent with the notion that causation flows from genes to phenotypes. We propose that the inference of causal direction can be extended to include relationships among phenotypes"*. Mendelian randomization within epidemiology clearly reflects similar thinking among transgenic animal researchers. Williams and Wagner consider that *"A properly designed transgenic experiment can be a thing of exquisite beauty in that the results support absolutely unambiguous conclusions regarding the function of a given gene or protein within the authentic biological context of an intact animal. A transgenic experiment may provide the most rigorous test possible of a mechanistic hypothesis that was generated by previous observational studies. A successful transgenic experiment can cut through layers of uncertainty that cloud the interpretation of the results produced by other experimental designs*" (Williams and Wagner 2000). The problems of interpreting some aspects of transgenic animal studies may also apply to Mendelian randomization within genetic epidemiology, however, and linked progress across the fields of genomics, animal experimentation and epidemiology will better define the scope of Mendelian randomization in future.

Categories of Mendelian randomization

There are several categories of inference that can be drawn from studies utilising the Mendelian randomization paradigm that have been previously outlined in some detail (Davey Smith and Ebrahim 2003, 2004). The most obvious and direct category is use of a genetic variant that is related to the level of exposure ("exposure propensity") or to an intermediate phenotype believed to influence disease risk. Less direct evidence can come from genetic variant–disease associations which indicate that a particular biological pathway may be of importance, perhaps because the variants modify the effects of environmental exposures. A few illustrative examples will be given here.

Exposure propensity

Milk intake and bone health

Osteoporosis may be related to habitual low calcium intake, but measuring this exposure accurately is difficult. It has been suggested that assessing the association between calcium exposure and bone health may be done by comparing people with and without lactase persistence, as this may provide a better index of long-term low calcium intake (Honkanen et al. 1996). Lactase persistence is an autosomal dominant condition and an LCT polymorphism, -13910 T/C near the lactase phlorizin hydrolase gene has been found. In post-menopausal women, the CC genotype is strongly associated with low dietary intake of calcium from milk, a 10% lower bone mineral density at hip and spine, and a greater risk of non-vertebral fractures (see Fig. 5a–c; Obermayer-Pietsch et al. 2004). This provides strong evidence that milk drinking improves bone health, especially because directly studying milk intake is potentially beset with problems of confounding, reverse causation (people with bone problems may be told to drink more milk) and measurement error. Indeed in another field claims of associations between milk drinking and reduced risk of CHD (Elwood et al. 1991; Ness et al. 2001) have been criticized for inadequately dealing with confounding and reverse causation (Shaper et al. 1991).

Alcohol intake and health

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 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 by inducing facial flushing, palpitations, drowsiness and other symptoms. There are very considerable differences in alcohol consumption according to genotype (Takagi et al. 2002). However, 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. Consequently, 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).

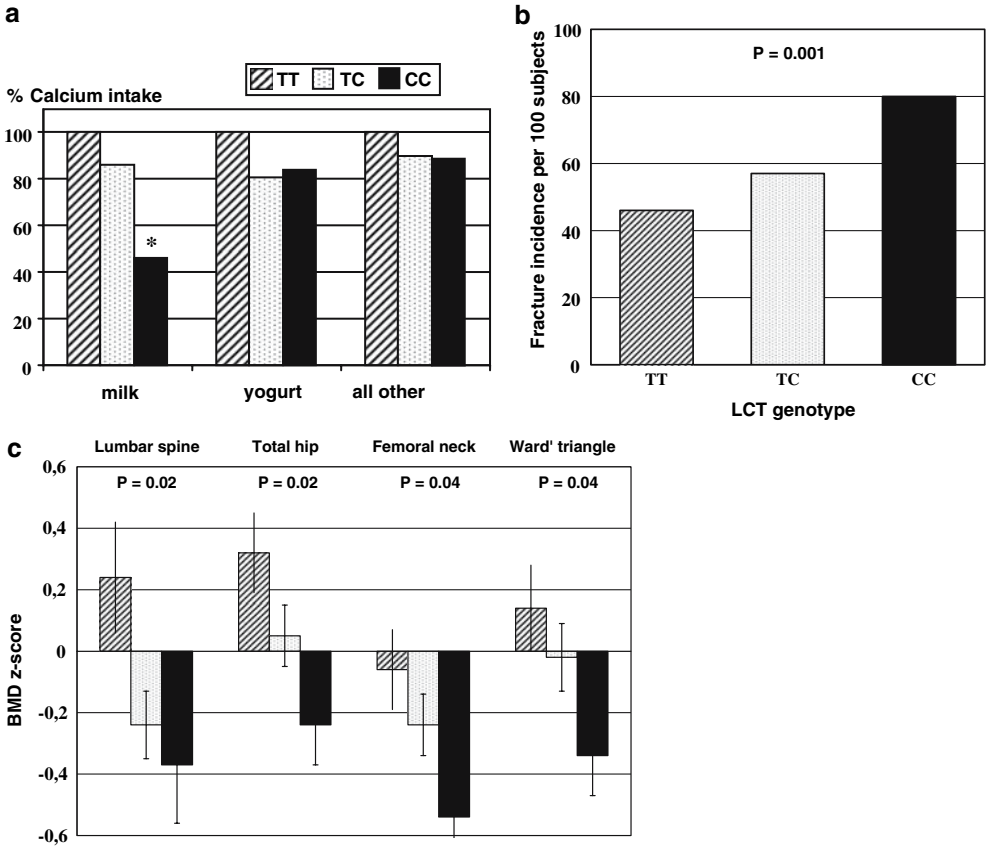

Fig. 5 **a** Milk drinking and fracture risk according to lactose persistence genotype. Individuals with genotype CC (*dark bars*) had lower calcium intake from milk (*$P = 0.004$) compared with TT (*dashed bars*), and TC (*shaded bars*) genotypes. **b** Fracture incidence per 100 subjects in post-menopausal women according to LCT genotypes. Individuals with genotype CC (*dark bars*) had a higher non-vertebral fracture incidence (*$P = 0.001$) than TC (*shaded bars*) and TT (*dashed bars*) genotypes, showing an increasing gene-dose effect towards these genotypes. **c** Bone Mineral Density *z*-score in post-menopausal women according to LCT genotypes. Indviduals with genotypes CC (*dark bars*) had a lower BMD score than TC (*shaded bars*) and TT (*dashed bars*) genotypes. From Obermayer-Pietsch et al. (2004)

Alcohol intake has also been postulated to increase the risk of oesophageal cancer; however, some have questioned the importance of its role (Memik 2003). Figure 6 presents findings from a meta-analysis of studies of ALDH2 genotype and oesophageal cancer risk (Lewis and Davey Smith 2005), clearly showing that people who are homozygous for the null variant, who therefore consume considerably less alcohol, have a greatly reduced risk of oesophageal cancer. Indeed this reduction in risk is close to that predicted by the joint effect of genotype on alcohol consumption and the association of alcohol consumption on oesophageal cancer risk in a meta-analysis of such observational studies (Gutjahr et al. 2001). When the heterozygotes are compared with the homozygous functional variant, an interesting picture emerges—the risk of oesophageal cancer is higher in the heterozygotes who drink rather less alcohol than those with the homozygous functional variant. This suggests that alcohol intake influences oesophageal cancer risk by increasing the level of acetaldehyde. The increased risk among heterozygotes is only apparent in those who drink some alcohol but metabolize it inefficiently, leading to high circulating levels of acetaldehyde. In this example the findings help to confirm both that alcohol is an environmentally modifiable risk factor for oesophageal cancer and suggest that acetaldehyde production is a mechanism through which alcohol has its effect.

Intermediate phenotypes

Genetic variants can influence circulating biochemical factors such as cholesterol, homocysteine, or fibrinogen levels. This provides a method for assessing causality in associations observed between these measures (*intermediate phenotypes*) and disease, and thus whether interventions to modify the intermediate phenotype could be expected to influence disease risk.

Cholesterol and coronary heart disease

Familial hypercholesterolaemia is a dominantly inherited condition in which many rare mutations (more than 700 DNA sequence variations; LDL receptor mutation cata-

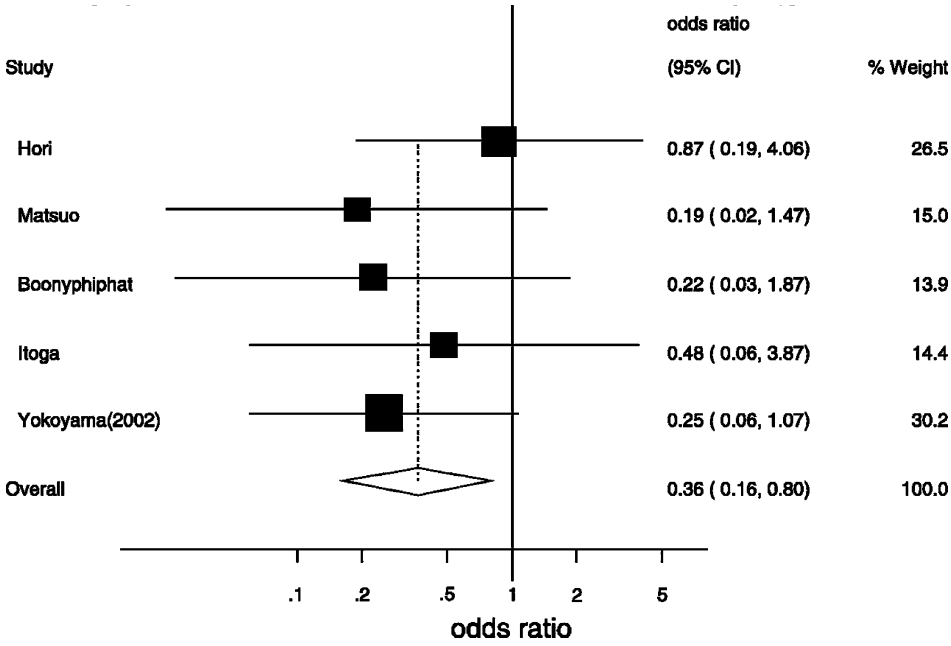

Fig. 6 Risk of esophageal cancer in individuals with the ALDH2*2*2 versus ALDH2*1*1 genotype. From Lewis and Davey Smith (2005)

logue 2003) of the low-density lipoprotein receptor gene (about 10 million people affected world-wide, a prevalence of around 0.2%), lead to high circulating cholesterol levels (Marks et al. 2003). The high risk of premature CHD in people with this condition was readily appreciated, with an early UK report demonstrating that by age 50 half of men and 12% of women had suffered from CHD (Slack 1969). Compared with the population of England & Wales (mean total cholesterol 6.0 mmol/l), people with familial hypercholesterolaemia (mean total cholesterol 9 mmol/l) suffered a 3.9-fold increased risk of CHD mortality, although very high relative risks among those aged less than 40 years have been observed (Scientific Steering Committee on behalf of the Simon Broome Register Group 1991). These observations, regarding genetically determined variation in risk, provided strong evidence that the associations between blood cholesterol and CHD seen in general populations reflected a causal relationship. The causal nature of the association between blood cholesterol levels and CHD has historically been controversial (Steinberg 2004). As both Steinberg (2005) and Færgeman (2003) discuss, many clinicians and public health practitioners rejected the notion of a causal link for a range of reasons. However from the late 1930s onwards evidence that people with genetically high levels of cholesterol had high risk for CHD should have been powerful and convincing evidence for the causal nature of elevated blood cholesterol in the general population.

With the advent of effective means of reducing blood cholesterol through statin treatment, there remains no serious doubt that the cholesterol–CHD relationship is causal. Among people without CHD, reducing total cholesterol levels with statin drugs by around 1 to 1.5 mmol/1 reduces CHD mortality by around 25% over 5 years (Heart Protection Study Collaborative Group 2002; Shepherd et al. 1995). Assuming a linear relationship between blood cholesterol and CHD risk, and given the difference in cholesterol of 3.0 mmol/1 between people with familial hypercholesterolaemia and the general population, the RCT evidence on lowering total cholesterol and reducing CHD mortality would predict a relative risk for CHD of around 2, as opposed to 3.9, for people with familial hypercholesterolaemia. However, the trials also demonstrate that the relative reduction in CHD mortality increases over time from randomization—and thus time with lowered cholesterol—as would be expected if elevated levels of cholesterol operate over decades to influence the development of atherosclerosis. People with familial hypercholesterolaemia would have had high total cholesterol levels throughout their lives, and this would be expected to generate a greater risk than that predicted by the results of lowering cholesterol levels for only 5 years. Furthermore, ecological studies relating cholesterol levels to CHD demonstrate that the strength of association increases as the lag period between cholesterol level assessment and CHD mortality increases (Rose 1982), again suggesting that long-term differences in cholesterol level are the important aetiological factors in CHD. As discussed above, Mendelian randomization is one method for assessing the effects of long-term differences in exposures on disease risk, free from the diluting problems of both measurement error and of only having short-term assessment of risk factor levels. This reasoning provides an indication that cholesterol-lowering efforts should be

lifelong rather than limited to the period for which RCT evidence with respect to CHD outcomes is available.

More recently, mutations in the gene coding for apolipoprotein B (apoB) have been found to produce a syndrome phenotypically indistinguishable from familial hypercholesterolaemia—Familial Defective ApoB (Soria et al. 1989; Tybjaerg-Hansen and Humphries 1992; Myant 1993). In a study of the Arg3500Gln mutation of the *APOB* gene, the basic principle behind Mendelian randomization can be demonstrated, in that Arg3500Gln heterozygotes had higher levels of total cholesterol but other CHD risk factors (including triglycerides, fibrinogen, glucose, body mass index and waist-hip ratio) did not differ from non-heterozygotes in the general population (Tybjærg-Hansen et al. 1998). The Arg3500Gln heterozygotes had a median 2.6 mmol/l higher blood cholesterol level and a high (but imprecise) odds ratio for CHD of 7.0 (95% CI 2.2 to 22) compared with the general population. As in the case of familial hypocholesterolaemia this is greater than that predicted by the RCT data, but again the differences in cholesterol by genotype will have been life-long, and the elevated CHD risk probably reflects the effects of long-term differences in cholesterol level.

Recently sequence variations in *PCSK9* associated with levels of LDL-cholesterol between 15 and 23% lower than levels in people without the mutant variants have been evaluated in the Atherosclerosis Risk in Communities study (ARIC), and considerably lower risks of CHD—between 47 and 88% lower—have been observed, depending on the level of LDL-cholesterol associated with each sequence variant (Cohen et al. 2006). Despite participants in ARIC having substantial burdens of other cardiovascular risk factors, these data indicate that *life-long* exposure to low levels of LDL-cholesterol (consistent with those achieved by statin treatment) is associated with markedly reduced risks of CHD, greater than the reductions observed for short-term cholesterol lowering in the statin trials. As other commentators have observed this is not surprising as atherosclerosis begins early in life, whereas statin treatment in later life would not be expected to achieve the same benefit (Brown and Goldstein 2006).

Identifying biological pathways for disease

The suggestion that taking aspirin reduces the risk of colon cancer originated from a case–control study exploring a large number of potential risk factors (Kune et al. 1988), but has been replicated in other studies (Sandler et al. 1998). Taking a Mendelian randomization approach provides one way of strengthening inference regarding the causal nature of this association. When examining variants in the gene coding for prostaglandin H synthase 2 (PTGS2), an enzyme involved in conversion of arachidonic acid to prostaglandin H_2 which is inhibited by aspirin (Lin et al. 2002), an association was found with reduced colon cancer risk. The investigators hypothesized that naturally occurring *PTGS2* variants might mimic long-term aspirin use. A larger study is required to confirm these exciting preliminary data. The data do, however, provide supportive evidence that aspirin (and other *PTGS2* inhibitors) protects against colon cancer, and that this protection is due to inhibition of the conversion of arachidonic acid to prostaglandin. Positive findings have been reported from two small randomized trials of aspirin in high-risk patients, providing further evidence in support of a causal role for aspirin (Sandler et al. 2003; Baron et al. 2003).

Modifiers of environment exposure

Sheep dip may be hazardous because of the organophosphates contained in it, but the vague symptoms farm workers attribute to exposure are not considered to be causal, but motivated by secondary gain from compensation (http://news.bbc.co.uk/1/hi/health/383003.stm, accessed 25 July 2007; http://news.bbc.co.uk/1/hi/health/537549.stm, accessed 25 July 2007). Conducting trials would be unethical and valid observational studies impossible, as reporting bias would be likely. Variants of the paraoxonase gene that produce forms of the enzyme paraoxonase with varying ability to detoxify organophosphates can be used to indicate the effects of different levels of sheep dip exposure. If organophosphates in sheep dip truly cause ill-health, then among people exposed to sheep dip a higher proportion of those with symptoms would be expected to carry the genetic variants related to less efficient detoxification, and this is what has been found (Cherry et al. 2002). Since it is unlikely that possession of the detoxification genotype is related to the tendency to report symptoms differentially, or to the desire for compensation, these findings provide evidence that sheep dip, and not compensation neurosis, is the cause of farm workers' symptoms.

Intergenerational influences—MTHFR polymorphisms and neural tube defects

Examining the effects of mother's genotype (independent of genotype of offspring) on the health outcomes of their offspring is a form of "intergenerational" Mendelian randomization, providing evidence on the role of intrauterine environment on the health of children. For example, periconceptual and early pregnancy folate deficiency are now known to be a cause of neural tube defects (NTDs), an effect confirmed by RCT evidence (MRC Vitamin Study Research Group 1991; Czeizel and Dudás 1992). The MTHFR 677C-->T polymorphism can be considered to mimic reduced folate and in a meta-analysis of case–con-

trol studies of NTDs, TT mothers had a twofold risk of having an infant with a neural tube defect compared with CC mothers (Botto and Yang 2000). The relative risk of a neural tube defect associated with the TT genotype in the infant was less than that observed with respect to maternal genotype, and there was no effect of paternal genotype on offspring neural tube defect risk. This suggests that it is the intra-uterine environment—influenced by maternal TT genotype—rather than the genotype of offspring that increases the risk of NTD (Davey Smith and Ebrahim 2003), and that higher maternal folate intake would reduce the risk of offspring NTDs, as found in the trials.

Implications of Mendelian randomization study findings

Establishing the causal influence of environmentally modifiable risk factors from Mendelian randomization designs informs policies for improving population health through population-level interventions. They do not imply that the appropriate strategy is genetic screening to identify those at high risk and application of selective exposure reduction policies. For example, the implications of studies on maternal MTHFR genotype and offspring NTD risk are that population risk for NTDs can be reduced through increased folate intake peri-conceptually and in early pregnancy. It does not suggest that women should be screened for MTHFR genotype; women without the TT genotype but with low folate intake are still exposed to preventable risk of having babies with NTDs. Similarly establishing the association between genetic variants (such as familial defective ApoB) associated with elevated cholesterol level and CHD risk strengthens causal evidence that elevated cholesterol is a modifiable risk factor for CHD for the whole population. Thus, even though the population attributable risk for CHD of this variant is small it usefully informs public health approaches to improving population health. It is this aspect of Mendelian randomization that illustrates its distinction from conventional risk identification and genetic screening purposes of genetic epidemiology.

Mendelian randomization and randomized controlled trials

Randomized controlled trials 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, however (Hingorani and Humphries 2005). Figure 7 illustrates this, drawing attention to the unconfounded nature of exposures proxied for by genetic variants (analogous to the unconfounded nature of a randomized intervention, and see Tables 2, 3 for explicit demonstration of the potential for confounding by intermediate phenotype but not by genotype), the lack of possibility of reverse causation as an influence on exposure-outcome associations in both Mendelian randomization and randomized controlled trial 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 proxied for exposure within any particular individual.

The analogy with randomized controlled trials is also useful with respect to one objection that has been raised with respect to Mendelian randomization studies. This is that the environmentally modifiable exposure proxied for

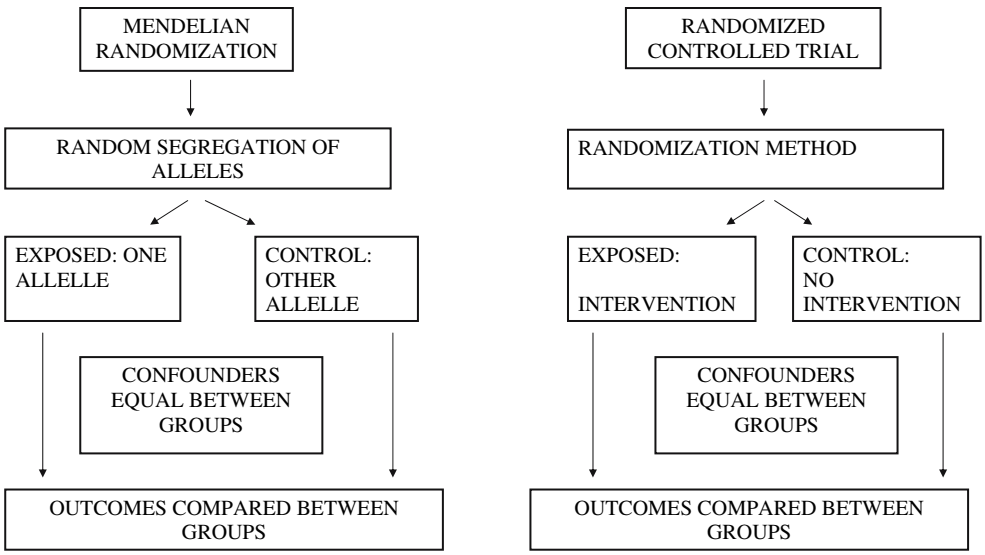

Fig. 7 Mendelian randomization and randomized controlled trial designs compared, after Hingorani and Humphries (2005)

by the genetic variants (such as alcohol intake or circulating CRP levels) are influenced by many other factors in addition to the genetic variants (Jousilahti and Salomaa 2004). This is of course true. However, consider a randomized controlled trial 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 vitiate the power of RCTs; neither does it vitiate the strength of Mendelian randomization designs.

Although Mendelian randomization approaches can be helpfully compared to RCTs, the analogy has been used to discuss several previously documented theoretical limitations of the approach (Nitsch et al. 2006). One area where there is now empirical data which supports caution in meta-analysis of Mendelian randomization studies—often necessary to achieve adequate power—is the variable relation between the MTHFR genotype and the CHD risk, dependent on the place of study, which may reflect differences in folate supplementation between populations leading to a gene–environment interaction (Lewis et al. 2005). As with any study design, it is crucial that investigators consider the ways in which underlying assumptions can be tested empirically.

Analysis of Mendelian randomization studies

At the heart of a Mendelian randomization study of the effects of an intermediate phenotype on disease are three relationships: genotype–intermediate phenotype; intermediate phenotype–disease; genotype–disease. Measuring these relationships by direct observation is straightforward but the effect sizes for the genotype–intermediate phenotype and genotype–disease associations tend to be very small and will be estimated with large imprecision if only small sample sizes are available. This limits the inferences that can be made from simple triangulation approaches in which the genotype–disease effect observed is compared to the expected effect derived from difference in intermediate phenotype.

The analytic approach can be illustrated using the example of homocysteine and CHD that we used in our original exposition of Mendelian randomization (Davey Smith and Ebrahim 2003). The first step is to calculate the MTHFR genotype–homocysteine association and from a meta-analysis of observational studies TT individuals had a 2.6 μmol/l higher homocysteine level compared with CC individuals. Using a meta-analysis of the observational studies of blood homocysteine levels and CHD, a 2.6 μmol/l higher homocysteine results in a relative risk of CHD of 1.13 (95% confidence intervals 1.08–1.19). A meta-analysis of studies relating MTHFR genotype to CHD risk, TT individuals experienced a risk of 1.16 (95% CI 1.05–1.28) compared to CC individuals. Thus, the two relative risk estimates—one subject to bias due to confounding and reverse causality and the other not—are similar from which it may be inferred that blood homocysteine level is causally related to CHD. More generally, an unbiased estimate of odds ratio of disease per unit change in intermediate phenotype (assuming linearity of the relation of intermediate phenotype difference and log odds ratio of disease) is given by OR (genotype–disease)$^{1/\Delta P}$ where ΔP is the difference in intermediate phenotype between genotypes (Minelli et al. 2004). In earlier accounts we suggested that an approach to assess the statistical significance of the difference between the genotype–disease and intermediate phenotype–disease difference was to use a *t*-test using the standard errors of both estimates as the denominator (Davey Smith and Ebrahim 2003; Davey Smith et al. 2004).

This simple triangulation approach has flaws in that it fails to take account of random error in estimation of the genotype–intermediate phenotype association. More recently, other analytic options have been advocated. Simulation approaches have been developed in which large numbers of randomly selected values of the relative risks of both the genotype–disease and intermediate phenotype–disease are used to assess bias in terms of the differences between the relative risks in a similar manner to estimating limits of agreement between two measures (Bautista et al. 2006). This approach has some important statistical limitations, however (Thomas et al. 2007) Instrumental variable methods have also been promoted as a means of analysis in which the genetic variant is treated as an instrument which is assumed to be associated with the disease only through its association with the intermediate phenotype (Timpson et al. 2005; Lawlor et al. 2007). Genetic variants that are not associated reasonably strongly (in statistical terms) with an intermediate phenotype will be "weak" instruments and will potentially yield biased estimates of the unbiased relative risk of the intermediate phenotype on disease.

A further issue, less often discussed, is that as the effect sizes are generally going to be small, even modest ascertainment biases (for example, selective loss of participants with specific genotypes) could lead to false inferences. Checking for Hardy–Weinberg equilibrium as a means of testing for possible ascertainment bias can be of value, and deviation from it should lead to a review of both accuracy of genotyping and possible selective losses in the sample studied.

Sample size issues

Sample size calculations need to consider the magnitude of both the effect of genotype on the modifiable risk factor that is being proxied for and the predicted effect of the modifiable risk factor on disease outcome. This often leads to very large studies being required, and failure to recognize this can lead to studies being uninformative. For example, in a report of a case–control study entitled "Elevated plasma fibrinogen. Cause or consequence of cardiovascular disease?" (Van der Bom et al. 1998), the relative risk of CHD for a 1 g higher blood fibrinogen level was 1.45 (95% CI 1.12–1.88), whereas the association between genotype and CHD risk was essentially null (relative risk 1.08 95% CI 0.71–1.65 for GA and AA genotypes compared with GG genotype). The authors interpreted these results by indicating that fibrinogen was not a cause of CHD. However, given the strength of the association between genotype and fibrinogen, with GA plus AA individuals having 0.17 g/l higher fibrinogen than GG individuals, the predicted risk according to genotype, given the observational association between fibrinogen and CHD, would be around 1.07. This is clearly not different from the estimated relative risk—indeed the point estimates are close, although there is a very wide confidence interval around the relative risk for genotype. Thus the authors' claim that their study suggests that fibrinogen is not causally related to the risk of CHD is not supported by evidence from their own study, although later larger studies and meta-analysis suggest their conclusion was correct (Davey Smith et al. 2005c; Keavney et al. 2006).

Table 4 shows the scale of sample sizes required to reliably detect the very small associations that are likely to exist between risk alleles and common diseases. The scale of these estimates indicates how much larger than most contemporary studies future collaborations will need to work on the weaker signals found in the WTCCC. However, large sample sizes alone do not protect from false inferences, as commentators on this work on fibrinogen make clear (Meade et al. 2006). The value of the Mendelian randomization approach is that it suggests scientific questions that need to be answered before the enormous costs of drug development—a fibrinogen lowering drug here—are entertained. For example, do animal models of fibrinogen gene variants show evidence of developmental canalization? Does the functional activity of fibrinogen differ in people with different fibrinogen gene variants? Do fibrinogen variants generate pleotrophic effects? More generally, if polymorphisms at more than one locus influence an intermediate phenotype then it may be possible to explore combinations of polymorphisms at different loci that produce differences in intermediate phenotype that are substantial enough to generate detectable effects on disease outcome. If the loci are not in linkage disequilibrium and thus segregate independently this could be termed "factorial Mendelian randomization", with interest being in the groups in which the combination of polymorphisms produce the most extreme difference in intermediate phenotype. If haplotypes that produce more extensive phenotypic differences than single SNPs could be studied, as they have been in the case of CRP and insulin resistance (Timpson et al. 2005), some relaxation of sample sizes would be possible. However, the quantum of variation in plasma CRP levels explained by CRP haplotypes is still small, in the range of 3–6%, and absolute differences in CRP level by haplotype (about 6–7 mg/l) (Kivimäki et al. 2007) are similar to the differences obtained in studies using single SNP variants (Casas et al. 2006).

Table 4 Sample size estimates for detection of odds ratios (OR) associations been susceptibility allele and disease for a power of 90% and $P = 0.001$

Odds Ratio	Frequency of susceptibility allele in controls					
	1%	5%	10%	20%	30%	40%
1.1	221927	46434	24626	13987	10759	9505
1.2	58177	12217	6509	3730	2896	2581
1.3	27055	5702	3051	1763	1380	1240
1.5	10604	2249	1213	712	566	516
2.0	3193	687	377	229	188	177
4.0	598	134	78	53	46	47

From Davey Smith et al. (2005c)

Problems and limitations of Mendelian randomization

Mendelian randomization shows considerable promise but has limitations that have been discussed in our first paper on the topic in some detail (Davey Smith and Ebrahim 2003) and are listed in Table 5. Several of these—need for large sample sizes, non-replication of findings, lack of functional genetic variants related to the pathway of interest—are common to all genetic epidemiological studies. In addition to these, the interpretation of findings from studies that appear to fall within the Mendelian randomisation remit can often be complex, as has been previously discussed with respect to *MTHFR* and folate intake (Davey Smith and Ebrahim 2003). As a second example consider the association of extracellular superoxide dismutase (EC-SOD) and CHD. EC-SOD is an extracellular scavenger of superoxide anions and thus genetic variants associated with higher circulating EC-SOD levels might be considered to mimic higher levels of antioxidants. However, findings are dramatically opposite to this—bearers of such variants have an increased risk of CHD (Juul et al. 2004). The explanation of this apparent paradox is that the higher circulating EC-SOD levels associated with the variant may arise from

Table 5 Limitations of Mendelian Randomisation, adapted from Davey Smith and Ebrahim (2003) and Bochud et al. (2007)

Failure to establish reliable genotype–intermediate phenotype or genotype–disease associations
• Genotyping errors
• Misclassification of phenotype
• Confounding by population structure
• Lack of power
• Chance
• Publication bias
Confounding of genotype–intermediate phenotype–disease associations
Pleiotropy and the multi-function of genes
Canalization and developmental stability
Lack of suitable polymorphisms for studying modifiable exposures of interest
Complexity of metabolic pathways
Large sample sizes required
Transmission ratio distortion
Parent of origin effects

movement of EC-SOD from arterial walls; thus, the in situ antioxidative properties of these arterial walls are lower in individuals with the variant associated with higher circulating EC-SOD. The complexity of these interpretations—together with their sometimes speculative nature—detracts from the transparency that otherwise makes Mendelian randomization attractive.

Mendelian randomization is not concerned with the question of whether or not 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 (Rose 1995; Berkowitz 1999; Baird 2000; Holtzman 2001; Strohman 1993). 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. Indeed in the phenocopy/genocopy example of pellagra and Hartnup disease discussed above, only a minority of the Hartnup gene carriers develop symptoms, but at a group level they have a much greater tendency to such symptoms and a shift in amino acid levels that reflect this (Scriver 1988, Scriver et al. 1987). These group-level differences create the analogy between Mendelian randomization and RCTs, outlined in Fig. 7.

Although the primary comparisons in a Mendelian randomization study involve triangulation of the effects observed between genotype–intermediate phenotype, intermediate phenotype–disease, and genotype and disease, it is clearly of importance to gather evidence of non-confounding by genotype, to examine a range of relevant outcomes to demonstrate any unexpected pleiotropy of the genetic variant under study, and to examine other intermediate phenotypes of relevance to the metabolic pathways involved to examine possible canalisation of the genetic effect.

Recent commentary on Mendelian randomization has highlighted two additional potential limitations related to transmission ratio distortion in which the distribution of alleles at a locus differs in surviving offspring from that expected from Mendelian principles (Bochud et al. 2007). This may arise during meiosis or as a result of selective survival after conception. The second problem is of parent of origin effects in which the function of an allele depends on which parent it was inherited from and is due to imprinting—an epigenetic effect. With increased knowledge of gene effects more data will be available to explore these issues in Mendelian randomization studies.

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.

Exploiting new opportunities

As indicated at the outset, the WTCCC genome wide association study will provide new impetus to expanding the range of genetic variants available for use as proxies for intermediate phenotypes in Mendelian randomization studies. Currently, Mendelian randomization studies are severely limited by the availability of genetic variants with clear effects relevant to environmental exposures of interest. For example, the *CYP27B1* SNP that is involved in vitamin D activation is related to only small differences in blood levels of vitamin D, limiting its use in studies attempting to clarify the role of vitamin D in common diseases (Hyppönen et al. 2007).

Further developments that will improve our chances of finding modifiable causes of common diseases are in the

large-scale population-based studies, either adequately powered case–control studies of categorical disease outcomes or intensively phenotyped studies in which genetic variants contributing to traits can be identified. In such studies considerably more effort needs to be put in to determine how best to use finite stored plasma resources for biomarker assays, and further consideration of improved cost-effective phenotyping of such population resources will be required. Of fundamental importance in the era of genome wide association studies is the need to standardize the reporting and cumulative presentation of findings so that robust knowledge is generated rather than noise due to a massive numbers of false positive associations—work that HuGENet is taking forward (Khoury et al. 2007).

Conclusions

The Mendelian randomization approach has developed rapidly over the past 5 years, and considerable discussion of its role in aetiological research has resulted. Proof of principle is now abundant. The need for large sample sizes remains a limitation currently, but is soluble with the greater recognition of the importance of collaborative studies which are now yielding substantive replicable findings for common genetic variants and common phenotypes such as obesity and diabetes (Frayling et al. 2007; Zeggini et al. 2007) These collaborations will provide the major infrastructure required for future well-powered Mendelian randomization studies.

References

Al-Delaimy WK, Rexrode KM, Hu FB, Albert CM, Stampfer MJ, Willett WC, Manson JE (2004) Folate intake and risk of stroke among women. Stroke 35:1259–1263

Ames BN (1999) Cancer prevention and diet: help from single nucleotide polymorphisms. PNAS 96:12216–12218

Baird P (2000) Genetic technologies and achieving health for populations. Int J Health Serv 30:407–24

Baron DN, Dent CE, Harris H, Hart EW, Jepson JB (1956) Hereditary pellagra-like skin rash with temporary cerebellar ataxia, constant renal amino-aciduria, and other bizarre biochemical features. Lancet 271:421–429

Baron JA, Cole BF, Sandler RS et al (2003) A randomized trial of aspirin to prevent colorectal adenomas. N Eng J Med 348:891–899

Bautista LE, Smeeth L, Hingorani AD, Casas JP (2006) Estimation of bias in nongenetic observational studies using "Mendelian Triangulation". Ann Epidemiol 16:675–680

Bazzano LA, Reynolds K, Holder KN, He J (2006) Effect of folic acid supplementation on risk of cardiovascular diseases: a meta-analysis of randomized controlled trials. JAMA 296:2720–2726

Berkowitz A (1999) Our genes, ourselves?. Bioscience 46:42–51

Berkson J (1946) Limitations of the application of fourfold table analysis to hospital data. Biometrics Bull 2:47–53

Bhatti P, Sigurdson AJ, Wang SS, Chen J, Rothman N, Hartge P, Bergen AW, Landi MT (2005) Genetic variation and willingness to participate in epidemiological research: data from three studies. Cancer Epidemiol Biomarkers Prev 14:2449–2453

Birge SJ, Keutmann HT, Cuatrecasas P, Whedon GD (1967) Osteoporosis, intestinal lactase deficiency and low dietary calcium intake. N Engl J Med 276:445–448

Bochud M, Chiolero A, Elston RC, Paccaud F (2007) A cautionary note on the use of Mendelian randomization to infer causation in observational epidemiology. Int J Epidemiol. doi:10.1093/ije/dym186

Botto LD, Yang Q (2000) 5, 10-Methylenetetrahydrofolate reductase gene variants and congenital anomalies: a HuGE review. Am J Epidemiol 151:862–877

Brennan P (2002) Gene environment interaction and aetiology of cancer: what does it mean and how can we measure it? Carcinogenesis 23(3):381–387

Brennan P (2004) Mendelian randomization and gene–environment interaction. Int J Epidemiol 33:17–21

Broer S, Cavanaugh JA, Rasko JEJ (2004) Neutral amino acid transport in epithelial cells and its malfunction in Hartnup disorder. Transporters 33:233–236

Brown MS, Goldstein JL (2006) Lowering LDL—not only how low, but how long? Science 311:1721–1723

Casas JP, Shah T, Cooper J, Hawe E, McMahon AD, Gaffney D, et al (2006) Insight into the nature of the CRP-coronary event association using Mendelian randomization. Int J Epidemiol 35:922–931

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

Cherry N, Mackness M, Durrington P et al (2002) Paraoxonase (PON1) polymorphisms in farmers attributing ill health to sheep dip. Lancet 359:763–764

Cheverud JM (1988) A comparison of genetic and phenotypic correlations. Evolution 42:958–968

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

Cohen JC, Boerwinkle E, Mosely TH, Hobbs HH (2006) Sequence variations in PSCK9, low LDL, and protection against coronary heart disease. N Engl J Med 354:1264–1272

Collins FS (1999) Medical and societal consequences of the Human Genome Project. N Engl J Med 341:28–37

Corazza GR, Benati G, Di Sario A et al (1995) Lactose intolerance and bone mass in postmenopausal Italian women. Br J Nutr 73:479–487

Correns CG (1900) Mendel's Regel über das Verhalten der Nachkommenschaft der Bastarde. Berichte der Deutschen Botanischen Gesellschaft 8:158–68. English translation, G. Mendel's law concerning the behavior of progeny of varietal hybrids. In: Stern and Sherwood, pp 119–32. WH Freeman and Co., San Francisco (1966)

Czeizel AE, Dudás I (1992) Prevention of the first occurrence of neural-tube defects by periconceptional vitamin supplementation. N Engl J Med 327:1832–1835

Davey Smith G (2006) Cochrane lecture: randomised by (your) god: robust inference from an observational study design. J Epidemiol Community Health 60:382–388

Davey Smith G (2007) Capitalising on Mendelian randomization to assess the effects of treatments. J R Soc Med 100:432–435

Davey Smith G, Ebrahim S (2002) Data dredging, bias, or confounding (editorial). BMJ 325:1437–1438

Davey Smith G, Ebrahim S (2003) 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? Int J Epidemiol 32:1–22

Davey Smith G, Ebrahim S (2004) Mendelian randomization: prospects, potentials, and limitations. Int J Epidemiol 33:30–42

Davey Smith G, Ebrahim S (2005a) What can Mendelian randomization tell us about modifiable behavioural and environmental exposures. BMJ 330:1076–1079

Davey Smith G, Ebrahim S (2005b) Folate supplementation and cardiovascular disease. Lancet 366:1679–1681

Davey Smith G, Ebrahim S (2007) Mendelian randomization: genetic variants as instruments for strengthening causal inference in observational studies. In: Vaupel JW, Weinstein M (eds) Bio-social surveys: current insight and future promise. The National Academies Press, National Research Council, Washington, DC

Davey Smith G, Phillips AN (1996) Inflation in epidemiology: 'The proof and measurement of association between two things' revisited. Br Med J 312:1659–1661

Davey Smith G, Harbord R, Ebrahim S (2004) Fibrinogen, C-reactive protein and coronary heart disease: does Mendelian randomization suggest the associations are non-causal? Q J Med 97:163–166

Davey Smith G, Ebrahim S, Lewis S, Hansell A, Palmer LJ, Burton P (2005a) Genetic epidemiology and public health: hope, hype, and future prospects. Lancet 366:1484–1498

Davey Smith G, Lawlor D, Harbord R, Timpson N, Rumley A, Lowe G, Day I, Ebrahim S (2005b) Association of C-reactive protein with blood pressure and hypertension: lifecourse confounding and Mendelian randomization tests of causality. Arterioscler Thromb Vasc Biol 25:1051–1056

Davey Smith G, Harbord R, Milton J, Ebrahim S, Sterne JAC (2005c) Does elevated plasma fibrinogen increase the risk of coronary heart disease?: evidence from a meta-analysis of Genetic Association Studies. Arterioscler Thromb Vasc Biol 25:2228–2233

Davey Smith G, Lawlor D, Harbord R, Timpson N, Day I, Ebrahim S (2007) Clustered environments and randomized genes: a fundamental distinction between conventional and genetic epidemiology. PLoS Medicine (in press)

Elwood PC, Yarnell JWG, Burr ML et al (1991) Epidemiological studies of cardiovascular disease: progress report VII. MRC Epidemiology Unit, Cardiff

Enomoto N, Takase S, Yasuhara M, Takada A (1991) Acetaldehyde metabolism in different aldehyde dehydrogenase-2 genotypes. Alcohol Clin Exp Res 15:141–144

Færgeman O (2003) Coronary artery disease: genes drugs and the agricultural connection. Elseveir, Netherlands

Fallon UB, Ben-Shlomo Y, Davey Smith G (2001) Homocysteine and coronary heart disease. Heart Online, March 14th. http://heart.bmjjournals.com/cgi/eletters/85/2/153

Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, Lindgren CM et al (2007) A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science 316:889–894

Gause GF (1942) The relation of adaptability to adaption. Q Rev Biol 17:99–114

Goldschmidt RB (1938) Physiological genetics. McGraw-Hill, New York

Gray R, Wheatley K (1991) How to avoid bias when comparing bone marrow transplantation with chemotherapy. Bone Marrow Transplant 7(suppl 3):9–12

Gutjahr E, Gmel G, Rehm J (2001) Relation between average alcohol consumption and disease: an overview. Eur Addict Res 7:117–27

Guy JT (1993) Oral manifestations of systematic disease. In: Cummings CW et al (eds) Otolaryngology—head and neck surgery, vol 2. Mosby, St Louis

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 318:1725–1729

Hasin D, Aharonovich E, Liu X, Mamman Z, Matseoane K, Carr L, Li TK (2002) Alcohol and ADH2 in Israel: Ashkenazis, Sephardics, and recent Russian immigrants. Am J Psychiatry 159:1432–1434

He K, Merchant A, Rimm EB, Rosner BA, Stampfer MJ, Willett WC, Ascherio A (2004) Folate, vitamin B6, and B12 intakes in relation to risk of stroke among men. Stroke 35:169–174

Heart Protection Study Collaborative Group (writing committee: Collins R, Armitage J, Parish S, Sleight, Peto R) (2002) MRC/BHF Heart Protection Study of cholesterol lowering with simvastatin in 20,536 high risk individuals: a randomised placebo-controlled trial. Lancet 360:7–22

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

Hingorani A, Humphries S (2005) Nature's randomised trials. Lancet 366:1906–1908

Holtzman NA (2001) Putting the search for genes in perspective. Int J Health Serv 31:445

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

Honkanen R, Kröger H, Alhava E, Turpeinen P, Tuppurainen M, Saarikoski S (1997) Lactose intolerance associated with fractures of weight-bearing bones in Finnish Women aged 38–57 years. Bone 21:473–477

Hyppönen E, Davey Smith G, Power C (2007) Vitamin D status and self-perceived health: conventional and Mendelian randomisation approaches. J Epidemiol Community Health 61(suppl1):A14

Jablonka-Tavory E (1982) Genocopies and the evolution of interdependence. Evol Theory 6:167–170

Jousilahti P, Salomaa V (2004) Fibrinogen, social position, and Mendelian randomisation. J Epidemiol Community Health 58:883

Juul K, Tybjaerg-Hansen A, Marklund S, Heegaard NHH, Steffensen R, Sillesen H, Jensen G, Nordestgaard BG (2004) Genetically reduced antioxidative protection and increased ischaemic heart disease risk: the Copenhagen city heart study. Circulation 109:59–65

Katan MB (1986) Apoliopoprotein E isoforms, serum cholesterol, and cancer. Lancet I:507–508 (reprinted IJE 2004;34:9)

Keavney B (2002) Genetic epidemiological studies of coronary heart disease. Int J Epidemiol 31:730–736

Kelada SN, Eaton DL, Wang SS, Rothman NR, Khoury MJ (2003) The role of genetic polymorphisms in environmental health. Environ Health Perspect 111:1055–1064

Khoury M, Beaty TH, Cohen BH (1993) Fundamentals of genetic epidemiology. Oxford University Press, Oxford, p 13, 126

Khoury M, Little J, Burke W (2004) Human genome epidemiology. Oxford University Press, Oxford

Khoury M, Davis R, Gwinn M, Lindegren ML, Yoon P (2005) Do we need genomic research for the prevention of common diseases with environmental causes? Am J Epidemiol 161:799–805

Khoury MJ, Little J, Gwinn M, Ioannidis JPA (2007) On the synthesis and interpretation of consistent but weak gene–disease association studies in the era of genome-wide association studies. Int J Epidemiol 36:439–445

Kivimäki M, Lawlor DA, Davey Smith G, Eklund C, Murme M, Lehtimäki T, Viikari JS, Raitakari OT (2007) Variants in the CRP Gene as a measure of lifelong differences in average C-reactive protein levels. The cardiovascular risk in young Finns study, 1980–2001. Am J Epidemiol 166:760–764

Kraut JA, Sachs G (2005) Hartnup disorder: unravelling the mystery. Trends Pharmacol Sci 26:53–55

Kune GA, Kune S, Watson LF (1988) Colorectal cancer risk, chronic illnesses, operations and medications: case control results from the Melbourne Colorectal Cancer Study. Cancer Res 48:4399–4404

Lawlor DA, Harbord R, Sterne JAC, Timpson N, Davey Smith G (2007) Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. Stats Med. doi: 10.1002/sim.3034

Lawlor DA, Davey Smith G, Bruckdorfer KR et al (2004) Those confounded vitamins: what can we learn from the differences between observational versus randomised trial evidence? Lancet 363:1724–1727

LDL receptor mutation catalogue. http://www.ucl.ac.uk/fh. Accessed 16 Dec 2003

Leimar O, Hammerstein P, Van Dooren TJM (2006) A new perspective on developmental plasticity and the principles of adaptive morph determination. Am Nat 167:367–376

Lenz W (1973) Phenocopies. J Med Genet 10:34–48

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 Epidemiol Biomarkers Prev 14:1967–1971

Lewis SJ, Ebrahim S, Davey Smith G (2005) Meta-analysis of MTHFR 677C->T polymorphism and coronary heart disease: does totality of evidence support causal role for homocysteine and preventive potential of folate? BMJ 331:1053

Li R, Tsaih SW, Shockley K, Stylianou IM, Wergedal J,Paigen B, Churchill GA (2006) Structural model analysis of multiple quantative traits. PLoS Genet 2:1046–1057

Lin HJ, Lakkides KM, Keku TO, Reddy ST, Louie AD et al (2002) Prostaglandin H Synthase 2 variant (Val511Ala) in African Americans may reduce the risk for colorectal neoplasia. Cancer Epidemiol Biomarkers Prev 11:1305–1315

Lipp HP, Schwegler H, Crusio WE, Wolfer DP, Leisinger-Trigona MC, Heimrich B, Driscoll P (1989) Using genetically-defined rodent strains for the identification of hippocampal traits relevant for two-way avoidance behaviour: a non-invasive approach. Experientia 45:845–859

Lower GM, Nilsson T, Nelson CE et al (1979) N-acetylransferase phenotype and risk in urinary bladder cancer: approaches in molecular epidemiology. Environ Health Perspect 29:71–79

Marks D, Thorogood M, Neil HAW, Humphries SE (2003) A review on diagnosis, natural history and treatment of familial hypercholesterolaemia. Atherosclerosis 168:1–14

McGrath J (1999) Hypothesis: is low prenatal vitamin D a risk-modifying factor for schizophrenia? Schizophr Res 40:173–177

McNamara JJ, Molot MA, Stremple JF, Cutting RT (1971) Coronary artery disease in combat casualties in Vietnam. JAMA 216:1185–1187

Meade TW, Humpries SE, De Stavola BL (2006) Commentary: fibrinogen and coronary heart disease—test of causality by "Mendelian" randomization by Keavney et al. Int J Epidemiol 35:944–947

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

Mendel G (1866) Experiments in plant hybridization. http://www.mendelweb.org/archive/Mendel.Experiments.txt

Minelli C, Thompson JR, Tobin MD, Abrams KR (2004) An integrated approach to the meta-analysis of genetic association studies using Mendelian randomization. Am J Epidemiol 160:445–452

Morgan TH (1913) Heredity and sex. Columbia University Press, New York

Morgan TH (1918) Physical basis of heredity

MRC Vitamin Study Research Group (1991) Prevention of neural tube defects: results of the Medical Research Council vitamin study. Lancet 338:131–137

Myant NB (1993) Familial defective apolipoprotein B-100: a review, including some comparisons with familial hypercholesterolaemia. Atherosclerosis 104:1–18

Ness AR, Davey Smith G, Hart C (2001) Milk, coronary heart disease and mortality. J Epidemiol Community Health 55:379–382

Newcomer AD, Hodgson SF, Douglas MD, Thomas PJ (1978) Lactase deficiency: prevalence in osteoporosis. Ann Intern Med 89:218–220

Nitsch D, Molokhia M, Smeeth L, De Stavola B, Whittaker JC, Leon DA (2006) Limits to causal inference based on Mendelian randomization: a comparison with randomized controlled trials. Am J Epidemiol 163:397–403

Obermayer-Pietsch BM, Bonelli CM, Walter DE, Kuhn RJ, Fahrleitner-Pammer A, Berghold A, Goessler W, Stepan V, Dobnig H, Leb G, Renner W (2004) Genetic predisposition for adult lactose intolerance and relation to diet, bone density, and bone fractures. J Bone Miner Res 19:42–47

Olby RC (1966) Origins of Mendelism. Constable London

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

Peto R (1976) Two properties of multiple regression analysis and regression to the mean (and regression from the mean). In: Fletcher CM, Peto R, Tinker CM, Speizer FE (eds) The natural history of chronic bronchitis and emphysema: an eight year study of early chronic obstructive lung disease in working men in London. Oxford University Press, Oxford, pp 218–223

Reynolds K, Lewis LB, Nolen JDL, Kinney GL, Sathya B, He J (2003) Alcohol consumption and risk of stroke: a meta-analysis. JAMA 289:579–588

Roderic TH, Wimer RE, Wimer CC (1976) Genetic manipulation of neuroanatomical traits. In: Petrinovich L, McGaugh JL (eds) Knowing thinking and believing. Plenum, New York

Rose G (1982) Incubation period of coronary heart disease. BMJ 284:1600–1601

Rose S (1995) The rise of neurogenetic determinism. Nature 373:380–382

Rothman N, Wacholder S, Caporaso NE, Garcia-Closas M, Buetow K, Fraumeni JF (2001) The use of common genetic polymorphisms to enhance the epidemiologic study of environmental carcinogens. Biochim Biophys Acta 1471:C1–C10

Sandler RS, Galanko JC, Murray SC, Helm JF, Woosley JT (1998) Aspirin and nonsteroidal anti-inflammatory gents and risk for colorectal adenomas. Gastroenterology 114:441–447

Sandler RS, Halabi S, Baron JA et al (2003) A randomized trial of aspirin to prevent colorectal adenomas in patients with previous colorectal cancer. N Engl J Med 348:883–890

Scientific Steering Committee on behalf of the Simon Broome Register Group (1991) Risk of fatal coronary heart disease in familial hyper-cholesterolaemia. BMJ 303:893–896

Scriver CR (1988) Nutrient–gene interactions: the gene is not the disease and vice versa. Am J Clin Nutr 48:1505–1509

Scriver CR, Mahon B, Levy HL (1987) The Hartnup phenotypeL Mendelain transport disorder, multifactorial disease. Am J Hum Genet 40:401–412

Shaper AG, Wannamethee G, Walker M (1991) Milk, butter and heart disease. BMJ 302:785–786

Shepherd J, Cobbe SM, Ford I et al for the West of Scotland Coronary Prevention Study Group (1995) Prevention of coronary heart disease with pravastatin in men with hypercholesterolemia. N Engl J Med 333:1301–1307

Slack J (1969) Risks of ischaemic heart disease in familial hyperlipoproteinaemic states. Lancet 2:1380–1382

Snyder LH (1959) Fifty years of medical genetics. Science 129:7–13

Soria LF, Ludwig EH, Clarke HRG, Vega GL, Grundy SM, McCarthy BJ (1989) Association between a specific apolipoprotein B mutation and familial defective apolipoprotein B-100. Proc Natl Acad Sci USA 86:587–591

Spearman C (1904) The proof and measurement of association between two things. Am J Psychol 15:72–101

Steinberg D (2004) Thematic review series: the pathogensis of atherosclerosis. An interpretive history of the cholesterol controversy: part 1. J Lipid Res 45:1583–1593

Steinberg D (2005) Thematic review series: the pathogensis of atherosclerosis. An interpretive history of the cholesterol controversy:

part II: the early evidence linking hypercholestrolemia to cornary disease in humans. J Lipid Res 46:179–190

Strohman RC (1993) Ancient genomes, wise bodies, unhealthy people: the limits of a genetic paradigm in biology and medicine. Perspect Biol Med 37:112–145

Takagi S, Iwai N, Yamauchi R, Kojima S, Yasuno S, Baba T, Terashima M, Tsutsumi Y, Suzuki S, Morii I, Hanai S, Ono K, Baba S, Tomoike H, Kawamura A, Miyazaki S (2002) Aldehyde dehydrogenase 2 gene is a risk factor for myocardial infarction in Japanese Mmen. Hypertens Res 25:677–681

Keavney B, Danesh J, Parish S, Palmer A, Clark S, Youngman L, Delépine M, Lathrop M, Peto R, Collins R The International Studies of Infarct Survival (ISIS) Collaborators (2006) Fibrinogen and coronary heart disease: test of causality by 'Mendelian randomization'. Int J Epidemiol 35:935–943

Thomas DC, Lawlor DA, Thompson JR (2007) Re: Estimation of bias in nongenetic observational studies using "Mendelian Triangulation". Ann Epidemiol 17:511–513

Thun MJ, Peto R, Lopez AD (1997) Alcohol consumption and mortality among middle-aged and elderly US adults. New Engl J Med 337:1705–1714

Timpson NJ, Lawlor DA, Harbord RM, Gaunt TR, Day INM, Palmer LJ, Hattersley AT, Ebrahim S, Lowe GDO, Rumley A, Davey Smith G (2005) C-reactive protein and its role in metabolic syndrome: mendelian randomization study. Lancet 366:1954–1959

Toole JF, Malinow MR, Chambless LE, Spence JD, Pettigrew LC, Howard VJ et al (2004) Lowering homocysteine in patients with ischemic stroke to prevent recurrent stroke, myocardial infarction, and death: the Vitamin Intervention for Stroke Prevention (VISP) randomized controlled trial. JAMA 294:565–575

Tybjaerg-Hansen A, Humphries SE (1992) Familial defective apolipoprotein B-100: a single mutation that causes hypercholesterolemia and premature coronary artery disease. Atherosclerosis 96:91–107

Tybjærg-Hansen A, Steffenson R, Meinertz H, Schnohr P, Nordestgaard BG (1998) Association of mutations in the apolipoprotein B gene with hypercholesterolemia and the risk of ischemic heart disease. New Engl J Med 338:1577–1584

Van der Bom JG, De Maat MPM, Bots ML, Haverkate F, De Jong PTVM, Hofman A, Kluft C, Grobbee DE (1998) Elevated plasma fibrinogen. Cause or consequence of cardiovascular disease? Arterioscler Thromb Vasc Biol 18:621–625

Wald DS, Law M, Morris JK (2002) Homocysteine and cardiovascular disease: evidence on causality from a meta-analysis. BMJ 325:1202–1206

Weimer RE (1973) Dissociation of phenotypic correlation: response to posttrial etherization and to temporal distribution of practice trials. Behav Genet 3:379–386

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

West-Eberhard MJ (2003) Developmental plasticity and evolution. Oxford University Press, Oxford

Wheatley K, Gray R (2004) Commentary: Mendelian randomization—an update on its use to evaluate allogeneic stem cell transplantation in leukaemia. Int J Epidemiol 33:15–17

Williams RS, Wagner PD (2000) Transgenic animals in integrative biology: approaches and interpretations of outcome. J Appl Physiol 88:1119–1126

Wolf U (1995) The genetic contribution to the phenotype. Hum Genet 95:127–148

Youngman LD, Keavney BD, 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

Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, Lango H et al (2007) Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. Science 316:1336–1341

Zoccali C, Testa A, Spoto B, Tripepi G, Mallamaci F (2006) Mendelian randomization: a new approach to studying epidemiology in ESRD. Am J Kidney Dis 47:332–341

Zuckerkandl E, Villet R (1988) Concentration—affinity equlivalence in gene regulation: convergence and environmental effects. Proc Natl Acad Sci USA 85:4784–4788