American Journal of Epidemiology

Vol. 189, No. 5
DOI: 10.1093/aje/kwaa020
Advance Access publication:
February 21, 2020

The *AJE* Classroom

Things Don't Always Go as Expected: The Example of Nondifferential Misclassification of Exposure—Bias and Error

Brian W. Whitcomb* **and Ashley I. Naimi**

* Correspondence to Brian W. Whitcomb, Department of Biostatistics and Epidemiology, School of Public Health and Health Sciences, University of Massachusetts Amherst, 715 N. Pleasant Street, Amherst, MA 01003 (e-mail: bwhitcom@umass.edu).

Initially submitted January 19, 2020; accepted for publication January 27, 2020.

The development of sophisticated methods for analysis of complex epidemiologic data is an exciting area of methodologic research. However, investigators frequently face more basic issues that challenge proper interpretation of study results. Nondifferential exposure misclassification is 1 such example and seems straightforward superficially. Complicated (and perhaps realistic) scenarios of nondifferential exposure misclassification, with correlated exposure and outcome measurement error, for example, have been described. However, even the simple case of nondifferential misclassification with a dichotomous exposure may be not so simple.

The term *nondifferential misclassification of exposure* is generally understood to bias estimates toward the null, but the meaning of this concept and the implications to research in practice are subjects of longstanding misunderstanding. In 2001, failure to acknowledge and accurately interpret the effects of exposure misclassification was described as commonplace in epidemiologic research (1), and certain points of confusion persist. In particular, the connection between the likelihood of exposure misclassification being related to the outcome and how misclassification occurs in a given data set and how bias from misclassification is related to error in a given study are prone to misperceptions.

In this article, we consider nondifferential misclassification of exposure in terms of likelihood vs. occurrence, and error versus bias to describe how misclassification can affect results and inferences. For simplicity, we restrict consideration to a prospective cohort study with dichotomous exposure misclassified (with probabilities sensitivity (SE) and specificity (SP)) and measured prior to incident outcomes. Also assuming a perfectly measured dichotomous outcome and no confounding or bias, we consider the following in regard to nondifferential misclassification of exposure:

1) What is meant by nondifferential?
2) How does nondifferential misclassification of exposure affect an individual study?
3) Counterintuitive effects of nondifferential misclassification of exposure.
4) How should nondifferential misclassification of exposure be addressed?

WHAT IS MEANT BY NONDIFFERENTIAL?

Nondifferential misclassification of exposure is described in vague and inconsistent terms in texts and published research that may create confusion about what, precisely, nondifferential references. Sometimes, the term refers to the degree of misclassification or the proportion of subjects misclassified. These descriptions imply nondifferential misclassification of exposure is the occurrence of misclassification of exposure that is the same in cases and noncases in a study. Examples where misclassification is exactly equivalent in data (as in Table 1, classification B) are often used to illustrate nondifferential misclassification of exposure.

However, interpreting nondifferential misclassification of exposure as exactly equal misclassification of exposure by outcome in a study is flawed. For one, it is impractical. Investigators would require complete data on both measured and true exposure status for all individuals in the study to determine nondifferential misclassification of exposure defined this way. This can be useful to illustrate misclassification; however, the notion of a data set that includes correct exposure, misclassified exposure, and outcomes for all individuals is implausible. More likely is this: Investigators design a study and measure exposure using an approach subject to misclassification that can reasonably be expected to be unrelated to outcomes. Nondifferential in this context means the probability of misclassification is the same in all study groups, for example, defined by outcome (2, 3). The probability of misclassification can be evaluated prior to data collection, on the basis of pilot data, study design, logic, and reason.

Table 1. Hypothetical Cohort Study Data With True Exposure Classification

Scenario[a]	Proportion Correctly Classified as Exposed		Exposure	Outcome				
	Cases	Noncases		No. of Cases	No. of Noncases	Total	Observed Risk	$\hat{RR}$
A	1	1	Actually exposed	100	900	1,000	0.100	2.00
			Actually unexposed	50	950	1,000	0.050	
B	0.90	0.90	Measured exposed	90	810	900	0.100	1.83
			Measured unexposed	60	1,040	1,100	0.055	
C	0.89	0.88	Measured exposed	89	792	881	0.101	1.85
			Measured unexposed	61	1,058	1,119	0.055	
D	0.92	0.87	Measured exposed	92	783	875	0.105	2.04
			Measured unexposed	58	1,067	1,125	0.052	

Abbreviation: RR, risk ratio.
[a] Classification A and varying scenarios with observed exposure as measured using a test with sensitivity = 0.9 and specificity = 0.1 Classifications B–D by proportion correctly classified on exposure status in cases and noncases.

HOW DOES NONDIFFERENTIAL MISCLASSIFICATION OF EXPOSURE AFFECT AN INDIVIDUAL STUDY?

The distinction between these 2 interpretations of nondifferentiality (i.e., equal occurrence vs. equal probability of misclassification) is more than just a matter of semantics. When dichotomous exposure is misclassified with the same exact SE and SP in cases and noncases, risks in the measured exposure groups appear more similar to one another, and magnitudes of measures of association will be underestimated. Simple formulas have been described to estimate the measure of effect had there been no misclassification (4). However, in practice, the nondifferentiality generally refers to the probability of misclassification rather than the realized misclassification proportion. In prospective studies with dichotomous exposure measured with error before and independent of outcomes, it is generally fair to conclude the study will be affected by nondifferential misclassification of exposure—exposure misclassification with equal probability in persons who go on to develop outcomes and those who do not. Nevertheless, it is highly unlikely that the occurrence of misclassification will be exactly equal between groups, due to chance.

Because nondifferential misclassification of exposure is defined in terms of probability, its realized impact on estimated measures of association in a given study is not completely predictable; instead, realized impact depends on manifestation (i.e., how misclassification occurs in that study). In a study affected by nondifferential misclassification of exposure but where misclassification is not exactly equal between cases and noncases, results can underestimate the true measure of effect (Table 1, classification C) if risks in observed exposure groups are more similar than when risks are compared between true exposure groups. This is the commonly understood case of bias toward the null. But, it is also possible for unequal exposure misclassification to occur between outcome groups, such that the result overestimates the truth (Table 1, classification D). It would be incorrect to describe this estimate as "biased away from the null." Rather, this particular study resulted in an overestimate, regardless of any bias. Bias is not the difference between a particular study result and the truth—that is error. Bias is a tendency—the difference between the true value of a parameter and the expected value based on (usually hypothetical) repetitions of a study. The potential for such counterintuitive findings was described at length by Jurek et al. (3). In practice, our challenge is the conduct and interpretation of results from single, nonhypothetical studies; this point remains underappreciated.

COUNTERINTUITIVE EFFECTS OF NONDIFFERENTIAL EXPOSURE MISCLASSIFICATION

The scenario in Table 1, classification D, provides a hypothetical example of study data that yield an overestimate despite nondifferential misclassification of exposure in probability. But, a single example does not provide information regarding how commonly such results may occur or why. Simulations are useful for this purpose. By generating data for many studies from the same underlying truth, simulations enable evaluation of errors in individual studies, and of bias, as the average error from repeated studies. So, we performed a simulation study generating 500 data sets, each with $n = 5{,}000$ for scenarios (s1–s25) with varying degrees of exposure-measurement error. Each data set included dichotomous outcome, dichotomous true exposure (X) that doubles outcome risk (i.e., risk ratio (RR) = 2.0), and measured exposure (Z) with nondifferential misclassification of exposure, with values determined with equal probability of misclassification (as reflected by SE and SP) for cases and noncases. Varying SE and SP, we calculated estimates for each of the 500 "studies" on the basis of true exposure ($\hat{RR}_X$) and measured exposure ($\hat{RR}_Z$), considered errors in individual studies. Geometric mean

Table 2. Simulation Study From 25 Scenarios of Varying Sensitivity and Specificity[a]

Scenario	Sensitivity	Specificity	Geometric Mean		Bias	% Studies $\hat{RR}_Z > \hat{RR}_X$[b]	Maximum[c]	
			$\hat{RR}_X$	$\hat{RR}_Z$			Max($\hat{RR}_X$)	Max($\hat{RR}_Z$)
1	1.00	1.00	1.99	1.99	0.00	0.0	2.86	2.86
2	1.00[d]	0.95[d]	2.00[d]	1.69[d]	−0.31[d]	1.6[d]	2.95[d]	2.69[d]
3	1.00[d]	0.80[d]	2.00[d]	1.36[d]	−0.64[d]	0.4[d]	2.98[d]	1.97[d]
4	1.00	0.60	1.97	1.22	−0.75	0.0	3.23	1.83
5	1.00[d]	0.40[d]	1.99[d]	1.16[d]	−0.83[d]	0.2[d]	3.15[d]	1.82[d]
6	0.95[d]	1.00[d]	1.99[d]	1.98[d]	−0.01[d]	46.2[d]	3.10[d]	3.00[d]
7	0.95[d]	0.95[d]	1.99[d]	1.66[d]	−0.33[d]	3.0[d]	2.99[d]	2.54[d]
8	0.95[d]	0.80[d]	2.00[d]	1.35[d]	−0.65[d]	0.2[d]	2.99[d]	1.99[d]
9	0.95	0.60	2.00	1.20	−0.79	0.0	3.05	1.76
10	0.95[d]	0.40[d]	1.96[d]	1.13[d]	−0.84[d]	0.2[d]	2.94[d]	1.59[d]
11	0.80[d]	1.00[d]	1.99[d]	1.94[d]	−0.05[d]	38.0[d]	3.03[d]	3.08[d]
12	0.80[d]	0.95[d]	1.98[d]	1.59[d]	−0.39[d]	4.0[d]	3.07[d]	2.52[d]
13	0.80	0.80	2.00	1.28	−0.73	0.0	3.20	1.79
14	0.80	0.60	1.99	1.14	−0.85	0.0	2.90	1.60
15	0.80	0.40	1.96	1.08	−0.88	0.0	3.27	1.55
16	0.60[d]	1.00[d]	1.99[d]	1.90[d]	−0.09[d]	37.2[d]	2.99[d]	2.98[d]
17	0.60[d]	0.95[d]	1.98[d]	1.51[d]	−0.48[d]	1.8[d]	2.92[d]	2.25[d]
18	0.60[d]	0.80[d]	1.98[d]	1.19[d]	−0.79[d]	0.2[d]	3.16[d]	1.75[d]
19	0.60[d]	0.60[d]	1.99[d]	1.07[d]	−0.91[d]	0.2[d]	2.81[d]	1.65[d]
20	0.60	0.40	1.99	1.00	−0.99	0.0	3.22	1.40
21	0.40[d]	1.00[d]	1.98[d]	1.85[d]	−0.13[d]	40.2[d]	2.77[d]	3.27[d]
22	0.40[d]	0.95[d]	1.96[d]	1.36[d]	−0.60[d]	1.8[d]	3.33[d]	2.17[d]
23	0.40	0.8	1.98	1.09	−0.89	0.0	2.87	1.58
24	0.40	0.6	2.01	1.00	−1.01	0.0	3.12	1.43
25	0.40	0.4	1.96	0.93	−1.03	0.0	3.25	1.32

Abbreviations: Max, maximum from 500 data sets; RR, risk ratio; $\hat{RR}$, estimated risk ratio.

[a] For each combination of sensitivity and specificity, 500 studies with $n = 5{,}000$ were generated with probability of exposure of 0.10 and a true risk ratio of 2. The maximum risk ratio using correctly measured exposure ($\hat{RR}_X$) is from an individual study and represents sampling variability. The maximum $\hat{RR}_Z$ represents a combination of misclassification and sampling variability.

[b] Percentage of 500 studies (each with $n = 5{,}000$) in which misclassification resulted in overestimates ($\hat{RR}_Z > \hat{RR}_X$).

[c] Maximum individual study estimate.

[d] Scenarios in which ≥1 of the 500 studies resulted in an $\hat{RR}$ using misclassified exposure that exceeded the estimate based on correctly measured exposure; bias was determined by comparison of geometric mean estimates from 500 data sets.

estimates were determined and compared to determine bias as the expected error across studies (Table 2).

For comparison, scenario 1 (s1) reflects no misclassification and thus no resultant bias; estimates vary due to sampling variability only. Bias was toward the null in all scenarios with nondifferential misclassification of exposure (i.e., s2–s25). With substantial misclassification, (e.g., s23–s25), biases were large and each of the 500 individual study results was an underestimate. However, in other scenarios, despite a bias toward the null, some individual study estimates based on misclassified exposure ($\hat{RR}_Z$) were further from the null than those based on true exposure ($\hat{RR}_X$). For example, with SE of 0.95 and SP of 1 (s6)—roughly corresponding to measurement of smoking status using serum cotinine—nondifferential misclassification of exposure resulted in a small bias toward the null; but, overestimates of effect were observed in 46% of studies ($n = 230$ of 500). These overestimates were generally small, but not always; in 1 of the data sets with SE of 0.4 and SP of 1 (s21), the RR_Z was 3.27, whereas the RR_X was 2.52.

HOW SHOULD NONDIFFERENTIAL MISCLASSIFICATION OF EXPOSURE BE ADDRESSED?

As shown in the simulation study, in the specific setting of a dichotomous exposure with nondifferential misclassification of exposure and errors uncorrelated with outcomes, the result is a bias toward the null, as observed

in a large number of study repetitions. Notably, even in circumstances where bias is always toward the null, our concern is interpretation of an individual study result in practice. Error in a study is not the same as bias as a tendency across studies, and there is no guarantee that an estimate from a study affected by nondifferential misclassification of exposure will be an underestimate. Probabilistic quantitative bias analysis has been described by various authors as an approach to provide insights about a plausible range of values for a measure of association. Using observed data and estimates for SE and SP, this kind of analysis can aid interpretation of findings by allowing for a degree of uncertainty in misclassification, as occurs in practice. The notion that nondifferential misclassification of exposure causes a bias toward the null and, therefore, results underestimate the true association is an oversimplification that is often not accurate. Taking additional complexities that can arise in other circumstances into account (e.g., dependent misclassification error and multilevel categorical exposures), caution is warranted to appropriately acknowledge exposure misclassification in practice when interpreting findings.

ACKNOWLEDGMENTS

Author affiliations: Department of Biostatistics and Epidemiology, University of Massachusetts Amherst, Amherst, Massachusetts (Brian W. Whitcomb); and Department of Epidemiology, University of Pittsburgh, Pittsburgh, Pennsylvania (Ashley I. Naimi).

Conflict of interest: none declared.

REFERENCES

1. Jurek AM, Maldonado G, Greenland S, et al. Exposure-measurement error is frequently ignored when interpreting epidemiologic study results. *Eur J Epidemiol*. 2006;21(12):871–876.
2. Last J. *A Dictionary of Epidemiology*. 4th ed. New York, NY: Oxford University Press; 2001.
3. Jurek AM, Greenland S, Maldonado G, et al. Proper interpretation of non-differential misclassification effects: expectations vs observations. *Int J Epidemiol*. 2005;34(3):680–687.
4. Copeland KT, Checkoway H, McMichael AJ, et al. Bias due to misclassification in the estimation of relative risk. *Am J Epidemiol*. 1977;105(5):488–495.