Effect of measurement error on epidemiological studies of environmental and occupational exposures

Ben G Armstrong

Abstract

Random error (misclassification) in exposure measurements usually biases a relative risk, regression coefficient, or other effect measure towards the null value (no association). The most important exception is Berkson type error, which causes little or no bias. Berkson type error arises, in particular, due to use of group average exposure in place of individual values. Random error in exposure measurements, Berkson or otherwise, reduces the power of a study, making it more likely that real associations are not detected. Random error in confounding variables compromises the control of their effect, leaving residual confounding. Random error in a variable that modifies the effect of exposure on health—for example, an indicator of susceptibility—tends to diminish the observed modification of effect, but error in the exposure can create a spurious appearance of modification. Methods are available to correct for bias (but not generally power loss) due to measurement error, if information on the magnitude and type of error is available. These methods can be complicated to use, however, and should be used cautiously as "correction" can magnify confounding if it is present.

(*Occup Environ Med* 1998;**55**:651–656)

Keywords: environmental; occupational; measurement; error

An epidemiological study generally requires for each study subject a measure of health outcome and of one or more potential explanatory variables. Explanatory variables often comprise one or a few variables of interest (occupational or environmental exposure) and potential confounders (smoking, age, etc). All of these are usually measured with some degree of error.

Should such error decrease our confidence in the results? If so, in just what way? Does it add to uncertainty in estimates of measures of effect? If so, is this extra uncertainty reflected in the usual statements of uncertainty—such as confidence intervals? Under what circumstances does it cause bias in a result, and can we know the direction and extent of bias, or even correct for it? Does it compromise the power of the study?

The purpose of this article is to summarise, in a manner accessible to the non-statistician, what is known about the effects of measurement error on the results of a study. The main text of the article is organised in three sections. The first distinguishes different types of error. The second describes the effects of error according to its type, first qualitatively and then where possible quantitatively. Finally there is an overview of when and how corrections can be made for the effects of error in the statistical analysis. To keep the article manageable, the main focus is on error in explanatory variables. The effects of errors in the health outcome (response variable), which are rather different, are mentioned briefly in the discussion.

Terms and notation

The term relative risk (RR) is used here in the statistical tradition, generically to include rate ratios, odds ratios, prevalence ratios, etc. (This usage is standard in statistics. Some epidemiologists restrict the meaning of relative risk to be the ratio of cumulitve incidence.) The term effect measure is used to denote a summary of the association between exposure and outcome—for example, relative risk or regression coefficient. The true exposure is denoted T and the approximate measure X, and the error E, with:

$$X = T + E.$$

The SD of T, X, and E are written σ_T, σ_X, and σ_E respectively .

Types of measurement error

The effects of measurement error depend critically on its type.

ERROR IN MEASURING WHAT?

We distinguish three categories of explanatory variable:
- Variable of interest (environmental or occupational exposure)
- Potential confounder (for example, active smoking or socioeconomic status)
- Potential effect modifier (markers of vulnerability to the effects of the variable of interest—for example, age).

Correspondence to:
Dr Ben G Armstrong, Environmental Epidemiology Unit, London School of Hygiene and Tropical Medicine, Keppel Street, London.

DIFFERENTIAL OR NON-DIFFERENTIAL
- Differential error varies according to the health outcome (recall bias in case-control studies leads to different error in cases and controls)
- Non-differential error does not depend on health outcome.

SCALE OF MEASUREMENT OF THE VARIABLE(S) WITH ERROR
- Categorical (qualitative), comprising: dichotomous (exposed *v* not exposed); polytomous (for example, high, medium, low)
- Numerical (concentration of particles in air in mg/m³, number of cigarettes smoked per day).

When occurring in categorical variables, measurement error is termed misclassification—study subjects may be classified incorrectly. Numerical variables can be made into groups, and thus become categorical variables. Conversely ordered polytomous variables can sometimes be treated as numerical.

Two further distinctions apply to error in numerical variables:

RANDOM OR SYSTEMATIC?
- Systematic: for example, all exposures overestimated by 2 units or by 20%
- Random: some exposures overestimated, some underestimated

Error often has some systematic and some random component.

CLASSICAL OR BERKSON?
This distinction is not well known and a little tricky to understand, but it has major implications for the effects of the error.
- Classical: The average of many replicate measurements of same true exposure would equal the true exposure.
- Berkson: The same approximate exposure (proxy) is used for many subjects; the true exposures vary randomly about this proxy, with mean equal to it.

Example:

A study investigates the relation of mean exposure to lead up to age 10 with intelligence quotient (IQ) in 10 year old children living in the vicinity of a lead smelter. The IQ is measured by a test administered at age 10. Consider two study designs for assessing exposure:

Design 1: Each child has one measurement made of blood lead, at a random time during their life. The blood lead measurement will be an approximate measure of mean blood lead over life. However, if we were able to make many replicate measurements (at different random time points), the mean would be a good indicator of lifetime exposure. This measurement error is thus **classical**.

Design 2: The children's place of residence at age 10 (assumed known exactly) are classified into three groups by proximity to the smelter—close, medium, far. Random blood lead samples, collected as described in design 1, are averaged for each group, and this group mean used as a proxy for lifetime exposure for each child in the group. Here the same approximate exposure (proxy) is used for all subjects in the same group, and true exposures, although unknown, may be assumed to vary randomly about the proxy. This measurement error is thus **Berkson type** error.

Another situation giving rise to Berkson error is when exposures are estimated from observed determinants of exposure with an exposure prediction model. Often error has both classic and random components, although one usually predominates.

DESCRIBING THE MAGNITUDE OF ERROR
The likely extent of misclassification of categorical variables is usually specified as probabilities of misclassification. For dichotomous variables, it is conventional to express these through the sensitivity (the probability of correctly classifying a truly exposed subject as exposed), and the specificity (the probability of correctly classifying a non-exposed subject as non-exposed) of the classification. Thus if sensitivity is 0.8 and specificity is 0.7, the probability of misclassifying an exposed subject as non-exposed is $1-0.8=0.2$, and the probability of misclassifying a non-exposed subject as exposed is $1-0.7=0.3$. For categorical variables of more than two levels, many different sorts of misclassification can occur, which can be specified in a matrix of misclassification probabilities.

The average magnitude of errors (classical or Berkson) in numerical variables can be described by their SD (σ_E) or variance (σ_E^2). Classical error is generally described by what is termed its coefficient of reliability, which is defined as the correlation of independent repeated measurements of exposure (ρ_{XX}). This may be shown to be equal to the square of the validity coefficient, which is the correlation between the true and approximate measurements (ρ_{XT}), and also to the ratio of variances of true and approximate exposures (σ_T^2/σ_X^2), or to a function of the ratio of SDs of error and of true exposures $(1/(1+(\sigma_E/\sigma_T)^2))$. $(\rho_{XX} = \sigma_T^2/\sigma_X^2 = \sigma_T^2/(\sigma_T^2 + \sigma_E^2) = 1/(1+\sigma_E^2/\sigma_T^2) = 1/(1+(\sigma_E/\sigma_T)^2))$.)

This usage of the term reliability is standard in technical discussions of measurement error, although it is used more generally among epidemiologists as a synonym for reproducibility or precision.

Effects of measurement error
We begin by considering the effects of non-differential error or misclassification in the exposure of interest, firstly on effect measures, then on the results of significance tests.

In general, random measurement error or misclassification leads to bias in effect measures (relative risks, regression coefficients, differences in means). This bias is usually downwards (towards the null), but there are important exceptions. With information on magnitude of measurement error and exposure variability (or prevalence), extent of bias can be estimated.

Table 1   Effect of non-differential misclassification on RRs in two groups

Non-differential misclassification: true RR=2.0			
Exposure sensitivity	*Exposure specificity*	*Proportion of exposure in the population*	*Observed RR*
0.6	0.9	0.1	1.34
0.6	0.9	0.5	1.42
0.6	0.99	0.1	1.79
0.6	0.99	0.5	1.54
0.9	0.9	0.1	1.48
0.9	0.9	0.5	1.73
0.9	0.99	0.1	1.89
0.9	0.99	0.5	1.82

EXPOSURE MEASURED ON A DICHOTOMOUS SCALE

Non-differential error always biases the effect measure toward the null value (there is a technical but unrealistic exception when the sum of sensitivity and specificity of exposure classification is <1, implying measurement that tends to reverse exposed and unexposed categories!)

Example:

A study of lung cancer in relation to proximity of residence to a coke oven classifies subjects (cases and populations) by distance of residence from the oven at the time of follow up—near=<4 km from oven; far=4–10 km. The incidence is compared in the two groups. Here there is misclassification due to migration—not all people living near the oven at the time of follow up will have lived there at the aetiologically relevant time. Thus if the true relative risk for subjects living in these areas throughout their lives were 1.5, the observed relative risk would tend to be less.

The extent of bias is dependent on sensitivity and specificity of the classification, and the proportion of truly exposed people among those not diseased, and may be calculated from them[1].

Suppose misclassification (migration) in the above example was such that 10% of the near group was in fact far at the time of relevant exposure, and vice versa, and that 10% of the population overall lived in the near area. The observed relative risk would then be 1.26.

Further examples are given in table 1, taken from Armstrong *et al* p 71.[1]

EXPOSURE MEASURED ON A POLYTOMOUS SCALE

Non-differential error biases downwards estimates of trend across ordered groups, but comparisons between specific categories can be biased in either direction.

Example:

Assume that in the previous example the near group was split into two: very near, and

quite near, with true relative risks, relative to far, of 2.0 and 1.3. If there is migration from very near to quite near, but not otherwise, observed risks for the two near groups, relative to the far group, will be closer together—for example 1.6 and 1.4, as seen on table 2. Thus here the quite near v far group relative risk is increased by misclassification.

EXPOSURE MEASURED ON A NUMERICAL SCALE

Classic errors bias regression coefficients (relative risks per unit exposure) towards zero. We say the association is attenuated. In fact, for linear regression the bias factor is equal to the coefficient of reliability ($\rho_{XX} = \rho_{XT}^2 = \sigma_T^2/\sigma_X^2$); with the observed regression coefficient

$$\beta_{OBSERVED} = \rho_{XX} \times \beta_{TRUE}$$

Lead-IQ example—design 1

Suppose that a regression of IQ on true lifetime mean blood lead has a regression with coefficient -2 (IQ reduced by 2 points per μg/dl blood lead). With classic measurement error with coefficient of reliability 0.5, this would be attenuated, on average, to 0.5×-2=-1.

From the alternative expressions for the coefficient of reliability (end of the section on types of measurement error), we see that it depends on the average magnitude of measurement error relative to the average magnitude of the true exposure (σ_E/σ_X). This implies that measurement error will have less effect if the true exposures are more spread out (σ_X is greater). Table 3 gives attenuation bias as function of the ratio of the SD of errors to that of true exposures (σ_E/σ_X). This is quite reassuring—error has to be relatively big to give serious bias.

For logistic and log linear (Poisson) regression coefficients the same qualitative result is true, and the quantitative one approximately so. Note that for logistic and log linear regression relative risk is linked to the regression coefficient by the formula RR=exp(β), thus:

$$RR_{OBSERVED} \cong (RR_{TRUE})^{\rho_{XX}}$$

If, in the children exposed to lead, we were to use as an outcome a child having IQ <80, and if the relative risk (odds ratio) increment per 10 μg/dl true blood lead (from logistic regression) was 1.5, then the observed RR is given by:

$$RR_{OBSERVED} \cong 1.5^{0.5} = 1.22$$

Berkson errors, however, lead to no bias in linear regression coefficients, and little or no bias in logistic or log linear regression coefficients.

Lead-IQ example—design 2:

In this grouped design the error is of Berkson type, so there is no bias in the regression coefficient. However, precision would be lost (width of confidence interval would be wider), and power would not be as great as without measurement error, or as in the biased design 1.

Table 2   The effect on non-differential misclassification on RRs in three exposure groups (example)

	RR		
	Very near	*Quite near*	*Far*
True	2.0	1.3	1.0
Misclassified	1.6	1.4	1.0

Table 3   Attenuation bias due to exposure measurement error in linear regression

Error σ_E/σ_X	0.0	0.1	0.2	0.3	0.4	0.5	0.75	1.0	1.5	2.0
Attenuation*	1.0	0.99	0.96	0.92	0.86	0.80	0.64	0.50	0.31	0.20

*Attenuation is the factor by which the naive regression slope will underestimate the true slope.

EFFECT ON SIGNIFICANCE TESTS

All types of non-differential random measurement error or misclassification reduce study power—the chance that a study will find a significant association if one is truly present. The extent of power loss can be measured if magnitude of measurement error and exposure variability (or for a dichotomous measure prevalence) are known. The approach to measuring power loss is essentially the same for dichotomous or numerical variables, being based on the result that the effective loss in sample size is equal to the coefficient of reliability of the measure[2].

Example:

A study is designed to have 80% power to detect a relative risk of 1.7 between truly exposed and truly unexposed people (80% of similar sized studies would find the association). If approximate measurements were used, the power would be less. If the measure of exposure has sensitivity= specificity=0.7 (50% controls exposed), then power would be just 20%. Thus exposure effects may be missed because of measurement error.

Despite the bias and power loss noted above, the p values obtained with the usual methods on data subject to random error or misclassification are valid. Spurious "significant" results (where there is in fact no association) are no more likely with than without measurement error.

Example:

A study finds an association between dust and loss of lung function, with p=0.02, but dust measurements were known to be subject to error. Providing that the error is non-differential, the low p value cannot be attributed to the measurement error.

CONFOUNDERS

The general rule is that errors in confounders compromise our ability to control for their effect, leaving residual confounding. The effect measure adjusted with the approximate confounder will on average lie between the crude, unadjusted effect measure and the effect measure adjusted with the true (unknown) confounder. The validity of significance tests on the effect of exposure are compromised.

A study of the relation of lung cancer to air pollution adjusts for smoking with a crude estimate of pack-years for each subject. Any confounding of the relative risk for lung cancer versus air pollution will be only partially controlled.

Example:

If crude $RR_{(crude)}$=1.50 (95%CI 1.20-1.88; p<0.001), and $RR_{(adjusted\ for\ true\ pack-years)}$=1.04 (95%CI 0.86-1.24; p=0.67), then the partially adjusted $RR_{(adjusted\ for\ approximate\ pack-years)}$ will in general lie between 1.50 and 1.04 and the partially adjusted p value will lie between 0.001 and 0.67.

The degree of residual confounding depends on the coefficient of reliability of the measure of the confounder. A coefficient of reliability of 0.5 will imply that about half the confounding present will be controlled, in the sense that the observed log(RR) (more generally the regression coefficient) will on average lie about halfway between the crude unadjusted log(RR) and the fully adjusted log(RR).

Continuing the same example:

If the coefficient of reliability of measured pack-years is 0.5, then log ($RR_{(adjusted\ for\ approximate\ pack-years)}$) will lie about half way between log ($RR_{(crude)}$) and log ($RR_{(adjusted\ for\ true\ pack-years)}$), which gives $RR_{(adjusted\ for\ approximate\ pack-years)}$=1.25 (95% CI 1.03 to 1.52; p=0.03).

There are a few exceptions. Entirely systematic error (everyone underreporting their smoking by 20%) will not usually compromise control of confounding. In special situations (when the effects of the confounder and the exposure of interest are additive) Berkson error (for example, use of group mean rather than individual pack-years of smoking) also leaves no residual confounding. Most importantly, if the variable suspected of confounding is in fact not associated with the exposure of interest (smoking not associated with air pollution) then there is no confounding or residual confounding, however strongly the variable is associated with the outcome (however bad the data on smoking, the observed association of lung cancer with air pollution is not biased).

Having to control for confounders, whether measured with error or not, increases somewhat the effect of error in the variable of interest on the relative risk of interest.

EFFECT MODIFIERS

An effect modifier is a variable that modifies the effect of the exposure of interest—for example, identifying subgroups vulnerable or resistant to the exposure. In statistical terms, we say that there is an interaction between the effect modifier and the exposure. Error in measuring effect modifiers tends to diminish effect modification. Vulnerable subgroups are thus made harder to identify.

Lead-IQ example:

Suppose diet modified the effect of lead on IQ, children with vitamin deficient diets having a regression slope of -3, and others a slope of -1. If diet is measured with error (misclassified), the apparent modification will tend to be less—for example the slope in vitamin deficient children might be -2.5, and that in others -1.5.

Even if the putative effect modifier is measured without error, error in the variable of interest can distort effect modification, and even create spurious modification. This may happen because the magnitude of error, and hence bias due to it, depends on the putative modifier. Even if this is not the case, the variation of exposure may depend on the putative modifier, in which case the bias due to measurement error will again depend on the putative modifier.

Example:

Suppose now that we investigated for a modification of the effect of lead on IQ by sex, which is measured without error, but lead is again measured with (classical) error. Suppose also that although the average error was the same for boys and girls, boys had more varied lead exposures than girls (σ_T is higher in boys than girls). In this case, if the true regression slope of IQ on lead is −2 for both boys and girls, the estimated slope will tend to be more attenuated for girls (say to −0.5) than for boys (say to −1.5). (For girls the SD σ_T is lower, and hence the attenuation bias σ_E^2/σ_T^2 is greater.)

Thus sex seems to modify the effect of lead on IQ, but does not in fact do so.

DIFFERENTIAL ERROR

Differential error can cause bias in the effect measure either upwards or downwards, depending on whether adverse outcomes are associated with overestimation or underestimation of exposure. Significance tests are not valid in the presence of differential error. For dichotomous exposure, the bias can be measured if the sensitivity and specificity of the approximate classification are known.

Example:

The association of exposure to use of a video display unit (VDU) with spontaneous abortion is investigated by means of a case-control study in which women are interviewed after a live birth or abortion, and asked about the number of hours a week that they spent using a VDU. The relative risk of spontaneous abortion in women using VDUs for ≥15 hours a week was 1.20 (95% CI 1.06 to 1.34). Due to media attention to the hypothesised association, women who had experienced spontaneous abortions may have been more likely to recall their VDU use fully. In this case, some or all of the excess of VDU users in the cases relative to the controls would be spurious, so that the true relative risk would be less than 1.20, possibly 1.00.

Correcting for measurement error

If there is information on the magnitude and type of error it is possible (but not always easy!) to allow for it in estimating the effect measure, at least for reasonably simple forms of measurement error. Sometimes, it is sufficient to invert the formulae already shown for deriving the effects of measurement error—for example:

$$\beta_{TRUE} = \beta_{OBSERVED}/\rho_{xx}, \quad RR_{TRUE} = (RR_{OBSERVED})^{(1/\rho_{xx})}.$$

In the lead-IQ study:

If we had found a regression coefficient ($\beta_{OBSERVED}$) of -1, and known that the coefficient of reliability of measurement (ρ_{xx}) was 0.5, then we could estimate

$$\beta_{TRUE} = -1/0.5 = -2$$

Other methods are available which refine this rather crude approach. The aim of these more complex approaches is usually to more exactly eradicate bias, use other sorts of information on measurement error, or to reflect in the estimate and confidence intervals uncertainty as to the magnitude of the error.

To obtain information on measurement error magnitude reliability studies (a sample of repeated independent measurements) or validity studies (a sample of gold standard measurements in parallel with the approximate measurements) are needed. These are not often available, and even if they are, much uncertainty remains unless they are large. If corrections are carried out on the basis of incorrect information on error magnitude, bias may be increased, rather than decreased. Corrections for attenuation can also magnify confounding or other information bias, rather than a true association. It may be sensible for researchers to always give the naive effect measure (using the approximate exposure in a regular analysis), even when

including effect measures corrected for measurement error. Also worth considering is calculating corrections under various assumptions, in the spirit of a sensitivity analysis.

Corrections will not in general affect the p value of a test of the null hypothesis of no association, nor will the power of the test be improved. Confidence intervals will in general, however, get wider.

In the lead-IQ study:

Suppose the regression coefficient of −1 had a 95% CI (−1.8 to −0.2), with p=0.01. Assuming a coefficient of reliability 0.05, the corrected coefficient is −2, the 95% CI (−3.6 to −0.4), and p=0.01, as before.

Discussion

For simplicity of presentation some assumptions and points of interpretation have been passed over. The most important of these are:

- Many of these results concern bias, which is an average effect if the study were to be repeated many times. In a large study the effect of measurement error will be close to the mean. However, in a single small sample, the effect may differ appreciably from this mean[3]. In these cases random error can sometimes even lead to an effect measure estimated from approximate exposures—that is, more extreme than that with the true exposure. It remains more likely, however, that if true exposure has an affect it is stronger than the estimate with the approximate measurement.[4]
- Correcting for bias due to measurement error is occasionally possible, but it is almost never possible to regain lost power by statistical fixes. Having more accurate exposures (or several approximate ones) is the only way this can be done.
- Increasing the sample size does not reduce measurement error bias—although it does increase power.
- We have assumed here that it is the relation between the true exposure and health outcome that is of interest. Sometimes this is not the case. If you wish to use the study to predict risks in subjects using the same approximate measure of exposure and drawn from the same population, then the naive regression estimate $\beta_{OBSERVED}$ is appropriate.
- Multiplicative error (proportional to the true exposure), with lognormal distribution of true exposures, is common in environmental and occupational epidemiology. Here measurement error changes the shape of the regression—for example, from a quadratic curve to a straight line.
- Random error in numerical measurements of health outcomes—for example, lung function—unlike random error in exposure outcomes, does not in general cause bias in effect measures, although it diminishes power. However, misclassification in dichotomous outcomes—for example, disease or no disease—does cause bias in effect measures, as well as diminishing study power. The bias is towards the null if

misclassification is not dependent on the exposure (non-differential with respect to exposure).

The impact of random non-differential exposure measurement error on inference about the size of an effect is fairly clear once a causal relation is assumed—the true effect of exposure is most likely to be greater than that estimated. The impact of measurement error on the evidence that such a study brings on whether a causal relation exists is more problematic. The following points should be considered:

- You should usually be more cautious, if there is measurement error, in concluding from a negative study that no causal association exists. The reduced power implies that missing a true underlying association is made more likely.
- You should not use the (uncorrected) confidence interval for relative risk (or other measure of effect) to indicate the highest risk that is compatible with the data. For example, an uncorrected confidence interval for a relative risk of (0.80 to 1.25) suggests that relative risks in excess of 1.25 can be excluded. With exposure measurement error, however, the true uncertainty is greater, so that a higher relative risk is possible.
- The results reviewed here are less helpful in deciding how measurement error should influence assessment of the evidence for a causal relation brought by a positive association. It is clear that such error should not lead us to discount entirely an observed association of exposure with disease. On the other hand, it cannot be assumed that a small non-significant or even significant estimated effect of exposure would be larger and more significant in the absence of exposure measurement error. Such small associations could be due to chance or to uncontrolled bias or confounding, in which case they would be no larger, on average, in the absence of measurement error.

Further reading

Most textbooks on epidemiology discuss the effect of misclassification of exposure on estimates of relative risk, and some give methods for calculating and correcting for bias due to measurement error. The book by Armstrong *et al*[1] (no relation!) on exposure measurement in epidemiology is the most accessible source for most of the results discussed in this article. There are an enormous number of articles on this topic published recently in the journals. A simple MEDLINE search showed 185 articles published since 1981 with the words "measurement error" and "epidemiology" in the title or abstract—most in the past 5 years. For statisticians, Carroll *et al*[5] have provided an excellent review monograph with a particular focus on logistic regression, and with most examples from epidemiology. More limited but more accessible reviews are given by Armstrong[6] (especially useful for further discussion of and references on Berkson errors and errors in confounders) and de-Klerk *et al*[7] (especially useful for results on the impact of error on comparisons of risk in quantiles of the exposure distribution).

1 Armstrong BK, White E, Saracci R. *Principles of exposure measurement in epidemiology*. Oxford: Oxford University Press, 1992 .
2 Lagakos SW. Effects of mismodelling and mismeasuring explanatory variables on tests of their association with a response variable. *Stat Med* 1988;**7**:257–74.
3 Sorahan T, Gilthorpe MS. Non-differential misclassification of exposure always leads to an underestimation of risk: an incorrect conclusion. *Occup Environ Med* 1994;**51**:839–40.
4 Wacholder S, Hartge P, Lubin JH, *et al*. Non-differential misclassification and bias towards the null: a clarification. *Occup Environ Med* 1995;**52**:557–8.
5 Carroll RJ, Ruppert D, Stefanski LA. Measurement error in nonlinear models. New York: Chapman and Hall, 1995.
6 Armstrong BG. The effects of measurement errors on relative risk regressions. *Am J Epidemiol* 1989;**132**:1176–84.
7 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**:705–12.

Correspondence and editorials

Occupational and Environmental Medicine welcomes correspondence relating to any of the material appearing in the journal. Results from preliminary or small scale studies may also be published in the correspondence column if this seems appropriate. Letters should be not more than 500 words in length and contain a minimum of references. Tables and figures should be kept to an absolute minimum. Letters are accepted on the understanding that they be subject to editorial revision and shortening.

The journal also publishes editorials which are normally specially commissioned. The Editor welcomes suggestions regarding suitable topics; those wishing to submit an editorial, however, should do so only after discussion with the Editor.