J Clin Epidemiol Vol. 44, No. 11, pp. 1223–1231, 1991
Printed in Great Britain.

0895-4356/91 $3.00 + 0.00

HOW INDEPENDENT ARE "INDEPENDENT" EFFECTS? RELATIVE RISK ESTIMATION WHEN CORRELATED EXPOSURES ARE MEASURED IMPRECISELY

ANDREW N. PHILLIPS[1]* and GEORGE DAVEY SMITH[2]

[1]Department of Public Health and Primary Care, Royal Free Hospital School of Medicine, Rowland Hill Street, London NW3 2PF and [2]Department of Epidemiology and Population Sciences, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, England

(Received in revised form 3 April 1991)

Abstract—A relative risk estimate which relates an exposure to risk of disease will tend to be estimated too close to unity if that exposure is subject to random measurement error or intra-subject variability. "Independent" relative risk estimates, for the effect of one exposure after adjusting for confounding exposures, may be biased in either direction, depending on the amount of measurement imprecision in the exposure of interest and in the confounders. We describe two methods which estimate the bias in multivariate relative risk estimates due to the effect of measurement imprecision in one or more of the exposure variables in the model. Results from the two methods are compared in an example involving HDL cholesterol, triglycerides and coronary heart disease. In this example, the degree of bias in relative risk estimates is shown to be highly dependent on the amount of measurement imprecision ascribed to the exposures. It is concluded that when two exposures are substantially correlated, and one or both is subject to sizeable measurement imprecision, a study in which exposures are measured only once will be inadequate for investigating the independent effect of the exposures. Where feasible, epidemiologists should seek study populations where the correlation between the exposures is smaller.

Errors-in-variables Measurement imprecision Adjusted relative risk estimates
Correlated exposures Independent effects

INTRODUCTION

Both intra-subject variability and technical measurement error of exposures lead to distortion of relative risk estimates in epidemiological analyses. If this variability/error is random and independent of the outcome (e.g. death) then the relative risk estimate associated with a given exposure will tend to be estimated too close to unity. Increased recognition of this has resulted in proposals regarding how to adjust parameter estimates to account for it [1–4]. In one method [2, 3] for adjusting logistic regression coefficients the regression coefficient of the "true" exposure values on the measured values is calculated (λ), and the logistic regression coefficient between the measured exposure and the outcome is then divided by λ. In general, these procedures require validation studies in which the relationship between actual measurements of an exposure and their "true" underlying values can be ascertained. However, if one makes the assumption that the measure of exposure is unbiased (i.e. the variability/error is random)

*All correspondence should be addressed to: Dr A. N. Phillips, MRC U.K. Centres for Coordinating Epidemiological Research on AIDS and HIV, Academic Department of Genito-Urinary Medicine, University College & Middlesex School of Medicine, James Pringle House, 73–75 Charlotte Street, London W1N 8AA, England.

then only estimates of measurement repeatability are required, not the relationship between measured and "true" values [3].

This describes the situation where one is considering the univariate association between an exposure which is measured imprecisely (we shall use this term from now on to refer to the random technical measurement error and/or intra-subject variability associated with an exposure) and risk of the outcome. Different degrees of measurement imprecision in correlated exposure variables will have complex consequences for the strengths (and even for the existence) of the apparent independent effects associated with the exposures in multivariate models. This has important implications for the study of diseases believed to be multifactoral in origin, since the main aim of many such investigations is to assess the importance of putative causal agents after taking account of already established risk factors.

In a 1959 article reviewing the evidence that smoking causes lung cancer, Cornfield and colleagues [5] addressed the suggestion that the association between the two was due to confounding by a factor related to both, such as urban residence. They stated that this could not be the case since the relative risk attached to smoking was greater than that attached to any of the potential confounding factors. This statement has been extended into the general rule that the relationship between an exposure and an outcome cannot be due to confounding by a factor more weakly associated with the outcome [6–9].

If the exposure measures were without technical error and took account of intra-individual variability then this rule is necessarily true. This is rarely what actually pertains when it is applied, however. In situations with measurement imprecision in exposures it could produce misleading conclusions. For a badly measured or highly variable cause a strongly correlated factor allowing stable and precise measurement could serve as a better proxy measure of exposure to the cause than its ostensible direct measurement. In this situation the non-causal factor would have a higher associated relative risk estimate than the actual cause. Clearly even in the absence of this extreme situation greater error or variability in measurement of the causal factor than of its non-causal correlate could lead to an apparent association between the latter and the outcome after adjustment for the former. This residual confounding would produce what appear to be independent effects associated with factors which actually have no such role [10].

It has been shown that random error and variability in the measurement of exposures in multiple logistic models can lead to either increased or decreased parameter estimates [11, 12]. As yet there has been inadequate attention paid to this problem, with most, but not all, of what has been written focusing upon the effect of misclassification of categorical variables in case–control studies [11–17]. Estimation problems due to strong correlations between continuous exposures (collinearity) in multiple logistic models have been studied, but the possible additional effect of measurement imprecision in those correlated exposures has, with some exceptions [16, 18], been ignored in this context. The estimation problem caused by collinearity in exposures becomes more acute if those exposures are measured imprecisely.

In this paper we describe two methods which provide multivariate relative risk estimates which have been "corrected" for the effect of imprecision in one or more of the exposures in the model. However, as we shall discuss further, such correction methods should not be seen as a panacea and these results are derived in order to approximately assess the extent to which multivariate relative risk estimates can be biased due to measurement imprecision of exposures, rather than as an encouragement for attempts to "correct" relative risk estimates in this way.

ASSESSMENT OF THE BIAS IN MULTIVARIATE RELATIVE RISK ESTIMATES DUE TO MEASUREMENT IMPRECISION IN EXPOSURES

In this section we attempt to quantify the amount by which a multivariate relative risk estimate (we refer to relative risk throughout this text although, strictly, it is the relative odds that is under consideration) may be biased due to failure to account for the fact that one or more of the exposures in the model may be measured imprecisely. As an example, we consider the effect of two serum lipids, HDL cholesterol and triglycerides on risk of coronary heart disease. Using a simulation method, we estimate the extent of this bias and give the "corrected" relative risk estimates. Another approach to the "correction" of relative risks has been described by Rosner and colleagues [3, 16]. This method is briefly described and comparison is made between the results obtained from the two methods.

Consider a situation where we wish to measure the effect of some risk exposure on a binary disease outcome after adjustment for confounding exposures using a multiple logistic model. Let us assume that the underlying "usual", or long-term average, level is the aspect of the exposure which is most important in determining risk of the disease (as opposed, for example, to the peak level reached over a period of time). Let us call these "usual" exposures $X_1 \ldots X_n$. Due to biological variability and/or measurement error the X_i $(i = 1, n)$ are each measured with some amount of random error, as Z_i. We assume that the random error associated with the X_i is uncorrelated. Without loss of generality, let us further assume that the X_i and Z_i have been transformed such that they have mean zero and standard deviation 1.

Let $\mathbf{X}$ and $\mathbf{Z}$ be $(n \times 1)$ vectors of the X_i and Z_i, respectively. For our multiple logistic model we have

$$\ln \frac{p'}{1-p'} = \alpha' + \boldsymbol{\beta}' \cdot \mathbf{X}$$

where p' is the estimated probability of disease and $\boldsymbol{\beta}'$ is the $(1 \times n)$ vector of "correct" logistic coefficients and

$$\ln \frac{p}{1-p} = \alpha + \boldsymbol{\beta} \cdot \mathbf{Z}$$

where p is the estimated probability of disease and $\boldsymbol{\beta}$ is the $(1 \times n)$ vector of naively estimated logistic coefficients.

Note that $\mathbf{X}$ is unknown, so $\boldsymbol{\beta}'$, the parameter we wish to know, cannot be estimated directly. Now consider the multivariate linear regression of $\mathbf{X}$ on $\mathbf{Z}$

$$\mathbf{X} = \lambda \cdot \mathbf{Z} + \mathbf{e}$$

where λ is an $(n \times n)$ matrix of regression coefficients and $\mathbf{e}$ an $(n \times 1)$ vector of error terms, is distributed as multivariate normal with mean 0. The intercept is zero since the X_i and Z_i have mean zero.

Rosner *et al.* suggest that the "corrected" logistic coefficients $\boldsymbol{\beta}'$ can be estimated by $\hat{\boldsymbol{\beta}}'$ where

$$\hat{\beta}' = \beta \cdot \lambda^{-1}$$

Given the assumption of multivariate normality, this method can only be applied to situations where there is measurement imprecision in (suitably transformed) continuous exposures, and not where there is misclassification of categorical risk factors. Another assumption made is that the probability of the disease outcome is small.

Simulation procedure

The simulation procedure has been carried out for the situation where we have "usual" exposures X_1 and X_2, which are measured as Z_1 and Z_2. As has been discussed, the measurement imprecision means that the association between Z_1 and the disease outcome will tend to be weaker than that between X_1 and the outcome. Similarly for Z_2 and X_2. In addition, the correlation between Z_1 and Z_2 will be smaller than that between X_1 and X_2. The simulation procedure is a method for adjusting Z_1 and Z_2 such that they have stronger associations with the disease outcome and a stronger association with each other. The adjustments of Z_1 and Z_2 result in variables X_1^* and X_2^* which mimic X_1 and X_2. They mimic X_1 and X_2 in that when the appropriate degree of measurement imprecision (i.e. random error) is added to them the resulting variables, Z_1^* and Z_2^*, have the same correlation with each other and, respectively, the same association with the disease outcome as Z_1 and Z_2. Once these mimics X_1^* and X_2^* are created, a logistic model can be fitted on these variables to obtain estimates of the "corrected" logistic coefficients.

We specify the measurement imprecision in terms of the Pearson correlation coefficient between the "usual" and the measured exposures. Let

$$r_1 = \text{corrn } (X_1, Z_1) \text{ and } r_2 = \text{corrn } (X_2, Z_2)$$

We assume that these values are known, or can be estimated. These r_i can be estimated as the square root of the intra-class correlation coefficient between replicate measures. When there are two replicates only, the intra-class correlation coefficient is equivalent to calculating the usual Pearson correlation coefficient between two measurements, provided each pair of measurements is counted twice, the second time in reverse order. Note that the correlation coefficients r_i indicate the measurement imprecision of an exposure and should not be confused with the correlation between two exposures.

Thus, in summary, the method is as follows. Given the measurement imprecision γ_i to be attributed to the exposures (where γ_i is normal with mean zero and variance such that corrn $(X_i^*, Z_i^*) = r_i$, $i = 1,2$) we adjust the Z_i to create X_i^* (mimics of X_i) such that

$$X_i^* + \gamma_i = Z_i^* \qquad (1)$$

where Z_i^* have the following properties

$$\text{corrn}\ (Z_1^*, Z_2^*) = \text{corrn}\ (Z_1, Z_2)$$

and

$$\beta_i = \beta_i^*$$

where β_i and β_i^* are multiple logistic coefficient estimates for Z_i and Z_i^*, respectively. Once the X_i^* have been created we can fit a logistic model on them to obtain our "corrected" logistic coefficient estimates.

Creation of X_i^* from Z_i

Z_1 was adjusted, to varying degrees, in three ways.

(i) to increase the association between Z_1 and the outcome: the level of Z_1 in all individuals who are cases was raised by a fixed amount and the level of Z_1 in individuals who are non-cases was lowered by the same amount (or vice versa if the association between Z_1 and disease outcome is negative).

(ii) to increase the correlation between Z_1 and Z_2: the value of Z_1 was moved closer to the value of Z_2 (or further away if correlation is negative). This was achieved by fitting the regression of Z_1 on Z_2, giving slope b. Each individual's value of Z_1 was then moved proportionately closer to their value of $b \cdot Z_2$.

(iii) random error (normally distributed) was added.

This process was carried out simultaneously for both Z_1 and Z_2 in order to create new variables X_1^* and X_2^*. An iterative process was used in order to reach a point of convergence. At every step variables were re-transformed to have a mean zero and variance of one. The point of convergence was the point where the parameters for (i), (ii) and (iii) were such that equation (1) was satisfied (in expectation). Since random error is added to create X_1^* and X_2^*, it is only possible to define a set of parameters such that on average, over multiple simulations, the X_1^* and X_2^* satisfy equation (1); i.e. the mean of the β_i equals the mean of the β_i^*. Similarly for the correlation between the Z_i^* and the Z_i. Therefore, for any given set of parameters, variables X_1^* and X_2^* were created 50 times in 50 separate simulations (and for each X_1^* and X_2^* a single corresponding Z_1^* and Z_2^*). If, on average over the 50 simulations, equation (1) is satisfied, then the means of the multiple logistic coefficient estimates for the X_1^* and X_2^* are taken as the "corrected" logistic coefficient estimates.

When multiple simulations are carried out with the same parameters then the correlation coefficient between Z_1^* and Z_2^* (and that between X_1^* and X_2^*) varies only to a negligible degree (i.e. error range of around 0.02). However, the logistic coefficient estimates vary quite substantially from simulation to simulation, due to the random component of the process. It is only for the logistic coefficient estimates that the random component causes any significant variation between simulations, due to the fact that these estimates are much less stable, being based on the number of cases of disease as opposed to the total number of subjects. The greater the measured correlation between the Z_i the greater the problem of collinearity in the estimation of the logistic coefficients for the Z_i and, especially, the X_i^* (because the collinearity is greater). In some cases more than 50 simulations would be required to reach a stable estimate for the β_i^* values.

Example—HDL cholesterol, triglycerides and heart attacks

One area in which measurement imprecision in confounded exposures arises is in the determination of the relative importance of triglycerides (TG) and high density lipoprotein cholesterol (HDLC) in the genesis of coronary heart disease (CHD). It has been postulated that high TG levels increase susceptibility to CHD whereas high HDLC levels offer protection. Measures of TG and HDLC are strongly inversely correlated [19]. In a review in 1980, Hulley *et al.* [20] considered the independence of the association between TG and risk of CHD reported by some studies to be due to the lack of inclusion of HDLC in the analyses. They demonstrated that the relative risk associated with TG fell markedly after adjustment for HDLC in a logistic model. A reference to a discussion by Cornfield [21] of attribution of cause was included to support the conclusion. This interpretation is now generally accepted [22]. The inter-relationships of TG, HDLC and CHD are used in the following example. There has been some dissent to the prevailing view regarding the relative lack of importance of TG [23], with the high variability of TG [24, 25] and the collinearity between TG and HDLC [26] being mentioned as potential sources of difficulty in establishing any independent effect associated with TG. We stress, however, that we are not here attempting to make any specific claims about the correct solution to this problem—we are merely using

this relatively well known example for the illustration of a general point.

The data for this example come from the British Regional Heart Study, a prospective study of risk factors for coronary heart disease in 7735 men aged 40–59 [27, 28]. 316 of the men suffered a heart attack during the follow-up period considered here. Non-fasting TG and HDLC were measured at screening. TG has been log (natural) transformed due to its skew distribution. The correlation coefficient between HDL and TG was -0.44. For the purposes of this example, we have taken the correlation coefficient between the "usual" TG level and a single measure of TG as 0.6. For HDL we have taken a value of 0.9. These values were chosen on the basis of previous work on the relative variability of measures of the two lipids [29]. The coefficients chosen would imply that the correlation between "usual" TG and "usual" HDL must therefore be substantially higher than the value of -0.44 found between HDL as measured and TG as measured. For clarity, we shall continue to use the terms X_i, Z_i, X_i^* and Z_i^*. For HDLC $i = 1$, for TG $i = 2$.

In Table 1 we show the univariate logistic coefficients for HDLC and for TG as well as the coefficients after adjustment for each other (all coefficients refer to a one standard deviation difference). Notice that the logistic coefficient estimate for TG has fallen markedly, from 0.26 to 0.12, after inclusion of HDLC in the model. Naive interpretation of these results would suggest that TG is a significantly less important risk factor for coronary heart disease than is HDLC since both the univariate and adjusted logistic coefficients for HDL cholesterol have a substantially greater magnitude than those for TG. Also in Table 1 are the results of 50 simulations as described above.

In each simulation the correlation coefficient between our mimics of "usual" TG and "usual" HDLC (X_1^* and X_2^*) was -0.81. This was the degree of correlation that was required such that when the appropriate amount of error was added (to create Z_1^* and Z_2^*) the correlation fell to -0.44. The values presented in Table 1 are means of the uncorrected logistic coefficients (i.e. coefficients for Z_1^* and Z_2^*) and means of the corrected logistic coefficients (i.e. coefficients for X_1^* and X_2^*, "usual" TG and "usual" HDL) over the 50 simulations. The uncorrected values are presented merely to show the degree to which the logistic coefficients for Z_1^* and Z_2^* agree with those for Z_1 and Z_2, as they should.

The "corrected" values for the univariate logistic coefficients are, as expected, further from zero than they were before correction, to a degree approximately proportional to the error in measurement of the respective variables. Thus the magnitude of the univariate coefficient for HDLC is slightly raised, from -0.39 to -0.44, and the magnitude of the univariate coefficient for TG is greatly raised, from 0.26 to 0.46. After adjustment for TG, the corrected estimate for HDLC is much closer to zero (about half the magnitude) than it was originally. This is because the effect of TG on risk and the association between TG and HDL cholesterol were previously underestimated. The resulting adjusted coefficient for TG is higher than it was previously, by almost 3-fold.

In order to compare these results with those from Rosner's method [16], we need to know the 2×2 matrix λ of regression coefficients of the "usual" HDL-C and "usual" TG on HDLC as measured and TG as measured (i.e. for X_1 and X_2 on Z_1 and Z_2). Let us assume that "usual" HDLC (X_1) and "usual" TG (X_2) are multivariate normal (with mean 0, SD 1). We can define X_1 and X_2 such that if error is added (normally distributed, mean 0, error in X_1 uncorrelated with error in X_2) to create Z_1 and Z_2—so that Z_1 is correlated with X_1 with $r = 0.9$, Z_2 is correlated with X_2 with $r = 0.6$ and that the correlation between Z_1 and Z_2 is -0.44—then (if Z_1 and Z_2 have SD = 1) the matrix is given by

$$\begin{pmatrix} 0.86 & -0.12 \\ -0.57 & 0.35 \end{pmatrix}$$

The values for the corrected logistic regression coefficients presented in Table 1 (row four) were obtained using this matrix and Rosner's formula. Overall the agreement is good. It should be noted that the assumption of bivariate normality holds reasonably well in this case.

In Tables 2 and 3, we show the situation for different amounts of measurement imprecision associated with HDLC and TG. In both cases, the two methods appear to give similar results. It is important to note from Table 2, where r_1 (HDL) is still taken as 0.9 but r_2 (triglycerides) is taken as 0.7 instead of 0.6, that there is a substantial difference in the "corrected" logistic coefficients for HDLC and for TG. This indicates that these "corrected" estimates are heavily dependent upon the degree of measurement imprecision in the exposures under consideration. Note that, in general, $r_1 r_2$ cannot be less than the correlation between Z_1 and Z_2 because

Table 1. Table of naively estimated logistic coefficient estimates for HDLC and TG (i.e. for Z_1 and Z_2) together with the mean (SD) logistic coefficient estimate over 50 simulations for X_1^* and X_2^* (i.e. "corrected" estimates) and for Z_1^* and Z_2^* where $r_1 = 0.9$, $r_2 = 0.6$. The average correlation between X_1^* and X_2^* was -0.81. Also given are "corrected" logistic coefficient estimates obtained by using Rosner's method. For comparison with the simulation procedure, these are obtained by multiplying the mean of the logistic coefficient estimates for the Z_1^* and Z_2^* by the matrix λ^{-1}

	Logistic coefficient estimates			
	TG		HDLC	
	Univariate	Adjusted for HDLC	Univariate	Adjusted for TG
From original data (Z_1 and Z_2)	0.26	0.12	−0.39	−0.32
Using simulation mean (SD) over 50 simulations				
Uncorrected	0.25	0.13	−0.38	−0.32
(Z_1^* and Z_2^*)	(0.05)	(0.06)	(0.04)	(0.05)
Corrected	0.46	0.35	−0.44	−0.15
(X_1^* and X_2^*)	(0.03)	(0.08)	(0.04)	(0.08)
Using Rosner's method	0.42	0.32	−0.42	−0.17

this would imply that the correlation between X_1 and X_2 were above one.

Comment on the two methods

We have assumed a model whereby exposures are measured with random error, due to measurement error/biological variability, and that those errors are uncorrelated between exposures. Rosner's method also deals with systematic (biased) error and allows for correlated errors. In this case, it is necessary to have data from a validation, not merely a replication, study. The other useful element in Rosner's method is that confidence limits have been derived for the "corrected" logistic coefficient estimates. These not only take into account the fact that exposures measured with error tend to have artefactually narrow confidence intervals but they also allow for the variation introduced due to the limited size of the validation study. Unlike Rosner's method, our simulation method is not dependent upon the distributions of the risk exposures in the model. Its disadvantage is that it can only be used when a limited number of exposures are being considered. However, it is rare that there are more than one or two strong confounders in any given situation. One further point should be noted about these two methods. If the multiple logistic model has been fitted for use in prediction of future cases of disease, then it would not be appropriate to apply the correction. If single measurements of the exposures are used to derive a predictive score from a logistic model then that score is optimal for use as a predictive score based on single exposure measurements on each individual.

DISCUSSION

In this paper we have demonstrated how differential degrees of measurement imprecision (whether due to error or biological variability) in exposures can distort the multivariate relative risk estimates attached to them. In this way,

Table 2. Table of naively estimated logistic coefficient estimates for HDLC and TG (i.e. for Z_1 and Z_2) together with the mean (SD) logistic coefficient estimate over 50 simulations for X_1^* and X_2^* (i.e. "corrected" estimates) and for Z_1^* and Z_2^* where $r_1 = 0.9$, $r_2 = 0.7$. The average correlation between X_1^* and X_2^* was -0.72. Also given are "corrected" logistic coefficient estimates obtained by using Rosner's method. For comparison with the simulation procedure, these are obtained by multiplying the mean of the logistic coefficient estimates for the Z_1^* and Z_2^* by the matrix λ^{-1}

	Logistic coefficient estimates			
	TG		HDLC	
	Univariate	Adjusted for HDLC	Univariate	Adjusted for TG
From original data (Z_1 and Z_2)	0.26	0.12	−0.39	−0.32
Using simulation mean (SD) over 50 simulations				
Uncorrected	0.25	0.12	−0.38	−0.33
(Z_1^* and Z_2^*)	(0.05)	(0.06)	(0.03)	(0.05)
Corrected	0.38	0.17	−0.43	−0.30
(X_1^* and X_2^*)	(0.03)	(0.06)	(0.04)	(0.06)
Using Rosner's method	0.36	0.16	−0.42	−0.31

Table 3. Table of naively estimated logistic coefficient estimates for HDLC and TG (i.e. for Z_1 and Z_2) together with the mean (SD) logistic coefficient estimate over 50 simulations for X_1^* and X_2^* (i.e. "corrected" estimates) and for Z_1^* and Z_2^* where $r_1 = 0.95$, $r_2 = 0.7$. The average correlation between X_1^* and X_2^* was -0.69. Also given are "corrected" logistic coefficient estimates obtained by using Rosner's method. For comparison with the simulation procedure, these are obtained by multiplying the mean of the logistic coefficient estimates for the Z_1^* and Z_2^* by the matrix λ^{-1}

	Logistic coefficient estimates			
	TG		HDLC	
	Univariate	Adjusted for HDLC	Univariate	Adjusted for TG
From original data (Z_1 and Z_2)	0.26	0.12	−0.39	−0.32
Using simulation mean (SD) over 50 simulations				
Uncorrected	0.24	0.11	−0.38	−0.32
(Z_1^* and Z_2^*)	(0.05)	(0.06)	(0.03)	(0.04)
Corrected	0.37	0.22	−0.41	−0.25
(X_1^* and X_2^*)	(0.04)	(0.06)	(0.02)	(0.05)
Using Rosner's method	0.35	0.19	−0.40	−0.27

exposures can appear to have "independent" effects on a disease outcome when this is not in fact the case. In the HDLC/TG example given, which modelled plausible degrees of measurement imprecision, the adjustments made lead to a dramatic change in relative risk estimates. HDLC, with the larger relative risk in univariate analysis, virtually displaced TG in a multiple logistic model. After associating a greater degree of measurement imprecision to TG than to HDLC this situation was completely reversed. The results before and after examining the possible effects of measurement imprecision would have very different implications for the attribution of causal primacy.

Even so, the "corrected" results were highly dependent on the amount of measurement imprecision attributed to HDLC and TG (compare Tables 1–3) and also involved making several assumptions (e.g. that the "usual" level is the most important aspect of the exposure and that the random error associated with the exposures are uncorrelated). This illustrates the problems with attempting to correct for measurement imprecision, when a study is being analysed. In some instances, when exposures are substantially correlated and one or more is measured with a sizeable amount of imprecision, it may have to be faced that the methodology of the study which has been carried out is inadequate to disentangle the independent contributions of the various exposures to the disease outcome. The HDLC/TG example would appear to be a case in point. We have used this example to illustrate how sensitive the "corrected" relative risk estimates are to the amount of imprecision attributed to the exposures when those exposures are substantially correlated, in order to show the importance of considering measurement imprecision in the design of epidemiological studies. The example is not intended to encourage the use of "correction" methods to generate "corrected" estimates from inadequately designed studies.

A further example of where measurement imprecision in correlated exposures causes a problem occurs in the investigation of cervical cancer. Cigarette smoking has been advanced as a contributory factor in the aetiology of cervical cancer [30]. It is generally accepted that a sexually transmitted infectious agent is involved in the genesis of this condition, but in epidemiological studies reported sexual behaviour is taken as a proxy measure of risk of contact with this agent. In the populations studied, cigarette smoking and reports regarding sexual behaviour (e.g. number of sexual partners and age at first intercourse) are strongly associated. Thus in one study [31] the relative odds for cervical cancer for current cigarette smokers, compared to never smokers, was 10.1, which fell to 3.4 after adjustment for confounders, which included number of sexual partners. It is likely that classification as a current smoker is a relatively precise procedure compared to the use of the number of sexual partners as a proxy for the chance of contact with the putative transmissible agent for cervical cancer. Only a small degree of imprecision in the indexing of risk of contact, through a question regarding number of sexual partners, would be necessary to produce an apparent independent effect of cigarette smoking.

In response to a comment along similar lines regarding these results [32], data from a different study were produced, which demonstrated a higher crude odds ratio for smoking 20 or more cigarettes a day than for reporting of 4 or more lifetime sexual partners [33]. Furthermore, in a logistic model the relative odds associated with

smoking was reduced by only 25%, whereas the relative odds associated with number of sexual partners was reduced by a greater amount, 31%. Thus, it was implied, an observer without preconceptions would consider smoking to be the factor more likely to be causally associated with cervical cancer.

These results can be seen as analogous to those in our example. When using the actual measurements HDLC was associated with a larger crude relative risk than TG, with the relative odds associated with TG also decreasing more in the logistic model. However, modelling a greater degree of measurement imprecision for TG than HDLC reversed these relationships. In the cervical cancer example neither the demonstration that the relative odds associated with smoking was larger than that associated with number of sexual partners, nor the fact that adjustment reduced the latter more than the former can be taken as necessarily supporting an important causal role for smoking. This conclusion is strengthened when consideration is given to the likely degree of measurement error in smoking and sexual behaviour data and to the role of reported sexual behaviour as a proxy measure for a more fundamental exposure. Again, in this instance, one could not have sufficient confidence in methods of "correction" to use them in order to make inferences about the aetiology of cervical cancer.

As we have said, rather than seek to correct for measurement imprecision at the analysis stage, epidemiologists should attempt to prevent it by improving the design of studies. It may be necessary, for example to take more than one measurement over a period of time for some exposures. Furthermore, since, as we have illustrated, the problems of measurement imprecision are most acute when two exposures are substantially correlated, epidemiologists faced with two correlated exposures measured imprecisely should seek populations where the correlation between the two exposures is smaller; i.e. where the confounding is broken. For the cervical cancer example, this would mean finding a culture where cigarette smoking is uncorrelated with sexual behaviour. The correction methods described in this paper may be useful for assessing whether measurement imprecision poses a significant problem in a particular situation. However, if a significant problem is identified, such as in the HDLC/TG example, then it would be dangerous to make inferences from such "corrected" relative risk estimates.

In conclusion, ignoring measurement imprecision in correlated exposures can lead to serious distortions of relative risk estimates from epidemiological studies. These can have important implications for the interpretation and implementation of the results of such studies. Epidemiologists designing future studies should anticipate measurement imprecision by collecting multiple measurements, where necessary, and attempting to minimalize correlations between exposures.

Acknowledgements—We are grateful to Professor AG Shaper for helpful suggestions and for permission to make use of the data from the British Regional Heart Study, which is a British Heart Foundation Research Group also supported by the Chest, Heart and Stroke Association and the Department of Health. This paper arose out of discussions with Juha Pekkanen at the **3rd Advanced Seminar in Cardiovascular Epidemiology and Prevention** in Corsendonk, Belgium (International Society and Federation of Cardiology, August 1989). Valuable comments on an earlier draft were received from Simon Thompson, Michael Hughes, Thierry Mertens, Goya Wannamethee and Ian White. We also thank the anonymous referees for several constructive comments and suggestions.

REFERENCES

1. Fuller WA, Hidiroglou MA. Regression estimates after correcting for attenuation. **J Am Stat Assoc** 1978; 73: 99–104.
2. De Klerk NH, English DR, Armstrong BK. A review of the effects of random measurement error on relative risk estimates in epidemiological studies. **Int J Epidemiol** 1989; 18: 705–712.
3. Rosner B, Willett WC, Spiegelman D. Correction of logistic regression relative risk estimates and confidence intervals for systematic within-person measurement error. **Stat Med** 1989; 8: 1051–1069.
4. MacMahon S, Peto R, Cutler J *et al.* Blood pressure, stroke and coronary heart disease, Part 1: Effects of prolonged differences in blood pressure—evidence from nine prospective observational studies corrected for the regression dilution bias. **Lancet** 1990; 335: 765–774.
5. Cornfield J, Haenszel W, Hammond EC, Lilienfeld AM, Shimkin MB, Wynder EL. Smoking and lung cancer: recent evidence and a discussion of some questions. **J Natl Cancer Inst** 1959; 22: 173–203.
6. Bross IDJ. Spurious effects from an extraneous variable. **J Chron Dis** 1966; 19: 637–647.
7. Schlesselman JJ. Assessing effects of confounding variables. **Am J Epidemiol** 1978; 108: 3–8.
8. Breslow NE, Day NE. Statistical methods in cancer research. Vol. 1: **The Analysis of Case–Control Studies.** Lyon: IARC, 1980.
9. Schlesselman JJ. **Case–Control Studies.** Oxford: Oxford University Press; 1982.
10. Savitz DA, Baron AE. Estimating and correcting for confounder misclassification. **Am J Epidemiol** 1989; 129: 1062–1071.
11. Fung KY, Howe GR. Methodological issues in case–control studies III: The effect of joint misclassification of risk factors and confounding factors upon estimation and power. **Int J Epidemiol** 1984; 13: 366–370.
12. Greenland S, Robins JM. Confounding and misclassification. **Am J Epidemiol** 1985; 122: 495–506.

13. Greenland S. The effect of misclassification in the presence of covariates. **Am J Epidemiol** 1980; 112: 564–569.
14. Tzonou A, Kaldor J, Smith PG, Day NE, Trichopolous D. Misclassification in case–control studies with two dichotomous risk factors. **Rev Epidemiol Sante Publ** 1986; 34: 10–17.
15. Liu K. Measurement error and its impact on partial correlation and multiple linear regression analysis. **Am J Epidemiol** 1988; 127: 864–874.
16. Rosner B, Spiegelman D, Willett W. Correction of logistic regression relative risk estimates and confidence intervals for measurement error: the case of multiple covariates measured with error. **Am J Epidemiol** 1990; 132: 734–745.
17. Armstrong BG, Whittemore AS, Howe GR. Analysis of case–control data with covariate measurement error: application to diet and colon cancer. **Stat Med** 1989; 8: 1151–1163.
18. Mosteller F, Tukey JW. **Data Analysis and Regression. A Second Course in Statistics** Philippines: Addison-Wesley; 1977.
19. Davis CE, Gordon D, La Rosa J, Wood PDS, Halperin M. Correlations of plasma high-density lipoprotein cholesterol levels with other plasma lipid and lipoprotein concentrations. **Circulation** 1980; 62 (Suppl. IV): IV-24–IV-30.
20. Hulley SB, Rosenman RH, Bawol RD, Brand RJ. Epidemiology as a guide to clinical decisions: The association between triglyceride and coronary heart disease. **N Engl J Med** 1980; 302: 1383–1389.
21. Cornfield J. Principles of research. **Am J Ment Defic** 1959; 64: 240–252.
22. Freedman DS, Gruchow HW, Anderson AJ *et al.* Relation of triglyceride levels to coronary artery disease: the Milwaukee cardiovascular data registry. **Am J Epidemiol** 1988; 127: 1118–1130.
23. Betteridge DJ. Lipoproteins and coronary heart disease. **J Human Hypertension** 1989; 3 (Suppl. 2): 13–25.
24. Riemersma RA. Coronary artery disease: raised cholesterol or triglycerides? **Int J Cardiol** 1984; 5: 193–194.
25. Austin MA. Plasma triglycerides as a risk-factor for coronary heart disease. The epidemiological evidence and beyond. **Am J Epidemiol** 1989; 129: 249–259.
26. Abbott RD, Carroll RJ. Interpreting multiple logistic regression coefficients in prospective observational studies. **Am J Epidemiol** 1984; 119: 830–836.
27. Shaper AG, Pocock SJ, Walker M, Cohen NM, Wale CJ, Thomson AG. British Regional Heart Study: cardiovascular risk factors in middle-aged men in 24 towns. **Br Med J** 1981; 283: 179–186.
28. Shaper AG, Pocock SJ, Walker M, Philips AN, Whitehead TP, Macfarlane PW. Risk factors for ischaemic heart disease: the prospective phase of the British Regional Heart Study. **J Epidemiol Commun Health** 1985; 19: 197-0-209.
29. Mogadam M, Ahmed SW, Mensch AH, Godwin ID. Within-person fluctuations of serum cholesterol and lipoproteins. **Arch Intern Med** 1990; 150: 1645–1648.
30. Winkelstein W. Smoking and cancer of the uterine cervix: hypothesis. **Am J Epidemiol** 1977; 106: 257–259.
31. Slattery ML, Robison LM, Schuman KL *et al.* Cigarette smoking and exposure to passive smoke are risk factors for cervical cancer. **JAMA** 1989; 261: 1593–1598.
32. Layde PM. Smoking and cervical cancer: cause or coincidence? **JAMA** 1989; 261: 1631–1633.
33. Winkelstein W. Smoking and cervical cancer: cause or coincidence. **JAMA** 1989; 262: 1631–1632.