Acta Genet Med Gemellol 41: 261-273 (1992)

Seventh International Congress
on Twin Studies

The Epidemiology of Genetic Epidemiology

J.L. Hopper

The University of Melbourne, Victoria, Australia

Abstract. Familial aggregation for disease is important; strong familial risk factors must exist even if the increased risk to a relative of an affected individual is modest. It is in practice difficult, however, to conduct studies in genetic epidemiology which conform to strict epidemiological principles. For twin studies there are two major questions: Are twins 'no different' from the population on which inference is to be made? Are study twins 'no different' to twins in the population? The importance of each question of bias depends on the scientific question, the trait(s) studied, and sampling issues. The strength of the twin design is its ability to refute the null hypothesis that genetic factors do not explain variation in a trait. Following the Popperian paradigm, alternate hypotheses should be considered in depth (both theoretically and empirically), with a design and sample size sufficient to exclude not just naive explanations. More sophisticated statistical techniques are now being applied, so the philosophy, assumptions, and limitations of statistical modelling must be appreciated. The concept of 'heritability' has, in the past, been misunderstood and misused. New advances in DNA technology promise to revolutionise epidemiological thinking, and so case-control-pedigree designs may well become standard tools. The strengths and limitations of studies based on related individuals as the sampling unit are discussed.

Key words: Epidemiology, Familial aggregation, Genetic dominance, Genetic epidemiology, Heritability, Popperian philosophy, Sampling, Shared environment, Twins.

INTRODUCTION

What is genetic epidemiology?

Epidemiology has been defined as the "study of the distribution and determinants of health-related states and events in populations" [32]. Morton [33] defined *genetic*

epidemiology as "a science that deals with etiology, distribution and control of disease *in groups of relatives* and with *inherited causes of disease* in populations". (In this context, "inherited" is meant to include biological and non-biological inheritance, the latter including also cultural inheritance [6]).

The feature common to the two disciplines is that study is in terms of the *population*, rather than the individual (family). Note that the additional italicised features of genetic epidemiology, however, are *familial* and not necessarily genetic. Genetic epidemiology developed in the main from population genetics, and consequently has been seen historically to be more a component of genetics rather than of epidemiology. This has been compounded by epidemiologists in general having ignored genetic factors, possibly due to their focus on disease determinants which have the potential to be modified.

In addition, epidemiologists would appear to have misunderstood the implications of even moderate familial aggregation in a disease. Even if the increased risk to a relative of an affected person is as small as 1.5 to 2, it is important. If having an affected relative increases one's risk for a disease by a factor of R, there could be an underlying familial risk factor which is associated with the disease by a risk ratio of $10R$ or more, that is, the strength of the underlying risk factor is an order of magnitude greater than the perceived increase in risk associated with an affected relative. This has, in theory, been demonstrated for the risk factor being defined by a single genetic locus [8,34], an 'environmental' exposure [30], or a continuous (genetic or environmental) variable [1,23], with similar conclusions. For example, a doubling of disease risk associated with an affected relative is consistent with a continuous risk factor which has an interquartile risk ratio for disease of 10 or 20, and a correlation between relatives of 1 or 0.5, respectively [23]. Therefore, efforts to understand why a disease 'runs in the family' are justified because they could uncover one or more genetic and/or environmental determinants for the disease which when combined have a substantial risk gradient. New advances in DNA technology promise to revolutionise epidemiological thinking.

Research in genetic epidemiology, however, poses greater practical problems than standard epidemiological studies. By definition, families form the unit of study and these must be *ascertained* through *probands*. For reasons of sampling and participation, it is difficult to conduct genetic epidemiology studies which conform to strict epidemiological principles. If conclusions are to be applicable to the population, the method of ascertainment must be unambiguously described in terms of the population from which probands are selected. (Linkage studies based on atypical and highly selected kinships, although essential for generating hypotheses, could be considered *not* to be part of true genetic epidemiology). Reports by probands of disease in their relatives are known to be subject to error, which may be quite substantial. Results, therefore, will be biased if recall depends on whether the proband is a case or control. Thus, the participation of, and not just information on, relatives is to be preferred. Consequently, the total sample size will be an order of magnitude or more greater than the number of probands.

Studies of twin pairs constitute the minimal sets of relatives and are possibly the easiest studies in genetic epidemiology. Because they have more genetic and environmental 'information' in common than other pairs of relatives, twin studies can make an immensely important contribution to genetic epidemiology.

SOME EPIDEMIOLOGICAL CONSIDERATIONS OF TWIN STUDIES

In genetic epidemiology, twins are of special interest because they offer an 'experiment of opportunity'. These pairs of individuals of the same age, who share all or, on average, half their genes can be studied from the viewpoint of (a) their similarity in disease state, (b) their similarity in disease determinants or risk factors, (c) a difference in their disease state, and (d) a difference in their exposure to risk factor(s). Each of these approaches can be used to address questions about disease aetiology. Twins are particularly useful for longitudinal studies, as they generally know they are of special scientific interest and so are usually not adverse to being approached more than once.

Two questions, however, must be addressed: (i) can twins be considered 'no different' from non-twin individuals in the population under consideration, with respect to the traits and issues of interest, and (ii) are the twins in a given study 'no different' from twins in the population, again with respect to the traits and issues under consideration? The importance of these questions of bias, however, depends on several factors. The first factor is the scientific issue under consideration. If inference is to be made about the genetic and environmental causes of a disease, then both the preceding questions must be answered in the affirmative. If, on the other hand, twins discordant for an exposure such as cigarette smoking are studied to determine associated effects on disease risk or other traits, then question (ii) is obviously not relevant, while question (i) may warrant consideration. The same would apply to a cohort study of twins in a longitudinal study, in which outcomes in time are to be related back to differences between and within twin pairs across the sample at baseline, or at previous time points. The second factor is demography, which is equally important as 'the population' must be defined in terms of time and space. Other factors relate to sampling, eg. what processes were used to ascertain the twins of the study sample? Was there population-based sampling? Were the twins identified from a registry, and if so, how were twins recruited onto the registry? Lastly, what was the actual response rate? Did it differ according to known characteristics of the twins, such as their sex, zygosity, educational status, and so on? Was it thought to be dependent on unmeasured characteristics? Finally and most important, if there was a differential response, how did this affect the study objectives and findings?

The Australian NHMRC Twin Registry is a listing of the names and addresses of twins (or their parents if twins are less than 18 years old) who have volunteered (or have been volunteered by their parents) to consider being involved in research projects. From 1978, almost 25,000 pairs of an estimated national twin population of 200,000 pairs have been registered. The proportion of twins registered is known to vary by age, sex, zygosity, and state of residence [2], and by some aspects of health status [18]. It is also considered to vary according to ethnicity, socioeconomic and educational status, and factors related to the strength of relationship between twins. The response rate of studies performed using this Registry can depend on the age, sex and other characteristics of the group of twins being approached as despite considerable efforts to maintain a current address listing, this cannot be assured especially among the younger age groups. *As discussed above, these issues may not necessarily be detrimental to study objectives, yet they need to be known and considered when interpreting study findings.*

To gain insight into the process of recruitment to the Australian Registry, consider

the following information obtained from a recent follow-up to the 1968 Tasmanian Asthma Study [27]. In 1968 a survey was carried out of all seven-year-old Tasmanian school children, which achieved a 98% response rate with 8,596 'probands' studied [16,17]. Amongst these, 91 pairs of twins (accounting for 2.14% of seven-year-olds) were identified, coming close to an expected 94 pairs based on national twin birth rates in 1961 [9]. Among the siblings of probands, 165 twin pairs were identified (1.56% of all siblings), which was less than the expected 243 pairs based on national twin birth rates. During 1991-92 a study was conducted which included an attempt to trace and to mail a health questionnaire to all these identified twin pairs. To date, 83% of twin probands and 85% of twin siblings compared to 76% of the 930 nontwin probands have been traced. Of these, completed questionnaires have been received from 63% of twin probands and 68% of twin siblings, compared to 76% of nontwin probands. Therefore there was no statistical difference in the overall response rate of twins and nontwins.

Identified twins were invited by mail to register with the Australian NHMRC Twin Registry. Prior to 1990, less than 15% of these twins were registered. Of all proband twins, 72% have now registered, compared to 60% of all twin siblings. The response rate to the health questionnaire was 69% and 75% in *registered* proband and sibling twins, respectively, compared to 50% and 58% in unregistered twins. Despite several approaches, over one-third of identified pairs have not registered, and although twins on the registry were more likely to complete the questionnaire, about one-quarter of these did not respond. Further analyses to determine factors which differentiate registering from non-registering twins have failed to reveal evidence related to sex, zygosity, or family size, but the registration rate appears to be higher in twins whose father was employed in a professional occupation in 1968, when the twins were children.

THE CLASSIC TWIN METHOD

The twin design owes its popularity to having the ability to refute, in an efficient and convincing manner, the null hypothesis that genetic factors do not explain variation in a trait. If the correlation between monozygotic pairs is significantly greater (in a statistical sense) than between dizygotic pairs of the same sex, then under certain assumptions, the alternate hypothesis that genetic factors do play a role in trait variation is preferred to the null hypothesis.

Possibly due to the epidemiological difficulties referred to above in conducting family studies, *replication* of twin studies is rare, while *refutation* of hypotheses generated from previous studies is virtually non-existent. Following the Popperian approach to science [35], alternate hypotheses should be considered in depth, both theoretically and empirically. There has been continuous debate in the epidemiological literature concerning the necessity or otherwise of, and the difficulties in, applying this paradigm to epidemiology; see eg. Greenland [15] and Rothman [38] for a collection of opinions. Many of the issues raised apply naturally to genetic epidemiology.

The design and sample size of twin studies should be such that more than just simplistic alternate explanations can be excluded with adequate statistical power. A study of relatively few monozygotic (MZ) and dizygotic (DZ) pairs may reveal that the correlation or disease concordance between MZ pairs is (statistically) greater than between DZ

pairs, and therefore consistent with a simple genetic model, under the assumptions of the classic twin model. There may be little statistical power, however, to test the basic assumptions of the model, and therefore it is tempting to transfer belief from the null hypothesis (ie. genetic factors do not play a role) to a specific alternate hypothesis (eg. additive genetic factors exist). Data sets consistent with this latter hypothesis may also be consistent with a range of alternate hypotheses, especially if the sample size is small. Weak 'goodness-of-fit' tests will only serve falsely to prop up belief.

The genetic component of variation (an additive component with or without a dominance component) predicts a specific pattern of covariation or correlation between relatives. This pattern, however, is not unlike what would be expected if trait similarity was determined solely by environmental factors shared by relatives. Whilst a neat theoretical foundation has been derived for the genetic model [13], there are almost limitless possibilities for the effects of common environments. Despite this, modelling of the common or shared environment by geneticists or by scientists with a particular interest in genetics has, in general, been simplistic and naive. The assumption that the strengths of effects common to twin pairs are the same for both MZ and DZ pairs is merely a *convenience* and has rarely been addressed or, if so, only superficially.

Careful examination of the assumption has, in our experience, been informative. For example, an approach which takes into account the cohabitational history of pairs of relatives in cross-sectional data [19] has revealed substantial changes in covariation with cohabitation. This was most evident for the 'environmental' trait, ie. lead level in blood [20], but it was also present in analyses of personality traits [22] and of alcohol consumption, depression and anxiety [7]. The latter study suggested that effects attributable to a common environment could depend on cohabitational status for MZ and DZ pairs in the same way, or in quite different ways, depending on the trait. These issues have been explored by Rose and others [37], and more recently using longitudinal data [41] also.

Sociologists, psychologists, and other scientists have accumulated substantial knowledge about factors which influence behaviour. Many of these factors have the potential to be common to members of the same family, at least while they are living together. In most studies, a number of these factors are measured by the researchers, usually by questionnaire. Evidence relevant to these factors may be available from blood or tissue samples eg. blood lead levels could be a useful indicator of degree of shared environment. Unfortunately, such evidence is almost universally ignored in analyses by twin reasearchers. *In general, modelling of the shared environment has not done justice to either the data at hand or to the researchers' biological and sociological knowledge.*

STATISTICAL MODELLING

Greater computational power has seen developments in the application of methods of statistical analysis. In particular, methods based on the maximum likelihood theory which require an iterative solution can now be carried out without undue computational delay, and can exhibit flexibility in modelling, not available in more restrictive approaches based on explicit solutions [19,21]. *Statistical modelling* has come into vogue, not only in epidemiology but also in the analysis of twin and family data, Eaves et al

[10]. The philosophy behind the underlying assumptions and the limitations of statistical modelling must therefore be recognised and appreciated.

There are major differences between statistical modelling and classic statistical inference. The former uses standard errors and confidence intervals as means of indicating the lack of precision of parameter estimates, and only rarely are they used for formal tests of *a priori* hypotheses. The emphasis, therefore, is away from 'statistical significance' in favour of trying to quantify effects. A standard error and/or a confidence interval should always be quoted for every parameter estimated. Furthermore, maximum likelihood theory enables calculation of the asymptotic variance-covariance matrix, whereby an understanding of the strength of confounding between effects can be assessed through examination of the correlations between parameters. Unfortunately, only on rare occasions is this important information presented in twin analysis publications, despite the fact that not assessing this information can have serious consequences for modelling, as demonstrated in Example 1 below.

The underlying assumptions behind models need to be detailed carefully, and to be appreciated. There are biological assumptions which are manifest in subsequent statistical assumptions, and there are statistical assumptions introduced out of convenience or tractability. Deviations from any of these could have a substantial influence on the conclusions of modelling. The 'sensitivity' of twin or pedigree models has only rarely been discussed.

One contentious issue revolves around the possible existence of non-additive genetic effects, as represented by the dominance component of variance (as distinct from dominant inheritance, which refers to binary traits). When R.A. Fisher derived the genetic and environmental decomposition of variance in his classic 1918 paper [12], he did not consider an environmental component common to relatives. This is no excuse, however, for future generations to have ignored this potentially important source of variation. As discussed above, in twin modelling it has been usual to treat it in a convenient manner. Typically, a common twin environment component is assumed to be the same for monozygotic as for dizygotic pairs and a constant independent of age, sex, cohabitational status, and so on. Until recently, it had not been explored by models or modellers. When this was done, interesting results appeared.

The appropriateness of a model's description of the data should be tested from a variety of perspectives, not just by a single test of "goodness of fit" with weak power, as has been the case in much modelling of twin data. Determination of the adequacy of fit of *all* reasonable models to the available data has been advocated [42]. In addition, descriptive measures should be derived which would support the conclusions of model fitting. As is shown by Example 2 in the section on Heritability below, this can reveal false conclusions from an otherwise naive interpretation of model fits.

Modelling has limitations. By its very nature it attempts to describe Nature in the most parsimonious manner, using as few assumptions and elements as can be discriminated from one another with the available data. The likelihood ratio test is often used to determine if more or less parameters are required in a definitive statistical model. The irony of this is that the bigger the data set, the more parameters will be needed to 'adequately' describe it. Simpler models based on smaller data sets are more likely not to be rejected by goodness-of-fit tests, which are really "badness-of-fit" tests. Therefore, provided one collects and analyses small data sets, there is little danger that model fits

will transgress the standard tests of fit. Consequently, there will be little reason to suspect the appropriateness of the parsimonious description. As discussed above, this is not good science according to the Popperian paradigm.

Finally it must be understood that fitting a model is not an end in itself. Selection of a 'best' model from a range of alternatives does not *prove* that the components of that model are *true* causes of variation, let alone the *only* ones.

Example 1:

Two common problems in the interpretation of statistical modelling of twin data are illustrated in the reported analysis of the body mass index (BMI) of twins who have been reared apart [40]. The paper presented evidence that genetic factors play a role in determining BMI by noting that the correlation between twin pairs was similar whether they were reared apart or together. Although the crude correlations and size of sample were tabled for MZ and DZ, male and female pairs reared together and apart, unfortunately neither standard errors nor confidence intervals were presented. Simple calculations of these, however, are revealing.

First, it is claimed that "nonadditive genetic variance made a *significant* contribution to the estimates of heritability, particularly among men". This statement is incorrect and appears to be due to the authors having been misled by their modelling. They claimed support for this result by stating that "the intra-pair correlations of monozygotic twins were more than twice those of dizygotic twins", yet did not examine the evidence carefully.

For men, the authors report in their Table 1 that $r_{MZA}=0.70$ ($n=49$), $r_{DZA}=0.15$ ($n=75$), $r_{MZT}=0.74$ ($n=66$), and $r_{DZT}=0.33$ ($n=89$), where T and A refer to reared together and apart, respectively. The results from model fitting in their Table 2 were: $\sigma_a^2=0.34\pm0.30$, $\sigma_d^2=0.62\pm0.16$, and $\sigma_e^2=0.42\pm0.03$. The significant σ_d^2 term (the estimate almost four times the standard error) appears to have been the basis for the above statement concerning the significant non-additive genes. Now these estimates of the variance components imply that for the pooled data $r_{MZ}=0.96/1.38=0.70$ and $r_{DZ}=0.325/1.38=0.24$. In this case, there were 164 DZM pairs in total, so based on the variance of an estimate of a true correlation ϱ being approximately $(1-\varrho^2)^2/(n-1)$, the standard error of r_{DZ}, s.e.(r_{DZ}), will be at least 0.07, and s.e.(r_{MZ}) at least 0.04.

Consider the natural test statistic $T=r_{MZ}-2r_{DZ}$, which takes the observed value $0.70-2\times0.24=0.22$. Now s.e.$(T)=\{\text{s.e.}(r_{MZ})^2+4\ \text{s.e.}(r_{DZ})^2\}^{1/2}>0.14$. Therefore the probability of rejecting the null hypothesis that there is no non-additive genetic variation, given the null hypothesis is true, will be greater than 0.05, even if the alternate hypothesis is one-sided, specifying that $\sigma_d^2>0$. This is in sharp contrast to the implication of the fitted model above which would suggest $p<0.001$ for this test. What has happened?

The answer lies in the authors' own words; they noted in their Statistical Methods section that estimates of variance components are not independent. Note that although the estimate of σ_d^2 appears to be significant, that of σ_a^2 is not. What is needed is knowledge about the change in log likelihood between fitting the model with $\sigma_d^2=0$ and the fitted model above; the argument in the paragraph above suggests the change would have been about 1, and not judged statistically significant by the likelihood ratio test.

This raises the question: can there be non-additive genetic variance without additive variance? For simplicity, suppose that at every loci involved in the trait, i, there are two alleles, a_i and A_i, for $i = 1,2,\ldots$. In this case the answer is yes, provided the mean trait value is the same for both homozygotes, a_ia_i, and A_iA_i, yet different from that for the heterozygote, A_ia_i. Is this biologically plausible? It certainly does not represent dominant or recessive genetic expression, and it would be of interest to know if this sort of expression has been observed, eg. in plant or animal data.

Second, it was concluded "that genetic influences on body-mass index are substantial, whereas the childhood environment has little or no influence". This is a misleading statement due to what is known as Type II error. That is, there is little statistical power to detect variation from the additive genetic model, as can be seen from the standard error of T. The study had less than a 50:50 chance of detecting a common environment effect even if it accounted for over 25% of variation, so it is an overstatement to conclude that the effect was "little" or non-existent.

VARIATION - ABOUT WHAT, AND DOES IT MATTER?

There is almost no discussion about the trait mean in the major text books, even excellent texts such as Falconer [11] and Bulmer [5]. The impression is given that the interest in pedigree analysis is only in the second order moments, and ratios of them, with the mean treated as if it is an immutable constant, μ. Variation cannot, however, be discussed without specifying the mean, or 'expected' value, about which the variation occurs. It is usual to express the expected value in terms of measured covariates, like age and sex, which are called fixed effects so as to distinguish them from the random effects of *un*measured covariates. Other *measured* covariates may influence trait mean, and they also could be familial. The correlation or covariation between related individuals in these *residuals* forms the basis of twin and pedigree analysis, so the interpretation depends on whatever factors have been used to model the mean. This is often not made explicit.

Adjusting for a familial covariate can have considerable influence on trait correlations. Let Y_1 and Y_2 be trait values of individuals 1 and 2, and suppose

$$E(Y_i|X_i = x_i) = \alpha_0 + \alpha_1 x_i \text{ for } i = 1,2,$$

$$\varrho = \mathrm{Corr}(Y_i, X_i) = \alpha_1 \sigma_X / \sigma_Y,$$

$$\varrho_Y = \mathrm{Corr}(Y_1, Y_2),$$

and suppose that

$$\varrho_X = \mathrm{Corr}(X_1, X_2)$$

is not necessarily zero. For example, if i and j are twins and X = age, $\varrho_X = 1$. The partial or adjusted correlation between Y_1 and Y_2, adjusting for the linear relationship with the covariate X, can be shown [29] to be

$$\varrho' = \mathrm{Corr}(Y_1, Y_2|X_1, X_2) = \varrho_Y + (\varrho_Y - \varrho_X)\varrho^2/(1-\varrho^2),$$

https://doi.org/10.1017/S0001566000002129 Published online by Cambridge University Press

provided $Cov(Y_i, X_j | X_i) = 0$ for $i \neq j$. Therefore it is straightforward to see that, if $\varrho_X = \varrho_Y$, the correlation is not changed by adjustment for covariate, while if $\varrho_X > \varrho_Y$, adjustment results in a lower correlation and if $\varrho_X > \varrho_Y$, adjustment results in a higher correlation.

Consider $\varrho_X = 1$; the correlation will always be decreased by adjustment for a covariate (like age in twin pairs) that is perfectly correlated for both individuals. If the unadjusted correlation, ϱ_Y, is high, adjustment for the covariate will have a small effect. If ϱ_Y is low, adjustment for even a weak association can have a big influence, especially in proportional terms. For example, if $\varrho_Y = 0.8$, as ϱ increases from 0 to 0.4, ϱ' decreases slightly from 0.80 to 0.76, while if $\varrho_Y = 0.2$, ϱ' decreases substantially from 0.20 to 0.06 and can become negative if ϱ increases further.

Consider $\varrho_X = 0$; the absolute value of the correlation between Y_1 and Y_2 is always increased by adjustment for a covariate that is *uncorrelated* between individuals. This effect is greatest when the unadjusted correlation, ϱ_Y, is high, and is small when ϱ_Y is low. For example, if $\varrho_Y = 0.8$, as ϱ increases from 0 to 0.4, ϱ' increases from 0.80 to 0.94, while if $\varrho_Y = 0.2$, ϱ' increases from 0.20 to 0.24.

HERITABILITY

Heritability has been defined as the ratio of the genetic component of variance to the total variance, expressed as proportion or as a percentage. It is akin to the epidemiologist's odds ratio, which is a ratio of two rates. It is very tempting to compare odds ratios and heritability estimates across studies. Both these measures, however, are functions of both the underlying disease process and the population under study. This has consequences for 'meta analysis', or 'overview', whereby attempts are made to combine estimates from a number of studies in a statistically valid way. This is not a process to be undertaken lightly, and not just because studies often vary in numerous methodological ways. A fundamental problem revolves around the question of why should the odds ratio or heritability be a constant; ie. a feature of the disease which is independent of the population, age, sex and other factors?

R.A. Fisher commented on heritability in an obscure 1951 publication [13], pointing out that whereas the genetic component of variance "... has a simple genetic meaning", the total variance "... includes errors of measurement, both controllable and uncontrollable, as well as the genetic variance". He concluded that "... information contained in the genetic component of variance is largely jettisoned when its actual value is forgotten, and it is only reported as a ratio of this *hotch potch* of a denominator".

If covariances between relatives and estimates of genetic components of variance are published, there is the possibility of making comparisons or poolings across studies. Consider for example a study of blood pressure in migrants to Melbourne, Australia, from the Greek island of Levkada, and of their sibling and other relatives who remained in Greece [36]. Analysis [24] revealed that the covariances between first-degree relatives were similar whether they lived in Melbourne or Levkada and were greater than the significant covariance between higher degree relatives. The total variance, however, was about 30% higher in Melbourne males than the total variance in Levkadan males and

in Melbourne females. Modelling gave a significant additive genetic component of variance, no detectable common environment or non-additive genetic variance, and an environmental component specific to individuals which was greater in Melbourne males. Further analysis showed that, within the power limitations of this study of 1,400 individuals, the genetic component of variance was no different between males and females, and more important, between Melburnian and Levkadan Greek-born relatives. That is, the genetic component of variance had been transported to a new environment upon migration where, at least among the males, different "uncontrollable" factors have acted to increase the total variance.

Example 2:

Analysis of bone mineral density measurements taken from 124 MZ and 47 DZ female twin pairs aged 25 to 80, adjusting for height, age and supposedly "environmental" covariates (smoking, calcium intake, exercise, alcohol, caffeine intake, childbearing patterns) led the authors to claim that "... not only gene interactions may exist, but also relatively few genes may be involved in the inheritance of bone mass" [39]. This was based on their contention that the correlation between MZ pairs, r_{MZ}, was greater than the estimate of "heritability", given by $H^2 = 2(r_{MZ} - r_{DZ})$, where r_{DZ} was the correlation between DZ pairs. Unfortunately, they failed to consider the role of chance in their results. Consider data from the lumbar spine, where $r_{MZ} = 0.80$ and $r_{DZ} = 0.19$. Therefore the observed value of H^2 was 1.22. As in Example 1, it can be shown that $s.e.(r_{DZ}) > 0.15$, therefore $s.e.(H^2) > 2 \times s.e.(r_{DZ}) > 0.30$, so clearly, the estimate of r_{MZ} is *not* significantly greater than the estimate of H^2. This test is further complicated by the fact that the two estimates are not independent. Furthermore, the s.e. of r_{DZ} is of such a magnitude that it does not appear that DZ pairs are significantly correlated; ie. this data casts doubt that genetic factors exist, let alone that there are major genes which exhibit interactions.

This example illustrates the importance of studying large numbers of DZ pairs if the correlation between these pairs is small. While r_{MZ} gives an assessment of the maximum proportion of variance that can be explained by genetic factors, this information is of little value if it cannot be shown that DZ pairs exhibit a significant correlation. The need to establish familial aggregation as an *a priori* condition for a role of genetic factors is often neglected by twin researchers in their enthusiasm to fit models.

THE FUTURE

Advances in DNA technology promise to revolutionise epidemiology [31]. Linkage studies identify putative genetic markers in highly selected and atypical pedigrees. There is therefore a need to establish studies which will test hypotheses concerning these markers in the *population*, and case-control-pedigree designs may become standard tools of epidemiology. That is, genetic epidemiology should provide a link between biomedical science and the public health consequences of its findings. It will also allow hypotheses concerning interactions between exposures and genetic susceptibility to be tested, using standard and well-established epidemiological methods.

An example of a case-control-family study which utilises this approach is a study of

breast cancer in Victorian families [25]. Cases are selected at random with strata from the Victorian Cancer Registry, which collects all cases in the state. Controls are from a random population sample or are the spouses or partners of breast cancer probands. All first-degree relatives and grandparents on both the paternal and maternal side, for both the proband and control, are in the study. A sequential sampling scheme is used to study first-degree relatives of any previously studied relative with breast cancer. Proxy information is collected for deceased subjects. Reported cancer cases in relatives are verified. Risk factor information is collected from all relatives by interview. Procedures have been developed to produce a proband response rate of over 80% and good cooperation from relatives. Blood samples are being collected for future testing of hypotheses related to specific markers and candidate DNA probes for breast cancer in this population-based sample of families. Twins are identified among probands and relatives. It is intended to apply this design to other populations and to family studies of other cancers.

A major contribution of twin research to science has come from studies of pairs discordant for a disease or for certain exposures or risk factors (eg. [3,14,26,28], and from studies of gene-environment interaction through applying interventions to MZ pairs [4]. These will become of even greater importance when putative genetic markers for specific diseases are identified.

The full potential of twins in epidemiology will only be realised when international collaborations occur, which will enable scientific questions to be addressed with substantial statistical power. International collaborations have become an integral part of epidemiology, and it is hoped that this will also be the case for genetic epidemiology and twin research in the near future.

Acknowledgements: This work was supported by the Australian National Health and Medical Research Council (NHMRC) and the Australian NHMRC Twin Registry. I would like to thank colleagues for their contribution, including John Mathews, Robert Elston, Ken Lange, Graham Giles, John Carlin, Murray Hannah and Greg Macaskill.

REFERENCES

1. Aalen OO (1991): Modelling the influence of risk factors on familial aggregation of disease. Biometrics 47:933-945.
2. Australian NHMRC Twin Registry (1991): 1990 Annual Report. Australian NHMRC Twin Registry, Melbourne.
3. Battie MC, Videman T, Gill K, Moneta GB, Nyman R, Kaprio J, Koskenvuo M (1991): Smoking and lumbar invertebral disc degeneration: an MRI study of identical twins. Spine 16:1015-1021.
4. Bouchard C, Tremblay A, Despres J-P, Nadeau A, et al. (1990): The response of long-term overfeeding in identical twins. New Engl J Med 322:1477-1482.
5. Bulmer MG (1980): The Mathematical Theory of Quantitative Genetics. Oxford: Clarendon Press.
6. Cavalli-Sforza LL, Feldman MW (1981): Cultural Transmission and Evolution: A Quantitative Approach. Princeton NJ: Princeton University Press.
7. Clifford CA, Hopper JL, Fulker D, Murray RM (1984): A genetic and environmental analysis of a twin family study of alcohol use, anxiety, and depression. Genet Epidemiol 1:63-79.

8. Day NE, Simons MJ (1976): Disease susceptibility genes - their identification by multiple case family studies. Tissue Antigens 8:109-119.
9. Doherty JDH, Lancaster PAL (1986): The secular trend of twinning in Australia, 1853-1982. Acta Genet Med Gemellol 35:61-76.
10. Eaves LJ, Eysenck HJ, Martin NG (1989): Genes, Culture and Personality: An Empirical Approach. London: Academic Press.
11. Falconer DS (1989): Introduction to Quantitative Genetics, (3rd ed.), Harlow: Longman.
12. Fisher RA (1918): The correlation between relatives on the supposition of Mendelian inheritance. Trans Roy Soc Edinb 52:399-433.
13. Fisher RA (1951): Limits to intensive production in animals. Brit Agric Bull 4:217-218.
14. Floderus B, Cederlof R, Friberg L (1988): Smoking and mortality: A 21-year follow-up based on the Swedish Twin Registry. Int J Epidemiol 17:332-340.
15. Greenland S (1987): Evolution of Epidemiologic Ideas: Annotated Readings on Concepts and Methods. Chestnut Hill, Ma., Epidemiology Resources.
16. Gibson HB, Silverstone H, Gandevia B, Hall GJL (1969): Respiratory disorders in seven-year-old children in Tasmania. Aims, methods and administration of the survey. Med J Aust 2:201-205.
17. Giles GG, Gibson HB, Lickiss N, Shaw K (1984): Respiratory systems in Tasmanian adolescents: a follow-up of the 1961 birth cohort. Aust NZ J Med 14:631-637.
18. Hay DA, Clifford CA, Derrick P, Hopper J, Renard B, Theobald TM (1990): Twin children in volunteer registries: biases in parental participation and reporting. Acta Genet Med Gemellol 39:71-84.
19. Hopper JL, Mathews JD (1982): Extensions to multivariate normal models for pedigree analysis. Ann Hum Genet 465:373-383.
20. Hopper JL, Mathews JD (1983): Extensions to multivariate normal models for pedigree analysis. II. Modeling the effect of shared environment in the analysis of variation in blood lead levels. Am J Epidemiol 117:344-355.
21. Hopper JL (1988): A review of FISHER. Genet Epidemiol 5:473-475.
22. Hopper JL, Culross P (1983): Covariation between family members as a function of cohabitational history. Behav Genet 13:459-471.
23. Hopper JL, Carlin JB (1992): Familial aggregation of a disease consequent upon correlation between relatives in a risk factor measured on a continuous scale. Am J Epidemiol (in press).
24. Hopper JL, Macaskill GT, Powles JP, Ktenas D (1992): Pedigree analysis of blood pressure in subjects from rural Greece and relatives who migrated to Melbourne, Australia. Genetic Epidemiology, 1992: 8 (in press).
25. Hopper JL, Giles GG (1992): A study of breast cancer in Victorian families. Cold Spring Harbor Conference on Cancer and Cell Biology (Abstract).
26. Hopper JL, Seeman E (1992): Tobacco use and reduced bone density: a co-twin control study. In Christiansen C, Overgaard K (eds) Osteoporosis 1990, Vol. 2. Copenhagen: Osteopress ApS, pp 1117-1118.
27. Jenkins M, Dalton M, Hopper JL, Carlin JB, Flander LB, Giles GG (1992): Follow up in 1991-92 of the 1968 Tasmanian Asthma Study. Faculty of Medicine Epidemiology Unit Report, The University of Melbourne.
28. Kaprio J, Koskenvuo M (1990): Cigarette smoking as a cause of lung cancer and coronary heart disease. A study of smoking-discordant twin pairs. Acta Genet Med Gemellol 39:25-34.
29. Kendall MG, Stuart A (1973): The Advanced Theory of Statistics, Vol. 2 (3rd ed.), Chapter 27. London: Griffin.
30. Khoury MJ, Beaty TH, Liang K-Y (1988): Can familial aggregation of disease be explained by familial aggregation of environmental risk factors? Am J Epidemiol 127:674-683.
31. Khoury MJ, Beaty TH, Flanders WD (1990): Epidemiologic approaches to the use of DNA markers in the search for disease susceptibility genes. Epidemiologic Reviews 12:41-55.
32. Last JM (1983): A Dictionary of Epidemiology. Oxford: Oxford University Press.
33. Morton NE (1982): Outline of Genetic Epidemiology. Basel: Karger.
34. Peto J (1980): Genetic predisposition to cancer. In Cairns J, Lyon JL, Skolnick M (eds): Banbury Report 4: Cancer incidence in defined populations. Cold Spring Harbor Laboratory, pp 203-213.
35. Popper KR (1968): The Logic of Scientific Discovery. (2nd ed.) New York: Harper & Row. Originally published as Logik der Forschung. Vienna: Springer, 1934.

36. Powles JP, Hopper JL, Macaskill GT, Ktenas D (1992): Blood pressure in subjects from rural Greece, comparing individuals migrating to Melbourne, Australia, with non-migrant relatives. J Hum Hypert (in press).
37. Rose RJ, Kaaprio J, Williams CJ et al (1990): Sibling contact and sibiling similarity: facts, issues, and red herrings. Behav Genet 20:763-778.
38. Rothman K (1988): Causal Inference. Chestnut Hill, Ma., Epidemiology Resources.
39. Slemenda CW, Christian JC, Williams CJ, et al (1991): Genetic determinants of bone mass in adult women: A reevaluation of the twin model and the potential importance of gene interaction on heritability estimates. Bone Min Res 6:561-567.
40. Stunkard AJ, Harris JR, Pedersen NL, McClearn GE(1990): The body mass index of twins who have been reared apart. New Engl J Med 322:1483-1487.
41. Williams CJ, Viken R, Rose RJ (1992): Likelihood-based analyses of longitudinal twin and family data: experiences with pedigree-based approaches. Behav Genet 22:215-224.
42. Wilson SR (1982): Heritability. J Appl Prob 19A:71-85.

Correspondence: Dr John L. Hopper, The University of Melbourne, Faculty of Medicine Epidemiology Unit, 151 Barry Street, Carlton, Victoria 3053, Australia.