Design and Interpretation of Studies of Differential Exposure Measurement Error

Emily White

From the Department of Epidemiology, University of Washington, and Cancer Prevention Research Program, Fred Hutchinson Cancer Research Center, Seattle, WA.

Received for publication April 26, 2002; accepted for publication November 13, 2002.

Differential exposure measurement error can have more adverse effects on estimates of exposure-disease associations than nondifferential measurement error, yet relatively little has been written about the design and interpretation of validity and reliability studies to assess differential measurement error. In this paper, a simple approximate equation is given for the effect of differential measurement error in a continuous exposure measure on the bias in the odds ratio. From this, it is shown that two parameters need to be estimated in validity/reliability studies in order to interpret the results in terms of the bias in the odds ratio in an epidemiologic study that will use the measure. The first is the correlation between the mismeasured and true exposure. The second is the differential bias (difference between cases and controls in the difference between mean measured and true exposure) relative to the true difference in exposure between cases and controls. It is shown that this latter parameter can be estimated in a method comparison study if one has a comparison measure that is unbiased or has nondifferential bias, so a perfect criterion measure is not needed. Researchers should consider measuring and reporting this parameter in validity/reliability studies when feasible.

epidemiologic methods; measurement error; reliability; validity

Validity and reliability studies of exposure measures that will be used in epidemiologic studies are important for several reasons. First, a validity or reliability study that precedes the epidemiologic study that will use the instrument can be used to decide whether the exposure measure is sufficiently accurate to be used in the epidemiologic study. Another use of validity/reliability studies is to estimate or correct for the impact of exposure measurement error on the results of the epidemiologic study, after that study has been completed (1–5).

Differential exposure measurement error between subjects with and without the disease under study is a major source of bias in epidemiologic studies. Sources of differential measurement error include differential recall or knowledge of exposures between cases and controls, the data collector’s knowledge of the subjects’ disease status, and the biologic effects of the disease or its preclinical phase. Although differential measurement error is a major concern in retrospective studies, because the subjects (and possibly data collectors) know both disease and exposure status, it could also occur in cohort studies, for example, if those with symptoms of disease change their “usual” diet before future diagnosis, or those with a strong family history of the disease have both more accurate knowledge of their family history and substantially higher future disease risk. Although nondifferential error generally biases the exposure-disease association toward the null, differential measurement error can cause the observed association between the measured exposure and the outcome to appear stronger than the true association or weaker than the true association, or it can lead to an association in the opposite direction, thus completely invalidating the results of the study (6, 7).

Despite this major concern about differential measurement error, most reliability and validity studies of exposures are not conducted to assess differential error, or they assess differential measurement error incompletely. For dichotomous expo-
sures, differential error can be assessed by the sensitivity and specificity of the exposure measure versus a criterion measure in a sample of cases and in a sample of controls, and simple formulas can be used to interpret these parameters in terms of bias in the odds ratio (6, 7). On the other hand, for continuous exposure measures, much of what has been written about differential error is mathematically complex (5) and therefore has not entered the mainstream of epidemiologic methods. However, if simplifying assumptions are made, then there is a simple approximate equation for the effect of differential measurement error on the odds ratio for a continuous exposure (6, 8). The primary aim of this paper is to show that the parameters in that equation are the key parameters that need to be estimated in a validity or reliability study to assess differential error in a continuous exposure and to discuss how one can design a validity/reliability study to estimate those parameters.

**EFFECTS OF MEASUREMENT ERROR ON THE ODDS RATIO**

**Definition of terms**

The exposure measurement error for an individual can be defined as the difference between his measured exposure and his true exposure. The true exposure is the agent hypothesized to cause the disease and would include a specified etiologic time period of interest. Differential exposure measurement error occurs when the measurement error differs between those with the disease and controls. A validity study assesses the relation of the mismeasured exposure to the true exposure in a population of interest. Such studies need a perfect measure of exposure (criterion measure), which is often not available or not feasible. A method comparison study refers here to a study in which a measurement method to be used in an epidemiologic study is compared with another, usually more accurate, but less than perfect method for measuring the same underlying true exposure. For example, a food frequency questionnaire might be compared with four 7-day food diaries for a group of subjects. Usually the more accurate measure cannot be used in the parent study because it is more burdensome to subjects (e.g., diaries), too costly (e.g., evaluation by experts), or not available for all subjects (e.g., medical or pharmacy records).

**A model of measurement error**

A common model of measurement error in a population is the following:

\[ X_i = T_i + b + E_i, \]

where

\[ \mu_E = 0 \]

and

\[ \rho_{TE} = 0. \]

This states that for a given individual \( i \), his observed measure \( X_i \) differs from his true value \( T_i \) by two types of measurement error. The first is systematic bias, \( b \), that would occur (on average) for all measured subjects in a population of interest. The second, \( E_i \), is the additional error that varies by subject.

For the population of interest, \( X, T, \) and \( E \) are variables with expectations (means over an infinite population) denoted by \( \mu_X, \mu_T, \) and \( \mu_E \), respectively, and variances denoted by \( \sigma^2_X, \sigma^2_T, \) and \( \sigma^2_E \). Because the average measurement error in the population is expressed as a constant, \( b \), it follows that \( \mu_E \), the population mean of the subject error, is zero. The model includes the assumption that the correlation coefficient of \( T \) with \( E, \rho_{TE} \), is zero; that is, that subjects with high true values do not have systematically larger (or smaller) errors than subjects with lower true values.

Two measures of measurement error are used to describe the validity of \( X \), that is, the relationship of \( X \) to \( T \) in the population of interest, based on the above model and assumptions. One is the bias or the average measurement error in the population:

\[ b = \mu_X - \mu_T. \]

The other parameter is a measure of the precision of \( X \), which is a measure of the variation of the measurement error in the population. One measure of precision is the correlation of \( T \) with \( X, \rho_{TX} \), termed the validity coefficient of \( X \). Under the above model, it can be shown (9) that

\[ \rho^2_{TX} = \frac{\sigma^2_T}{\sigma^2_X}. \]

\( \rho_{TX} \) is assumed to range between zero and one; that is, for \( X \) to be considered to be a measure of \( T \), \( X \) must be positively correlated with \( T \).

Bias and precision are independent concepts. A measure can be biased but perfectly precise; for example, an accurate scale that is calibrated to measure all subjects exactly 2 kg too light would have a bias of –2 kg and \( \rho_{TX} \) = 1.0. A measure can be unbiased, that is, yield the correct average in a population but lack precision. For example, a scale could be correct on average while overestimating the weight of some subjects and underestimating the weight of others. In this situation, \( b = 0 \), while \( \rho_{TX} \) would be less than one.

Figure 1 demonstrates the effect of measurement error on the distribution of \( X \) in a population. The bias in \( X \) causes a shift in the distribution of \( X \) compared with \( T \), such that the means differ by \( b \):

\[ \mu_X = \mu_T + b. \]

The imprecision of \( X \) (measured by \( \rho_{TX} \)) causes a greater variance or dispersion of the distribution of \( X \) compared with that of \( T \) (9):

\[ \sigma^2_X = \sigma^2_T / \rho^2_{TX}. \]

**Effects of differential measurement error on the odds ratio**

Measurement error is not an inherent property of an instrument but, rather, is a property of the instrument administered...
in a particular way to a specific population. Therefore, measurement error can differ between those with the disease of interest and a control group. Differential measurement error would have effects on the observable means and variances of the exposure variable within the disease and control groups (as above) and, more importantly, would bias the measure of exposure-disease association, for example, the odds ratio function (the odds of disease at each level of exposure vs. the odds at a reference level \( r \)).

Extending the measurement error model above to the two groups (\( D \) for disease group and \( C \) for control group) and using \( X \) to represent the exposure measure to be used in the epidemiologic study:

\[
\begin{align*}
X_{1D} &= T_{1D} + b_{1D} + E_{1D} \\
X_{1C} &= T_{1C} + b_{1C} + E_{1C}.
\end{align*}
\]

In this model, differential exposure measurement error occurs when \( b_{1C} \), the bias in the exposure measure in the control group, differs from \( b_{1D} \), the bias in the disease group, or when the precision of \( X_{1c} \) differs from that of \( X_{1D} \).

Figures 2 and 3 give an example of differential measurement error, specifically differential bias between cases and controls. In the figures, the true mean exposure in the disease group, \( \mu_{TD} \), is greater than the true mean exposure in the control group, \( \mu_{TC} \), which leads to a positive slope in the true odds ratio function (OR = \( f(T) \)). In this example, exposure is overestimated in the control group (positive bias) so the distribution of \( X_{1C} \) is shifted to the right relative to \( TD \), and the exposure is underestimated among those with disease (negative bias) so that the distribution of \( X_{1D} \) is shifted to the left relative to \( TD \). This leads the observable odds ratio curve (OR = \( f(X) \)) to cross over the null value of one: It indicates less disease risk with increasing exposure, rather than the true increasing disease risk.

One measure of differential bias is the ratio of the observed mean difference in exposure between cases and

---

**FIGURE 1.** The effect of measurement error (\( b \)) on the distribution of an exposure. \( T \), true exposure; \( X \), exposure measured with error; \( \mu_T \), population mean of \( T \); \( \mu_X \), population mean of \( X \). (Adapted from Armstrong et al. (6)).

---

**FIGURE 2.** Example of the effect of differential measurement error (differential bias, \( b \)) on the distributions of exposure among the disease group and the control group. \( TD \) and \( TC \) are the true exposures among the disease group and the control group, respectively; \( XD \) and \( XC \) are the exposures measured with error among the disease group and the control group, respectively; \( \mu_X \), population mean. (Adapted from Armstrong et al. (6)).
controls to the true mean difference in exposure, which will be termed factor $A$:

$$A = \frac{\mu_{X_1D} - \mu_{X_1C}}{\mu_{TD} - \mu_{TC}} = 1 + \frac{b_{1D} - b_{1C}}{b_{1D} - b_{1C}}$$  \hspace{1cm} (1)

If $A$ is positive, it gives the proportion over- or underestimation; for example, $A = 1.5$ means that $X_1$, the exposure measured with error, overestimates the true case-control mean difference in exposure by 50 percent. If $A$ is negative, then the mean difference in exposure between cases and controls has changed signs; that is, if the disease were truly associated with higher levels of exposure on average, it would appear, based on the use of measure $X_1$, that the disease was associated with lower mean exposure.

To make possible a simple equation for the effects of differential measurement error on the odds ratio, one needs to make certain assumptions. Results can be derived for unmatched case-control studies under the following assumptions: 1) $X_{1D}$ and $X_{1C}$ are modeled as above with for each group; 2) $T_D$ and $T_C$ are normally distributed with means $\mu_{TD}$ and $\mu_{TC}$, respectively, and the same variance, $\sigma_T^2$; and 3) $E_{1D}$ and $E_{1C}$ are normally distributed with mean zero and common variance, $\sigma_{E1}^2$. The last two assumptions mean that there is only differential bias, and the precision, $\rho_{TX}$, is the same for cases and controls. (The equations below also are based on the assumption that the only source of error in the odds ratio is measurement error in the exposure.)

The above assumptions lead to a logistic model for the probability of disease ($P$) as a function of true exposure $T$, with a true logistic regression coefficient $\beta_T$ (10):

$$\log \left[ \frac{P}{(1-P)} \right] = \alpha_T + \beta_T T,$$

where

$$\beta_T = (\mu_{TD} - \mu_{TC})/\sigma_T^2.$$

In this model, the odds ratio function can be expressed as a single parameter representing the true odds ratio for any $u$-unit increase in $T$:

$$OR_T = e^{\beta_T u}.$$

With measurement error in the exposure variable $X$, the assumptions also lead to a logistic model (8):

$$\log \left[ \frac{P}{(1-P)} \right] = \alpha_X + \beta_X X_1,$$

where

$$\beta_X = (\mu_{X_1D} - \mu_{X_1C})/\sigma_X^2.$$

Then, the observable odds ratio ($OR_\rho$) for a $u$-unit increase in $X$ can be expressed in terms of $OR_T$ (if $\beta_T \neq 0$) as follows:

$$OR_\rho = OR_T^{\beta_X},$$  \hspace{1cm} (2)

where $A$ is defined in equation 1 above. $OR_\rho$ differs from $OR_T$ because of the effect of differential bias (expressed by $A$) and lack of precision (expressed by $\rho_{TX}^2$). The effect of $\rho_{TX}^2$ is more predictable because it can only range from zero to one. However, as noted above, the factor $A$ can be any magnitude and either positive or negative. If $A \rho_{TX}^2$ is between zero and one, the observable odds ratio will be attenuated toward the null value of one compared with the true odds ratio; if $A \rho_{TX}^2$ is greater than one, the observable odds ratio will be biased away from the null; and if $A$ is

---

FIGURE 3. Example of the effect of differential measurement error (differential bias) on the odds ratio curve. OR = $f(T)$ is the true odds ratio function for exposure $T = t$ versus reference level $T = r$, and OR = $f(X)$ is the observable odds ratio function for exposure $X = x$ versus reference level $X = r$. (Adapted from Armstrong et al. (6)).
less than zero, the observable odds ratio crosses over the null value of one from the true odds ratio. For example, if the exposure measurement were perfectly precise \((\rho_{TX_i} = 1)\), then values of \(A\) of \(-0.5, 0.0, 0.5, 1.0,\) and \(1.5\) can be interpreted as biasing a true odds ratio of 4 to an observable odds ratio of 0.5, 1, 2, 4, and 8, respectively. When \(\rho_{TX_i} < 1\), then each of these observable odds ratios would be biased toward one.

Note that, when there is nondifferential measurement error and the assumptions above hold, equation 2 can be simplified (10, 11) to the following:

\[
\text{OR}_A = \text{OR}_{T}^{\rho^2_{TX_i}}.
\]  

This equation shows that, under nondifferential error, the observable odds ratio for any \(u\)-unit increase in \(X_i\) is attenuated toward the null value of one compared with the true odds ratio for a \(u\)-unit increase in \(T\) by the power \(\rho^2_{TX_i}\). The bias in the odds ratio (the attenuation) under nondifferential measurement error is not a function of the bias in \(X_i\) (because the same constant bias is added to exposure for both disease and control subjects). Thus, the focus of validity/reliability studies for nondifferential error is on estimation of \(\rho_{TX_i}\). This makes their design and interpretation substantially different from studies of differential measurement error, which need to assess \(A\) as well as \(\rho_{TX_i}\). Therefore, this paper focuses below on the measurement and interpretation of factor \(A\).

**DESIGN OF STUDIES TO MEASURE DIFFERENTIAL MEASUREMENT ERROR**

**A model for method comparison studies**

In a method comparison study, each person in a sample from a population of interest is measured twice, with \(X_i\), the measure to be used in the epidemiologic study, and \(X_2\), a comparison measure that measures the same true exposure. To assess differential measurement error, the study needs to be conducted on two samples: a sample of cases and a sample of controls.

For a given subject \(i\), two (continuous) exposure measurements, \(X_{i1}\) and \(X_{i2}\), are obtained. The simple model of measurement error above can be extended to \(X_{i2}\) measured in the two groups (disease and control):

\[
X_{i2D} = T_{iD} + b_{2D} + E_{i2D}
\]

\[
X_{i2C} = T_{iC} + b_{2C} + E_{i2C}.
\]

This model states that the second measure on the \(i\)'th subject in the disease group, \(X_{2D}\), is equal to his true value, \(T_{iD}\) (same true value as the first measure), plus the bias of the second instrument in the disease group, \(b_{2D}\), plus his error on the second measure, \(E_{2D}\). Similarly, \(b_{2C}\) is the bias of the comparison measure in the control group.

In a method comparison study, information is collected on \(X_i\) and \(X_2\) for each subject but not on \(T\). Such a study can yield estimates of the mean of \(X_i\) and \(X_2\) in the diseased group (\(\bar{X}_{1D}\) and \(\bar{X}_{2D}\)) and in the control group (\(\bar{X}_{1C}\) and \(\bar{X}_{2C}\)) and of the correlation between \(X_i\) and \(X_2\) in each group. The primary question is: How can a study be designed so that these estimable parameters can be used to estimate the parameters in equation 2?

**Selection of the comparison measure for estimation of differential bias in \(X_i\)**

Method comparison studies often cannot provide information on the bias in \(X_i\). Only if \(X_2\) is perfect or if \(X_2\) is an unbiased measure of \(T\) \((b_2 = 0)\), then:

\[
\hat{b}_1 = \bar{X}_1 - \bar{X}_2.
\]

Thus, a comparison measure \(X_2\) could be selected if it is correct on average (unbiased), even if it is not perfect.

As described above, a meaningful measure of differential bias is factor \(A\) (equation 1). The difference between the bias in \(X_i\) between cases and controls, \(b_{1D} - b_{1C}\), can be measured not only when \(X_2\) is perfect or \(X_2\) is unbiased \((b_{2D} = b_{2C} = 0)\) as above, but also when there is bias in the comparison measure \(X_2\) but it is nondifferential, that is, if \(b_{2D} = b_{2C}\). Then, under the simple measurement error model (6):

\[
(b_{1D} - b_{1C}) = (\bar{X}_{1D} - \bar{X}_{2D}) - (\bar{X}_{1C} - \bar{X}_{2C}).
\]

Therefore, if the comparison measure is carefully selected, a method comparison study can assess the differential bias between cases and controls in the measure of interest, \(X_i\). Ideally, to determine the accuracy of an instrument, measurements from the instrument would be compared with those from a perfect measure of exposure in a validity study. This would yield estimates of both the bias and the validity coefficient among cases and among controls. Almost all techniques for measuring and adjusting for differential measurement error assume that a perfect comparison measure of exposure is available (5). However, as shown above, when a perfect measure is not available or feasible, a study can assess differential bias in \(X_i\), if the comparison measure \(X_2\) is selected that is unbiased (among cases and among controls) or that can be assumed to have nondifferential bias (equal bias for cases and controls). In each case, the differential bias in \(X_i\) can be estimated by equation 4. For example, to assess differential bias in mother’s recall of a child’s birth weight (\(X_i\)) between mothers of children with leukemia and control mothers, \(X_1\) could be compared with medical records (\(X_2\)) among cases versus a similar comparison among controls. Although there may be error in the medical records, any bias is unlikely to be different between cases and controls. Thus, a good choice for a comparison measure is prospectively collected information recorded before disease diagnosis (or more accurately, prior to the period during which preclinical disease could influence exposures) if such information is available on at least a subset of cases and controls. Unfortunately, because such comparison measures are often not available, studies of differential measurement error are not always feasible.

**Analysis of differential bias**

The value of \(t\) from a two-sample \(t\) test on the variable \((X_{1i} - X_{2i})\) computed for each subject can be used to compute
a confidence interval for the difference in \( b_1 \) between the case and control groups. However, a judgment as to whether there is differential measurement error should not rely on statistical significance. A statistically significant difference between cases and controls would imply differential error, but a nonsignificant difference might still indicate an important degree of differential measurement error that was not significant given the sample size of the validation study.

**Interpretation of differential bias**

As outlined above, factor \( A \) is a useful parameter for describing the effect of differential bias on the mean exposure difference between cases and controls and, under certain assumptions, the effect of differential bias on the odds ratio (equation 2). Estimation of factor \( A \) (equation 1) requires an estimate of \((\mu_{TX} - \mu_{TC})\) as well as an estimate of \((b_{1D} - b_{1C})\) by equation 4. When \( X_2 \) is perfect, unbiased, or has nondifferential bias:

\[
\mu_{TX} - \mu_{TC} = \bar{X}_{2D} - \bar{X}_{2C}.
\]  
(5)

If the parent epidemiologic study has been completed, it may be more accurate to estimate \((\mu_{TX} - \mu_{TC})\), using the mean of \( X_1 \) in the disease group \((\bar{X}_{1D})\) and nondisease group \((\bar{X}_{1C})\) from the parent study, and the differential bias from the method comparison study:

\[
\mu_{TX} - \mu_{TC} = (\bar{X}_{1D} - \bar{X}_{1C}) - (b_{1D} - b_{1C}).
\]  
(6)

So a method comparison study in which \( X_2 \) is perfect, unbiased, or has nondifferential bias can be used to estimate \( A \) from equation 4 and equation 5 or 6. This can be used to understand the effect of differential measurement error on the odds ratio from equation 2, using an estimate of factor \( A \) and the true exposure. This paper shows that method comparison studies that use a comparison measure \((X_2)\) does not have differential bias can be used to estimate \( A \).

Another way to use equation 2 to interpret the effects of differential error is to estimate the true odds ratio after completion of the parent epidemiologic study. The true odds ratio \((OR_T)\) could be estimated from the observed odds ratio \((OR_O)\) by solving equation 2 for \( OR_T \) (8):

\[
OR_T = OR_O^{1/A'p_{TX}}.
\]  
(7)

Cautions about the use of such adjustment equations are discussed below.

**Estimation of \( p_{TX} \)**

To fully understand the effects of differential measurement error, one also needs to estimate \( p_{TX} \), preferably separately for the disease and control groups. Detailed discussions of the design and interpretation of validity/reliability studies to estimate \( p_{TX} \) or of the design of validation substudies used to correct for the effects of lack of precision of the exposure in the parent epidemiologic study have been presented (3, 6, 12–22). Briefly, \( p_{TX} \) can be estimated directly from a validity study in which the comparison measure \( X_2 \) is perfect or can be calculated under certain assumptions when the errors in \( X_1 \) and \( X_2 \) are uncorrelated. Under some situations, one could use the same study to estimate \( p_{TX} \) as well as \( A \), even when a perfect comparison measure is not available. If \( X_1 \) is not perfect but does not have differential bias and is more precise than \( X_2 \), and if the errors in \( X_1 \) and \( X_2 \) are uncorrelated, then information about both \( A \) and \( p_{TX} \) can be gained (6).

**DISCUSSION**

This paper discusses the design and interpretation of validity/reliability studies to assess differential error in an exposure measure \((X_1)\), with emphasis on differential bias rather than on differential precision. Specifically, the focus is on the measurement and interpretation of a parameter, \( A \), which can be interpreted as the effect of differential bias on the mean exposure difference between cases and controls. Furthermore, under simplifying assumptions, factor \( A \) can be interpreted in terms of the effect of differential measurement error on the odds ratio, if one has empirical or hypothetical estimates of the true odds ratio and of the validity coefficient, \( p_{TX} \) (the correlation between the imperfect exposure measure and the true exposure). This paper shows that method comparison studies that use a comparison measure \((X_2)\) does not have differential bias can be used to estimate \( A \).

To understand the effects of differential measurement error, one also needs to estimate \( p_{TX} \), separately for the disease and control groups. Differential precision also biases odds ratio estimates. If there was nondifferential bias in \( X_1 \) but \( p_{TX} \) differed between cases and controls, the shape of the odds ratio function could change. For example, the observable odds ratio function could be U-shaped when, in reality, disease frequency increases monotonically with increasing exposure (23). However, studies that report only differential measurement error in terms of the validity coefficient for each of the case and control groups do not measure the important effects of differential bias.

There are several limitations to the results presented. The first is that the results are based on a simple additive error model, with a constant additive error (bias) within each group and an additive subject error. If part of the error is proportional to the true value, that is, \( X_1 \) or \( X_2 \) captures only some fixed proportion of the exposure or similarly if the scales of \( X_1 \) and \( X_2 \) are in different units, then the results do not hold. For example, one could not generally use a biochemical measure as a comparison to assess bias or differential bias in a food frequency measure of intake because these would be on different scales.

The adjustment equations (equations 2 and 7) are presented to give researchers tools to interpret the magnitude of factor \( A \) in terms of the magnitude of its effect on the odds ratio, under simplifying assumptions. These assumptions would not necessarily hold in real applications. Specifically, these equations were based on the assumption of equal variances of \( T \) and \( E \) across cases and controls, which implies equal \( p_{TX} \) for the two groups. If one has a perfect measure of exposure, \( T \), in a validity study, one can test these assumptions directly. Otherwise, if the variance of \( X_i \) varies substantially between the two
groups (in the full epidemiologic study), or if the correlation coefficients between $X_1$ and $X_2$ in a validity/reliability study vary substantially between the two groups, this suggests that these assumptions do not hold. The normality assumptions used in the derivation of equations 2 and 7 mean that these equations become poorer approximations as the exposure deviates further from a normal distribution.

The simple adjustment equations given also do not take into consideration the imprecision in the estimates of $\rho_{XY}$ and $A$ (and for equation 7, the observed odds ratio). For example, the results from equation 7 when the observed odds ratio is slightly less than one would lead to a very different estimate of the true odds ratio than if the observed odds ratio were slightly greater than one. Also, these equations ignore the effects of covariates in the model. The bias in the covariate-adjusted odds ratio would depend on the multivariate measurement error structure of the main exposure and covariates, and an accurate adjustment of the odds ratio would generally require information from a validity study with perfect measures of the main exposure and covariates. There are several statistical methods that can be used for adjustment for differential measurement error that make fewer distributional assumptions, provide confidence intervals, allow other measurement error models, and/or allow adjustment for measurement error in the covariates as well (8, 15, 24–28) (see Thürigen et al. (5) for a review). Unless these techniques are used, the emphasis should be on understanding the possible degree of bias in the observed odds ratio due to the exposure measurement error rather than on the estimated true odds ratio.

Despite these limitations and cautions, the approaches given in this paper may provide added insight into the design and interpretation of studies of differential measurement error. If more such studies are conducted, this will help epidemiologists to understand which exposure-disease associations or study designs are particularly vulnerable to differential error.

ACKNOWLEDGMENTS

The author thanks Mari Nakayoshi for manuscript preparation.

REFERENCES

APPENDIX

Example of the Design and Interpretation of a Study of Differential Exposure Measurement Error

As an example, a reliability study was conducted using a nested case-control study within a cohort study to assess differential bias between breast cancer cases and controls in a retrospective food frequency questionnaire (FFQ) estimate of dietary fiber intake ($X_1$) (29). Women in the cohort study completed an FFQ in 1986 covering their diet in the past year. This prospective (prediagnostic) FFQ estimate of fiber intake was used as the comparison measure ($X_2$). Women who developed breast cancer over the next 2 years and selected controls completed another FFQ in 1989. That retrospective FFQ ($X_1$) asked about their diet in 1985, so it covered approximately the same time period as the 1986 FFQ. The results for mean grams of fiber intake are shown in appendix table 1.

Because there is reasonable certainty that any bias in $X_2$ is equal for cases and controls, the differential bias in $X_1$ can be estimated from equation 4 as:

$$b_{1D} - b_{1C} = (20.0 - 19.5) - (20.2 - 20.5) = 0.8 \text{ g},$$

and $A$ (equation 1) can be estimated using equation 5 as:

$$A = 1 + \frac{0.8}{19.5 - 20.5} = 0.2.$$

Thus, only 20 percent of the estimated true difference between cases and controls in fiber intake was observed on the retrospective questionnaire. This can also be interpreted to mean that, if the true odds ratio for dietary fiber and breast cancer were, for example, 0.25 for a 10-g increase in fiber intake, and if the validity coefficient of dietary fiber intake from the FFQ, $\rho_{TX_1}$, were estimated to be 0.6 (for both cases and controls for the retrospective questionnaire), then the differential measurement error would lead the observable odds ratio to be (by equation 2):

$$\text{OR}_O = 0.25^{0.2} \times 0.6^7$$

$$= 0.91.$$

This could be compared with the attenuation of the odds ratio due to nondifferential measurement error in the prospective study, which would lead to an observed odds ratio of 0.61 (from equation 3), if $\rho_{TX}$ = 0.6 for the prospective FFQ. Thus, in this example, a strong protective relation of fiber on risk of breast cancer (OR = 0.25) would be attenuated to an observed OR = 0.61 in the cohort study, but the relation would be almost completely obscured because of the differential measurement error in the retrospective study (OR = 0.91).

It should be noted that, in this example, some of the assumptions used in the derivation of equations 2 and 3 do not hold. The correlation coefficients between the retrospective and prospective FFQs differed between cases and controls ($r = 0.43$ and 0.64, respectively), which suggests that $\rho_{TX}$ was not equal for cases and controls, and dietary values are unlikely to be normally distributed. Nonetheless, parameter $A$ did approximate the difference between the odds ratio as observed in the actual retrospective study conducted (1.08 for the highest quintile of fiber) versus the risk ratio of 0.62 observed in the actual prospective study (29), (i.e., if OR$_O$ from equation 3 were substituted into equation 2).

APPENDIX TABLE 1. Comparison of fiber intakes by breast cancer cases and controls as reported prospectively in 1986 and retrospectively in 1989

<table>
<thead>
<tr>
<th>Mean fiber intake (g)</th>
<th>Cases (n = 300)</th>
<th>Controls (n = 602)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1989 FFQ* ($X_1$)</td>
<td>20.0</td>
<td>20.2</td>
</tr>
<tr>
<td>1986 FFQ ($X_2$)</td>
<td>19.5</td>
<td>20.5</td>
</tr>
</tbody>
</table>

* FFQ, food frequency questionnaire.