STATISTICS IN MEDICINE, VOL. 11, 953–961 (1992)

BIAS IN RELATIVE ODDS ESTIMATION OWING TO IMPRECISE MEASUREMENT OF CORRELATED EXPOSURES

ANDREW N. PHILLIPS*

Department of Public Health and Primary Care, Royal Free Hospital School of Medicine, Rowland Hill Street, London NW3 2PF, U.K.

AND

GEORGE DAVEY SMITH

Department of Epidemiology and Population Sciences, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, U.K.

SUMMARY

A series of graphs is presented that show the estimated degree of bias in logistic coefficient estimates for two correlated continuous exposures measured with imprecision. These graphs indicate that even when the correlation coefficient between the exposure of interest and a correlated exposure is as low as 0·2, imprecision in the measurement of the latter exposure can result in at least as serious bias in the logistic coefficient estimate for the exposure of interest as measurement imprecision in the exposure of interest itself. The implications for the design and interpretation of epidemiological studies are discussed.

INTRODUCTION

Inferences about the aetiology of disease are often made from logistic coefficient estimates. Logistic models of risk factors for myocardial infarction have been used as a source of information as to which, for example, of the serum lipids HDL cholesterol and triglycerides are likely to be causally related to coronary heart disease.[1–5] The multiple logistic coefficient estimate for HDL cholesterol has often been found to be of much greater magnitude than that for triglycerides, with which HDL cholesterol is highly negatively correlated, when both are included in the same model. This fact has been considered as evidence that HDL-cholesterol, and not triglycerides, is important in the genesis of coronary heart disease (CHD).[1,4,5]

If a single measurement of an exposure poorly characterizes individuals' true or 'usual' exposure, either due to limitations in measuring techniques or because the exposure varies naturally within individuals, the relative odds estimate for the exposure will be biased.[6,7] Mis-characterization to some degree occurs in virtually all cases. If this mis-characterization is

* Current address: MRC UK 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, U.K.

0277–6715/92/070953–09$05.00
© 1992 by John Wiley & Sons, Ltd.
Received December 1990
Revised October 1991

random (from now on we use the term *measurement imprecision* to define this random mischaracterization for continuous exposures) and non-differential with respect to the outcome, then, for the situation where there is only one exposure variable in the model (that is, a univariate logistic model) this bias will always tend to be towards zero. This will still be the case for the *multiple* logistic coefficient estimate, as long as the exposure of interest is not substantially correlated with other exposures. In these circumstances, where there are no important potential confounding exposures, the bias is unlikely to result in inferences which are seriously misleading. That is to say, only the magnitude of the effects of exposures will be estimated incorrectly, not their direction. If, however, exposures are substantially correlated and measured with varying degrees of precision, then the multiple logistic coefficient estimates can be biased in *either* direction.[8,9] Most seriously, the sign of the logistic coefficient estimate can be wrongly estimated. It is in this situation, when two exposures are substantially correlated, that measurement imprecision can have the most severe consequences. This may be the case in our HDL cholesterol/triglycerides example. The two lipids are correlated with coefficient approximately $-0{\cdot}4$ and, furthermore, triglycerides are known to vary within individuals more than HDL cholesterol. It is possible, therefore, that if this bias could be removed, HDL cholesterol would emerge as being less strongly related to risk of CHD than triglycerides. This would have significant implications for the study and treatment of CHD. For example, some physicians have suggested that consideration be given to intervention with specific HDL-cholesterol-raising agents in men with low total cholesterol and low HDL-cholesterol.

Methods for estimating the degree of bias in such situations have been proposed for both continuous exposure variables, with which we are concerned, and for misclassification of categorical exposure variables. Based on these estimates, methods of 'correcting' for the bias have been advanced.[8,10–13] We have previously used one of these methods (described below),[10] which we compared with a simulation approach of our own, to estimate the degree of bias in the HDL/triglycerides example.[14] The simulation approach was based on studying the effect on multivariate relative odds estimation of adding random error to the exposure variables. The two approaches gave similar results. However, we found that the degree of bias was extremely sensitive to the amount of measurement imprecision ascribed to the correlated exposures HDL cholesterol and triglycerides. A change of just 0·1 (from 0·6 to 0·7) in the reliability coefficient (defined as the correlation between the true and measured exposures) ascribed to triglycerides, with no change in that ascribed to HDL (0·9), resulted in complete reversal in the apparent relative 'independent' importance of the two lipids. The 'corrected' logistic coefficient estimate, corresponding to a one standard deviation change in triglycerides, moved from 0·32 to 0·16, while that for HDL moved from $-0{\cdot}17$ to $-0{\cdot}31$. Because of this sensitivity we concluded that methods for 'correcting' estimates for measurement imprecision may be useful for assessing whether measurement imprecision represents a major problem in a particular situation. If such is the case, however, the methods could not be said to give reliable 'corrected' estimates of the 'independent' effect of the exposures. This reflects the fact that corrected estimates will tend to be more inaccurate as the amount of imprecision in exposure measurement increases. In circumstances where measurement imprecision is a major problem, as it is in the lipids/CHD example, we suggest the fact might have to be faced that the methodology of the study which has been carried out is inadequate to answer the question. A study in which subjects' exposures are more precisely measured at baseline is required.

Thus, the situation where serious problems can arise is when two exposures are substantially correlated and one or both is measured with a substantial amount of imprecision. The question in both cases is what is 'substantial'? In this paper we attempt to answer this by presenting a series of graphs which show the degree of bias in logistic coefficient estimates for the situation where there

are two risk exposures correlated to varying degrees, and measured with varying degrees of imprecision.

BIAS IN RELATIVE ODDS ESTIMATES FOR CORRELATED EXPOSURES MEASURED IMPRECISELY.

The 'corrected' logistic coefficient estimates are derived using the method of Rosner *et al.*[10,14] We assume that there is normal random 'error' in the measurement of two continuous exposures, X_1 and X_2, so that they are measured as Z_1 and Z_2, respectively. It is assumed that the errors are uncorrelated. For our multiple logistic model we have

$$\ln\{p'/(1-p')\} = \alpha' + \beta_{1c}.X_1 + \beta_{2c}.X_2$$

where p' is the estimated probability of disease and β_{1c} and β_{2c} are the 'correct' logistic coefficients and

$$\ln\{p/(1-p)\} = \alpha + \beta_{1u}.Z_1 + \beta_{2u}.Z_2$$

where p is the estimated probability of disease and β_{1u} and β_{2u} are the 'uncorrected' logistic coefficients.

Note that the X's are unknown, so β_{1c} and β_{2c}, the parameters we wish to estimate, cannot be estimated directly. Now consider the bivariate linear regression of X_1 on Z_1 and Z_2 and X_2 on Z_1 and Z_2:

$$X_1 = a_1 + \lambda_{11}.Z_1 + \lambda_{12}.Z_2 + e_1$$

and

$$X_2 = a_2 + \lambda_{21}.Z_1 + \lambda_{22}.Z_2 + e_2$$

where (e_1, e_2) are distributed as bivariate Normal.

Rosner *et al.*[10] suggest that the 'corrected' logistic coefficients β_{1c} and β_{2c} can be estimated by $\hat{\beta}_{1c}$ and $\hat{\beta}_{2c}$ where

$$(\hat{\beta}_{1c}, \hat{\beta}_{2c}) = (\beta_{1u}, \beta_{2u}) \begin{pmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{pmatrix}^{-1}.$$

Given the assumption of bivariate 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.

Since interest is typically focused on the effect of one particular exposure and other exposures are included owing to a possible confounding role, we refer to the two exposures as the 'exposure of interest' and the 'confounding' exposure. Strictly, the word 'confounding' implies a certain causal structure. We use the word as a convenient way to describe an exposure which is correlated with the exposure of interest, and is therefore only potentially a confounder. The graphs show the bias in the coefficient estimates for the exposure of interest. This is calculated as the 'uncorrected' logistic coefficient estimate *minus* the 'corrected' logistic coefficient estimate. The coefficient estimates correspond to a one standard deviation difference in the exposure. The measurement imprecision is expressed in terms of reliability coefficient, which was defined earlier. This can be estimated by the square root of the intraclass correlation coefficient between replicate measures on the same individual.[6] When there are two replicates only, the intraclass 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.[6]

RESULTS

Figure 1 shows the amount of bias in the logistic coefficient estimate for the exposure of interest (vertical axis is 'uncorrected' logistic coefficient estimate minus 'corrected' logistic coefficient estimate) according to different levels of the logistic coefficient estimate for the exposure of interest (horizontal axis) and the logistic coefficient estimate for the confounder (diagonal lines). This graph is for a correlation between the exposure of interest (as measured) and the confounder (as measured) of 0·2, and reliability coefficients of 0·7 and 0·9 for the exposure of interest and the confounder, respectively. The first point to note is that, since the diagonal lines are so close together, the logistic coefficient estimate for the confounder appears to make a relatively small difference to the bias in the coefficient estimate for the exposure of interest, in this instance. However, since the exposure of interest is measured with a sizeable degree of imprecision (reliability coefficient 0·7), there is a substantial amount of bias in its coefficient estimate. For an 'uncorrected' coefficient estimate (horizontal axis) of 0·4, for example, the bias is around − 0·2. In other words, the 'correct' coefficient estimate is around 0·6. For a coefficient estimate of − 0·4, on the other hand, the bias is around + 0·2. In both instances this represents an *underestimate* by one-third of the magnitude of the logistic coefficient. In this example the correlation between the exposure of interest and the confounder is relatively small *and* the measurement imprecision of the confounder is small. Thus this case is similar to the univariate situation. This is consistent with the observation that the bias in the logistic coefficient estimate for the exposure of interest tends to be *towards zero.*

Figure 2 shows a series of graphs in the same format as that in Figure 1 for various correlations (increasing from 0·1 to 0·4 across the rows) between the exposure of interest and the confounder and for various degrees of measurement error ascribed to the two exposures. The different diagonal lines, which represent different estimates for the ('uncorrected') logistic coefficient for the confounder, are shown for values − 0·4, 0·0 and + 0·4 only. In every case the lowest diagonal line is for − 0·4 and the highest for + 0·4. Clearly, the closer the diagonal lines are to the horizontal 'zero bias' line, the smaller the amount of bias. It can immediately be seen that, in every case, the bias is lowest when the logistic coefficient estimate for the confounder is zero, and that the bias increases with the magnitude of the 'uncorrected' coefficient estimate. On each graph, the shaded area represents that region where the bias is such that the 'uncorrected' logistic coefficient estimate *is in the wrong direction.* This is when the bias is most serious. For example, consider the situation where the Pearson correlation coefficient between the exposures is 0·4 and the measurement imprecision of both the exposure of interest and the confounder is given by a reliability coefficient of 0·7. For 'naive' logistic coefficient estimates of 0·15 for the exposure of interest and 0·4 for the confounder, the bias is given as + 0·39. Thus, the 'correct' logistic coefficient estimate for the exposure of interest is − 0·24, rather than the + 0·15 estimated. Clearly, this is potentially an extremely misleading bias.

In studying the graphs, first consider the change in the pattern of bias as, with increasing correlation coefficient between the two exposures, we move from left to right across the graphs. No matter what the measurement imprecision in the exposures, the diagonal lines tend to move further apart, and thus from the zero bias line. Thus, as expected, the bias increases with increasing correlation between the exposure of interest and the confounder, whatever the measurement imprecision ascribed to the two exposures (unless the reliability coefficient of either exposure is 1, which is uncommon). In keeping with this, the shaded area, representing situations

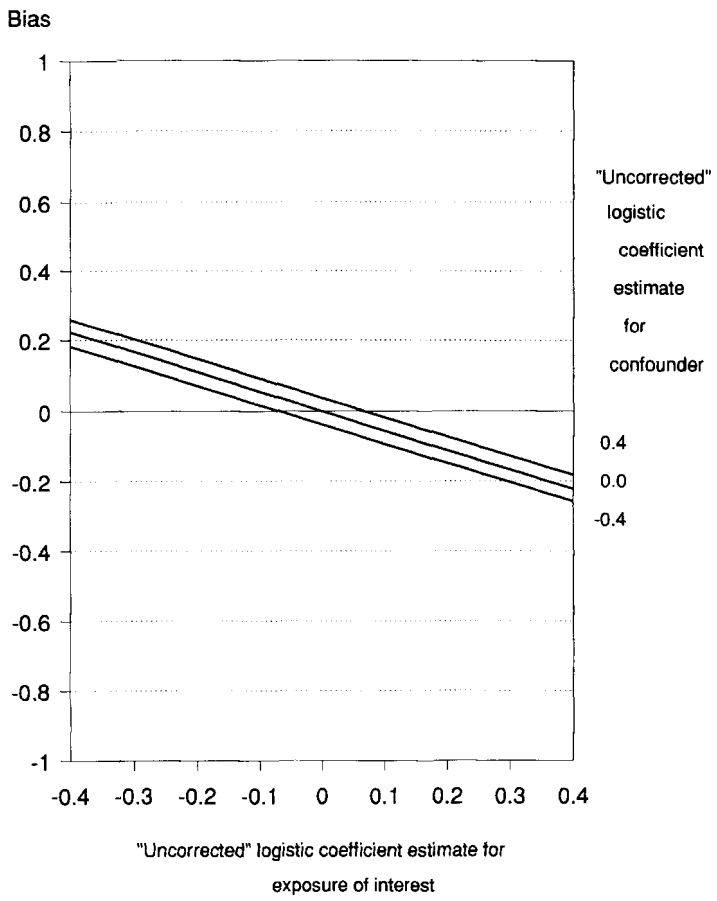

Figure 1. Graph showing degree of bias in logistic coefficient estimate according to the size of the 'uncorrected' coefficient estimate for the exposure of interest (horizontal axis) and the confounder (diagonal lines). Correlation between exposures 0·2, reliability coefficients for exposure of interest and confounder 0·7 and 0·9, respectively

where the coefficient estimate is in the wrong direction, grows larger as the correlation between the exposures grows larger.

Let us now consider differences between graphs in the four rows, for any given column; we consider variation in measurement imprecision of exposures at fixed correlations between exposures. When the both exposures are measured quite precisely (reliability coefficient 0·9 for both – row 1) the bias tends to be relatively small, especially when the correlation between exposures is small. Increasing measurement imprecision of the exposure of interest, at a constant level of measurement imprecision for the confounder, that is moving from row 1 to row 2 or from row 3 to row 4, has the effect of increasing the *angle* of the diagonal lines, but not their separation to any great degree. The size of the shaded area increases only slightly. On the other hand, increasing the measurement imprecision for the confounder, at a constant level of measurement imprecision for the exposure of interest, that is moving from row 1 to row 3 or from row 2 to row 4, has the effect of *separating* the diagonal lines, but not greatly changing their angle. Here the shaded area increases substantially.

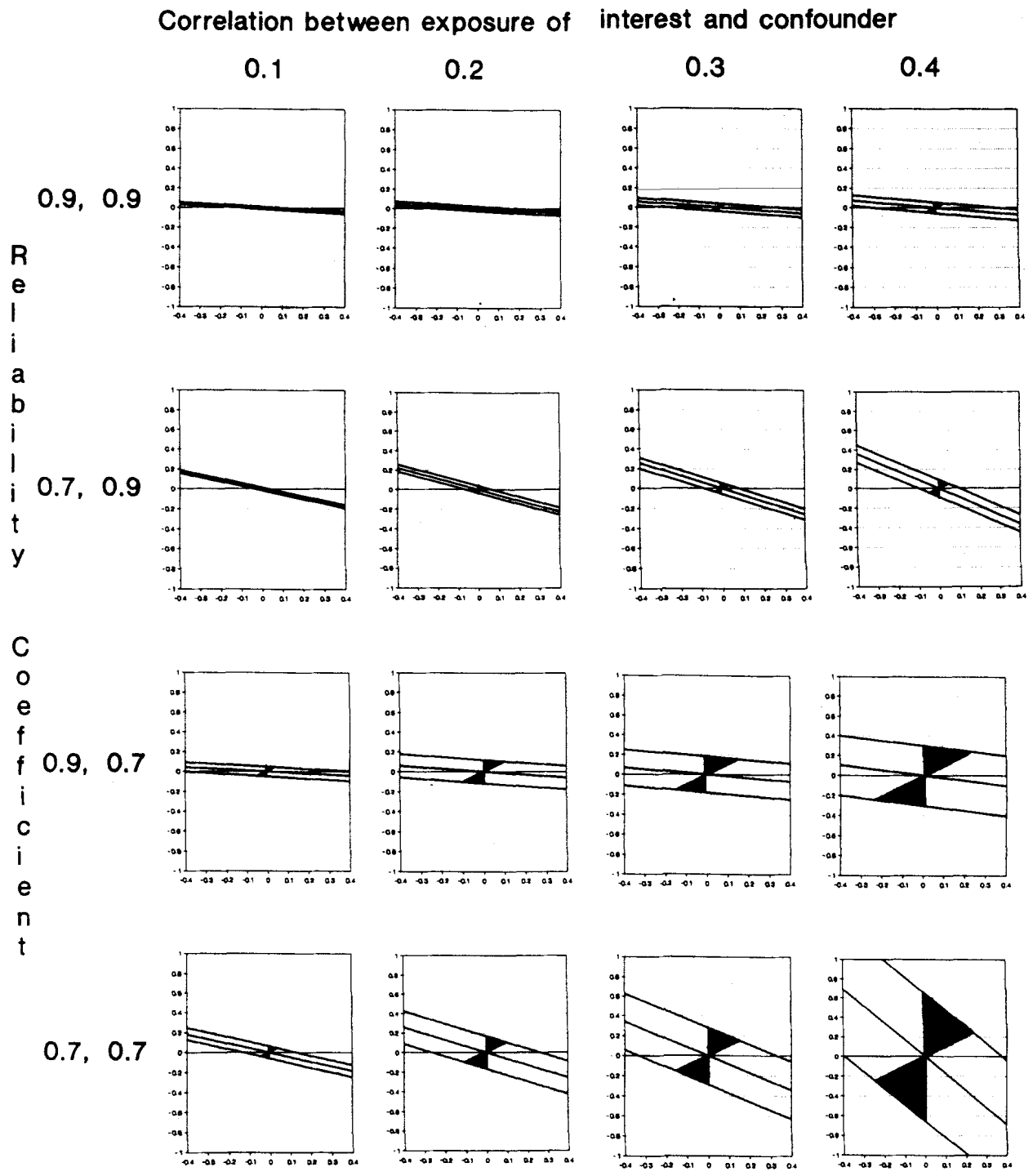

Figure 2. Series of graphs indentical to that in Figure 1 but for different correlations between the exposures and for different amounts of measurement imprecision ascribed to the exposures. The first reliability coefficient is for the exposure of interest, the second for the confounder. The shaded triangles cover regions where the bias results in the logistic coefficient estimate being in the wrong direction

Thus, increasing measurement imprecision in the exposure of interest creates increasing bias, but this is mainly underestimation of the magnitude of the coefficient. The sign of the 'uncorrected' coefficient estimate is almost always correct, as reflected by the fact that the shaded area is small, as long as the measurement imprecision of the confounder is small. However, when there is significant measurement imprecision for the confounder the shaded area tends to be quite large, even if there is little measurement imprecision for the exposure of interest. It appears, therefore, that measurement imprecision in a *confounding* exposure is potentially *more* serious than measurement imprecision in the exposure of interest. If there is substantial measurement imprecision in the confounder (for example, a reliability coefficient of 0·7) then, even if there is little measurement imprecision for the exposure of interest (reliability coefficient 0·9) and the correlation coefficient between the exposure of interest and the confounder is as low as 0·2, the bias can still be important (see graph on row 3, column 2). In a situation where the 'uncorrected' logistic coefficient estimate for the confounder is 0·4, a coefficient of $-0{\cdot}17$ for the exposure of interest would be estimated as 0·0 without 'correction'.

DISCUSSION

The graphs presented in Figure 2 give useful guidance as to the situations in which measurement imprecision in correlated continuous exposures presents a serious problem. They illustrate how measurement imprecision in a correlated exposure will often have more important consequences than measurement imprecision in the exposure of interest, at least in the situation where the exposures are continuous. The situation with misclassification of categorical variables has been more widely studied.[8,9,12,13,15,16–18] In this situation some authors have suggested that misclassification of confounders can be as important as misclassification of the variable of interest.[8,9] None, to our knowledge, has illustrated that error in confounding variables is actually the more serious potential problem, although Kupper has reported a similar result when surrogate exposures are used.[15] Although we have made use of methods of 'correction' of logistic coefficient estimates in order to illustrate the approximate amount of bias that can arise in different situations we would advocate caution in the adoption of these techniques, as the 'corrected' estimates can prove extremely unstable in some situations.[14] The greater the factor of correction in the estimates, the greater the probable error in the final estimates. Thus, only when small degrees of adjustment are involved, relative to the magnitude of the 'uncorrected' estimates, should 'corrected' values be accepted with any confidence.

When two correlated exposures are measured with a significant amount of imprecision, a study in which each exposure is measured only once must be considered to be of inadequate design for determining the independent contributions of the two exposures to risk. The optimal solution would be to find a study population in which the correlation between the exposures is greatly reduced or non-existent. In this situation measurement error in either exposure becomes less of a problem. Breaking confounding in this manner is possible when exposures are associated for cultural rather than biological reasons. For example, sexual behaviour is considered to be a risk factor for cervical cancer. Some commentators[19] also consider smoking to be a cause of cervical cancer. Smoking and sexual behaviour are related in many societies, however, which leads to confounding of the association between cervical cancer and smoking. In theory it should be possible to study the relationship within a society in which the correlation between smoking and sexual behaviour is weak or non-existent.

In many situations – such as the HDL/triglyceride example discussed above – the correlation between two exposures is probably universal and populations cannot be found in which the association does not exist. In these cases the reduction of measurement imprecision in exposures

becomes vitally important. Although in some circumstances reduction in measurement imprecision is impossible, in many instances epidemiologists could obtain improved measurement precision if this aspect of study design were given a higher priority compared with issues such as sample size and sample selection. For example, more precise characterization of subjects' concentrations of biochemical and haematological variables could be obtained if these were measured more than once on each subject and the mean taken. This might mean that the overall size of a study would have to be smaller, owing to cost considerations. However, the loss in precision of relative odds estimates resulting from the smaller study size would often be more than compensated by the decrease in bias which results from the lower measurement imprecision of the exposure variables.

In some situations, when measurement imprecision cannot be improved, observational epidemiological studies may not be able to sort out whether the effects associated with correlated exposures are independent of each other. Intervention studies may be able to alter one but not the other exposure, providing a better basis for making such decisions. It would, however, be preferable to recognize the limitations of available methodologies than to accept the 'independence' of many effects, which may merely be the result of residual confounding.[20]

ACKNOWLEDGEMENTS

We thank Mr. Simon Thompson, Dr. Michael Hughes and Dr. Stephen Senn for useful comments and suggestions on an earlier draft. We are also grateful to two anonymous referees who made several constructive comments on this paper which have resulted in a significant improvement.

REFERENCES

1. Hulley, S. B., Rosenman, R. H., Bawol, R. D. and Brand, R. J. 'Epidemiology as a guide to clinical decisions: The association between triglyceride and coronary heart disease', *New England Journal of Medicine*, **302**, 1383–1389 (1980).
2. Fraser, G. E. *Preventive Cardiology*, Oxford, New York, 1986.
3. Gordon, D. J. and Rifkind, B. M. 'High-density lipoprotein – the clinical implications of recent studies', *New England Journal of Medicine*, **321**, 1311–1316 (1989).
4. Gordon, T. 'The diet-heart idea. Outline of a history', *American Journal of Epidemiology*, **127**, 220–225 (1988).
5. Jacobs, D. R., Mebane, I. L., Bangdiwala, S. I., Criqui, M. H. and Tyroler, H. A. 'High density lipoprotein cholesterol as a predictor of cardiovascular disease mortality in men and women: the follow-up study of the lipid research clinics prevalence study', *American Journal of Epidemiology*, **131**, 32–47 (1990).
6. De Klerk, N. H., English, D. R. and Armstrong, B. K. 'A review of the effects of random measurement error on relative risk estimates in epidemiological studies', *International Journal of Epidemiology*, **18**, 705–712 (1989).
7. Rosner, B., Willett, W. C. and Spiegelman, D. 'Correction of logistic regression relative risk estimates and confidence intervals for systematic within-person measurement error', *Statistics in Medicine*, **8**, 1051–1069 (1989).
8. Fung, K. Y. and Howe, G. R. 'Methodological issues in case-control studies III—The effect of joint misclassification of risk factors and confounding factors upon estimation and power', *International Journal of Epidemiology*, **13**, 366–370 (1984).
9. Greenland, S. and Robins, J. M. 'Confounding and misclassification', *American Journal of Epidemiology*, **122**, 495–506 (1985).
10. Rosner, B., Spiegelman, D. and Willett, W. 'Correction of logistic regression relative risk estimates and confidence intervals for measurement error: the case of multiple covariates measured with error', *American Journal of Epidemiology*, **132**, 734–745 (1990).
11. Byar, D. P. and Gail, M. H. (Eds) Workshop on errors-in-variables', *Statistics in Medicine*, **8**, 1027–1179 (1989).
12. Kaldor, J. and Clayton, D. 'Latent class analysis in chronic disease epidemiology', *Statistics in Medicine*, **4**, 327–335 (1985).

13. Savitz, D. A. and Baron, A. E. 'Estimating and correcting for confounder misclassification', *American Journal of Epidemiology*, **129**, 1062–1071 (1989).
14. Phillips, A. N. and Davey Smith, G. 'How independent are "independent" effects? Relative risk estimation when correlated exposures are measured imprecisely', *Journal of Clinical Epidemiology*, **44**, 1223–1231 (1991).
15. Kupper, L. L. 'Effects of the use of unreliable surrogate variables on the validity of epidemiologic research studies', *American Journal of Epidemiology*, **120**, 643–648 (1984).
16. Greenland, S. 'The effect of misclassification in the presence of covariates', *American Journal of Epidemiology*, **112**, 564–569 (1980).
17. Tzonou, A., Kaldor, J., Smith, P. G., Day, N. E. and Trichopoulos, D. 'Misclassification in case-control studies with two dichotomous risk factors', *Revue Epidemiologie et Sante Publique*, **34**, 10–17 (1986).
18. Duffy, S. W., Rohan, T. E. and Day, N. E. 'Misclassification in more than one factor in a case-control study: a combination of Mantel–Haenszel and maximum likelihood approaches', *Statistics in Medicine*, **8**, 1529–1536 (1989).
19. Winkelstein, W. 'Smoking and cervical cancer – current status: a review', *American Journal of Epidemiology*, **131**, 945–957 (1990).
20. Davey Smith, G. and Phillips, A. N. 'Declaring independence. Why we should be cautious', *Journal of Epidemiology and Community Health*, **44**, 257–258 (1990).