TABLE A. Properties of the error distributions in a technical report on measurement error (2)

Error distribution	Type	$E(x\|z)$*	$\mathrm{Var}(x\|z)$*,†
I	Normal	z	h^2
II	Normal	z	k^2z
III	Gamma	z	k^2z
IV	Log-normal‡	wz	$w^2(w^2-1)\,z^2$

* $E(x|z)$ and $\mathrm{var}(x|z)$ denote the mean and variance of the measured exposures, x, conditional on the true exposures, z.

† h, k, and w are constant parameters that indicate the magnitude of the error.

‡ The log-normal distributions were parameterized so that the distribution of log x, conditional on z, followed a normal distribution with mean log z and variance s^2, $w = \exp(s^2/2)$.

IV). Second, for a specified true dose, misclassification to values that are far from z is less likely than is misclassification to values close to z. The assumption of these properties would seem reasonable for many exposures encountered in epidemiologic studies. This investigation did not address bias or systematic error, since these effects can be readily determined if the nature of the bias is known.

For every situation evaluated, I found that the estimated regression coefficients for both linear and log-linear models were biased toward the null, with the magnitude of the bias depending on several different factors related to both the error distributions and the true exposure distributions. I also found that the shape of the dose-response function could be distorted and that variances of estimated parameters could be underestimated.

Dosemeci et al. (1) recognize that the misclassification patterns in their examples are extreme, and although the examples probably do not represent common occurrences, they are important. Investigators need to know that exceptions to bias toward the null can occur, and they need to attempt to understand both systematic and random errors in dose measurements that are used in epidemiologic studies. It is also important, however, to recognize that it is systematic error that largely accounts for the trend reversals shown in Dosemeci et al.'s paper and that, for a wide range of situations involving only random or nonsystematic error (2), the true effect is, as expected, biased toward the null.

References

1. Dosemeci M, Wacholder S, Lubin JH. Does nondifferential misclassification of exposure always bias a true effect toward the null value? Am J Epidemiol 1990;132:746–8.
2. Gilbert ES. The effects of errors in the measurement of continuous exposure variables on the assessment of risks. Richland, WA: Battelle, Pacific Northwest Laboratory, 1988. (Technical report no. PNL-6578, available from the author).

Ethel S. Gilbert
Epidemiology and Biometry Department
Battelle, Pacific Northwest Laboratory
P.O. Box 999
Richland, WA 99352

THE AUTHORS CLARIFY AND REPLY

In reference to our recently published report (1), we would like to note that we made one statement which might be misconstrued by our readers and presented one example which is not as striking as we had implied. These oversights *do not* affect the main points of our work. We also add a paragraph which may help readers understand our point.

We stated our agreement with the point made commonly in the literature that nondifferential misclassification of exposure in a 2 × 2 table could not change the direction of the estimate of the odds ratio. However, a qualification must be added to this statement—the requirement, noted by Rogan and Gladen (2) in the context of screening for disease, that the sum of the specificity and the sensitivity be greater than 1. A simple counterexample to the general statement is the situation in which all exposed subjects are misclassified as unexposed and all unexposed subjects are misclassified as exposed; the observed odds ratio is then the reciprocal of what would be observed without misclassification. Rogan and Gladen's rule, generalized to K levels, is that a classification scheme must meet the following condition to be reasonable: for $i, j \leq K$ and $j \neq i$,

Pr (classifying a subject into category i|the subject is truly in category i) >
Pr (classifying a subject into category i|the subject is truly in category j);

this means that the scheme must classify individuals who are truly in category i correctly with a higher probability than it classifies individuals who are truly in any other category into category i. It also means that the largest element in each column of the misclassification matrix must be the diagonal element.

Our second example in table 2 (1, p. 747) has a misclassification scheme that does not meet the Rogan-Gladen rule, and thus it is not remarkable that the trend changes direction. Nonetheless, the first example in table 2 establishes that it is possible for a positive trend to appear to be negative with nondifferential misclassification, even when the Rogan-Gladen rule is met.

We would like to add a point which may clarify the basis of the main finding of our paper for

some readers. When the estimate of the odds ratio for a category at an intermediate level of risk compared with the baseline increases or changes direction, it is because of misclassification into or out of a category *other than the two categories being compared.* The fact that this does not happen with $K = 2$ may explain why our findings were overlooked in earlier work.

Since sending this letter to the *Journal*, we have received the correspondence printed above. We believe that Dr. Brenner (3) and Dr. Peeters (4) are independently making points that are similar to the one we made above. We agree with Dr. Brenner (3) that the phenomenon we demonstrated in table 1 of our paper (1, p. 747) behaves like a "selection bias" acting in the analysis stage and resulting from misclassification. Dr. Gilbert (5) points out that the anomalous estimates of trend in our table 2 (1) are "systematically" misleading because the measurements are negatively correlated with the true levels. Her point suggests that validation studies which might reveal a negative correlation between gold standard and error-prone measurements can be useful in alerting investigators to the possibility of serious distortions such as those we observed, even when the presumed misclassification matrix indicates that the Rogan-Gladen rule is satisfied.

In conclusion, we wish to reiterate our views, which we believe are in accord with those of the correspondents, that reversal of the direction of trend is rare but possible, and that caution needed in interpreting results in the presence of misclassification, even when it is known to be nondifferential.

REFERENCES

1. Dosemeci M. Wacholder S. Lubin JH. Does non differential misclassification of exposure always bias a true effect toward the null value? Am J Epidemiol 1990;132:746–8.
2. Rogan WJ. Gladen B. Estimating prevalence from the results of a screening test. Am J Epidemiol 1978;107:71–6.
3. Brenner H. Re: "Does nondifferential misclassification of exposure always bias a true effect toward the null value?" (Letter). Am J Epidemiol 1991 134:438–9.
4. Peeters PHM. Re: "Does nondifferential misclassification of exposure always bias a true effect toward the null value?" (Letter). Am J Epidemiol 1991;134:439–40.
5. Gilbert ES. Re: "Does nondifferential misclassification of exposure always bias a true effect toward the null value?" (Letter). Am J Epidemiol 1991, 134:440–1.

Mustafa Dosemeci
Sholom Wacholder
Jay H. Lubin
Epidemiology and Biostatistics Program
National Cancer Institute
Bethesda, MD 20892

RE: "ASSOCIATIONS ARE NOT EFFECTS"

Dr. Diana Petitti raised a number of questions about intentions and interpretations of causality in observational epidemiology in her recent editorial (1). I concur with her argument that the presentation of statistical results should be free of causal terminology. For example, analyses do not identify "protective effects" but rather yield risk ratios below 1.0. The use of "predictor" as a term for the independent variables in a regression equation is similarly arguable, except in the narrowest sense of statistical rather than causal prediction. I would go further in discouraging the separation of descriptive statistics from analytic statistics—all statistical analyses (including hypothesis tests and associated p values) are at best descriptions of data (and not always very good ones at that), with no implications regarding causality. Thus, there is a need to present statistical results in the most neutral and revealing manner possible, assisting rather than constraining the reader in drawing causal inferences from the data.

I strongly disagree, however, with Dr. Petitti's suggestion that only intervention studies can "claim to have as a purpose the assessment of the 'effect' of exposure" (1, p. 101). As Rothman argues, "Some scientists are reluctant to speak so blatantly about cause and effect, but in statements of hypothesis and in describing study objectives such boldness serves to keep the real goal firmly in focus and is therefore highly preferable to insipid statements about 'association' instead of 'causation' " (2, p. 77). Thus, in introducing a study, one should be honest and direct in stating the common goal of gathering data that will provide evidence regarding a causal association In the Discussion, the results should be evaluated from that perspective: To what extent do the study results support or refute the presence of a causal association? Claims that a study is merely "hypothesis-generating" or authors' statements that "our interest is in statistical associations, not causation" serve only to discourage critical evaluation of the applicability of the data to causal inferences. Studies that have a limited capability of addressing causal associations may still be worthy of conduct and publication, depending on the state of prior knowledge. Studies that were truly concerned only with examination of statistical patterns and that had no bearing on causality would rarely be of epidemiologic value.