Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Epidemiology. Author manuscript; available in PMC 2015 November 01.

Published in final edited form as:

Epidemiology. 2014 November ; 25(6): 902–909. doi:10.1097/EDE.0000000000000166.

Potential sensitivity of bias analysis results to incorrect assumptions of nondifferential or differential binary exposures misclassification

Candice Y. Johnson[a,b], **W. Dana Flanders**[a], **Matthew J. Strickland**[a], **Margaret A. Honein**[b], and **Penelope P. Howards**[a]

[a]Rollins School of Public Health and Laney Graduate School, Emory University, Atlanta, GA

[b]National Center on Birth Defects and Developmental Disabilities, Centers for Disease Control and Prevention, Atlanta, GA

Abstract

Background—Results of bias analyses for exposure misclassification are dependent on assumptions made during analysis. We describe how adjustment for misclassification is affected by incorrect assumptions about whether sensitivity and specificity are the same (nondifferential) or different (differential) for cases and non-cases.

Methods—We adjusted for exposure misclassification using probabilistic bias analysis, under correct and incorrect assumptions about whether exposure misclassification was differential or not. First, we used simulated datasets in which nondifferential and differential misclassification were introduced. Then, we used data on obesity and diabetes from the National Health and Nutrition Examination Survey (NHANES) in which both self-reported (misclassified) and measured (true) obesity were available, using literature estimates of sensitivity and specificity to adjust for bias. The ratio of odds ratio (ROR; observed odds ratio divided by true odds ratio) was used to quantify magnitude of bias, with ROR=1 signifying no bias.

Results—In the simulated datasets, under incorrect assumptions (e.g., assuming nondifferential misclassification when it was truly differential), results were biased, with RORs ranging from 0.18 to 2.46. In NHANES, results adjusted based on incorrect assumptions also produced biased results, with RORs ranging from 1.26 to 1.55; results were more biased when making these adjustments than when using the misclassified exposure values (ROR=0.91).

Conclusions—Making an incorrect assumption about nondifferential or differential exposure misclassification in bias analyses can lead to more biased results than if no adjustment is performed. In our analyses, incorporating uncertainty using probabilistic bias analysis was not sufficient to overcome this problem.

Bias analysis (sensitivity analysis) has been proposed as an improvement over the qualitative descriptions of study limitations and potential sources of bias typically provided by investigators, in which potential effects of systematic error, and not only random error,

Corresponding author: Candice Y. Johnson, National Institute for Occupational Safety and Health, Centers for Disease Control and Prevention, 1090 Tusculum Ave MS R-15, Cincinnati, OH 45226. Tel: (513) 841-4454. cyjohnson@cdc.gov.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

are quantified.[1] The quantitative nature of these analyses allows a more transparent assessment of the potential direction and magnitude of bias and also guards against the tendency of investigators to favor causation over bias as the most likely explanation for observed results.[2,3] Some investigators have advocated greater incorporation of quantitative analyses for exposure misclassification and other forms of bias, [4–8] and many examples are now available in the published literature.[9–14]

Bias analysis for exposure misclassification involves identifying potential sources of misclassification, estimating bias parameters (e.g., sensitivity [Se] and specificity [Sp]) from validation studies or literature reviews, and using this information to adjust study results. Often, this adjustment is accomplished using simple algebraic manipulations of the contingency table. Probabilistic bias analysis extends this basic approach by allowing the investigator to assign a probability distribution to each bias parameter, sample randomly from the distribution, and perform the bias analysis repeatedly to produce a distribution of the adjusted measure of association. These probabilistic methods allow investigators to acknowledge uncertainty in choice of bias parameters and are more frequently used now that they are available in widely used software such as SAS, Stata, and Excel.[2,9,15]

There is discussion in the literature on choosing values or distributions of sensitivity and specificity for bias analyses of exposure misclassification.[2,3] Relatively less emphasis has been given to the importance of correctly specifying in the analysis whether misclassification is nondifferential or differential. In most studies, it is unclear whether nondifferential misclassification (sensitivity and specificity are the same for cases and non-cases) or differential misclassification (sensitivity and specificity differ between cases and non-cases) is the more appropriate assumption, unless internal validation data are available – in which case sensitivity and specificity can be estimated directly, albeit often with error. It has previously been shown that assuming nondifferential misclassification in a bias analysis when misclassification is truly differential can produce a result further from the truth than the unadjusted estimate.[16,17] Investigators might be hesitant to assume differential misclassification unless outcome-specific estimates of sensitivity and specificity are available or the investigator has other data specifying how they differ between cases and non-cases. In the literature, there are examples of bias analyses that use assumptions of nondifferential misclassification only[13,14] or both nondifferential and differential misclassification.[10–12] The potential effects of other types of incorrect assumptions have not been explored in depth.

The purpose of this study is to illustrate the potential sensitivity of bias analysis results to incorrect assumptions of nondifferential or differential misclassification of a binary variable. Through simulations, we create datasets with nondifferential and differential exposure misclassification, and we adjust for this misclassification using correct and incorrect assumptions about misclassification in probabilistic bias analyses. We then use data from the National Health and Nutrition Examination Survey (NHANES) to show how incorrect assumptions can affect the results of a bias analysis in an epidemiologic study. In the simulated datasets and in NHANES, both true and misclassified versions of exposure are known, and so we can evaluate the success of our adjustments.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

EXAMPLE 1: SIMULATED DATA

Methods

We began with 2 datasets with 10,000 simulated study participants each. In the first, we assigned a 10% disease prevalence and a 10% exposure prevalence among both cases and non-cases (odds ratio (OR) = 1.00). In the second, we assigned a 10% disease prevalence, 12.5% exposure prevalence among cases, and a 10% exposure prevalence among non-cases, to create a weak association between exposure and disease (OR = 1.29) (Table 1). No specific study design is implied in our simulations (i.e, cohort, case-control, cross-sectional), and we use the OR as our measure of association because it can be calculated for all study designs.

We then made 4 copies of each dataset, introducing a different type of exposure misclassification in each (2 nondifferential and 2 differential exposure misclassification scenarios). We used Bernoulli trials to randomly misclassify the simulated study participants under a given sensitivity and specificity. For the 2 nondifferential misclassification datasets ("exactly nondifferential" and "approximately nondifferential" misclassification, to be further discussed below), cases and non-cases were misclassified with Se = 0.90 and Sp = 0.90 independent of disease status. For differential misclassification, we created one dataset in which cases had higher sensitivity and specificity than non-cases (Se_{case} = 0.95, Sp_{case} = 0.95, $Se_{non\text{-}case}$ = 0.90, $Sp_{non\text{-}case}$ = 0.90) and one in which cases had lower sensitivity and specificity than non-cases (Se_{case} = 0.85, Sp_{case} = 0.85, $Se_{non\text{-}case}$ = 0.90, $Sp_{non\text{-}case}$ = 0.90). For convenience, misclassification scenarios in which cases have more accurate classification than non-cases will be referred to as "differential A misclassification" and scenarios in which cases have less accurate classification than non-cases will be referred to as "differential B misclassification".

We then adjusted for exposure misclassification in these datasets to determine if we could obtain a less biased estimate of the true OR, even if using incorrect assumptions about nondifferential or differential misclassification in the bias analysis. To adjust for misclassification, we used a common algebraic method that involves back-calculating expected cell counts for the correctly classified 2 x 2 contingency table given cell counts for the misclassified contingency table and estimates of sensitivity and specificity (Table 2).[3] Because the method uses contingency tables to adjust for bias, it can be used for other measures of association calculated from contingency tables, such as the risk difference or risk ratio.

We implemented a probabilistic analysis by specifying triangular distributions for sensitivity and specificity.[2,3] Triangular distributions were chosen over other options such as trapezoidal distributions because they allow specification of a single estimate with the highest probability of being chosen, rather than a range of values; this most closely matches our scenarios in which we designated a single value as most likely. Sensitivity and specificity values (as described below) were used as the modes of their respective triangular distributions. The maximum and minimum values for the triangular distributions were assigned to be +/− 0.05 of the mode for sensitivity and specificity. The distributions were truncated when necessary so all values fell between 0.5 and 1.00, inclusive. At each of

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

10,000 iterations, one value of sensitivity and one value of specificity were randomly chosen from the triangular distributions for cases and again for non-cases to calculate the misclassification-adjusted OR.

For each of the misclassified datasets in the bias analysis, we adjusted for misclassification using 4 assumptions: exactly nondifferential, approximately nondifferential, differential A, and differential B misclassification. We use the term “exactly nondifferential” to mean that we used identical values of sensitivity and specificity for the cases and the non-cases. “Approximately nondifferential” means that the sensitivity and specificity triangular distributions were the same for cases and non-cases, but sensitivity and specificity values were chosen independently from the same triangular distribution, so that values could differ between cases and non-cases by chance (i.e., numerically there might be differential misclassification).

We chose values of sensitivity and specificity for the adjustments to be as close to the true values as possible so as not to conflate the effects of choosing incorrect values of sensitivity and specificity with the effects of making incorrect assumptions about nondifferential or differential misclassification. However, under scenarios where incorrect assumptions were made, no “true” values of sensitivity and specificity existed, and so we had to make alternate assumptions. The values of sensitivity and specificity used in each bias analysis adjustment (see eTable 1) were calculated as follows.

Making correct assumptions—When a correct assumption was made, sensitivity and specificity were calculated directly from the simulated population. For nondifferential misclassification, sensitivity and specificity were calculated from the whole population (cases and non-cases together). For differential A and B misclassification, sensitivity and specificity were calculated separately for cases and non-cases, and these values were used in the analysis.

Making incorrect assumptions—When assuming nondifferential misclassification, sensitivity and specificity were calculated from the whole population (cases and non-cases together). When assuming differential A, sensitivity and specificity for non-cases were calculated directly from the non-cases; sensitivity for cases was assumed to be sensitivity for non-cases + 0.05, and specificity for cases was assumed to be specificity for non-cases + 0.05. When assuming differential B, sensitivity and specificity for non-cases were calculated directly from the non-cases; sensitivity for cases was assumed to be sensitivity for non-cases – 0.05, and specificity for cases was assumed to be specificity for non-cases – 0.05.

We calculated the ratio of the misclassification-adjusted OR to the true OR (calculated using the known true exposure status). We will refer to this metric as the ratio of odds ratios (ROR); ROR=1.00 signifies no bias. Results are presented as the median ROR and 95% simulation interval (SI). The 95% SI represents the 2.5^{th} and 97.5^{th} percentiles of the ROR distribution generated by simulation. If the sensitivity and specificity values chosen in the analysis produced a negative OR, we did not include it in the calculation of the OR, ROR, or 95% SI. All analyses were conducted in SAS version 9.3 (SAS Institute, Cary, NC) (see eAppendix 1 for SAS code).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Results

For both scenarios of OR = 1.00 and OR = 1.29, when correct assumptions about misclassification were made, results were on average unbiased or nearly unbiased (median ROR range = 1.00 to 1.01). The exception was the situation when misclassification was truly approximately nondifferential; in this scenario, no adjustment provided an unbiased result, given our assumptions (Table 3). There was little difference in results between scenarios in which the true association was null or weakly positive, although there was slightly less bias when the association was non-null. When misclassification was truly nondifferential, the assumption of either differential A or differential B misclassification produced biased results on average, but the biases were in different directions. When differential A was the true type of misclassification, all incorrect assumptions underestimated the magnitude of the association. When differential B was the truth, all incorrect assumptions overestimated the association. The 95% SIs included the true (null) value under some assumptions but not others.

Compared with the misclassified (unadjusted) ORs, the misclassification-adjusted ORs were closer to the truth only when the correct assumption was made for exactly nondifferential, differential A, or differential B misclassification (i.e., results were unbiased). Under all incorrect assumptions and for all assumptions when the truth was approximately nondifferential, the adjusted OR was farther from the truth than the unadjusted estimate.

EXAMPLE 2: DATA FROM AN EPIDEMIOLOGIC STUDY

Methods

In this example, we investigate an association between obesity and diabetes in NHANES using literature estimates of sensitivity and specificity, thereby combining the effects of making an incorrect assumption of nondifferential and differential misclassification with potential misspecification of the sensitivity and specificity distributions.

We included non-pregnant women aged 18 to 49 participating in NHANES between 1999 and 2010. NHANES uses a complex, multistage, probability sampling design to select participants from the civilian, non-institutionalized population of the United States.[18] NHANES participants complete an in-person interview during which they self-report height and weight. One or two weeks later, they visit a mobile examination center during which their height and weight are measured. Women with missing values for self-reported or measured height or weight were excluded from our analysis.

Obesity (exposure) was defined as body mass index ≥ 30 kg/m^2, calculated as weight in kilograms divided by squared height in meters. The reference group for all our analyses is non-obese women (body mass index <30 kg/m^2). We will refer to obesity status calculated from self-reported height and weight as “self-reported obesity” (misclassified exposure), and obesity status calculated from measured height and weight as “measured obesity” (true exposure). For the purposes of this example, we assume that measured obesity is measured without error.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Self-reported diagnosis of diabetes (outcome) was obtained by questionnaire. In our analysis, no distinction was made between type 1 and type 2 diabetes. Women who reported "borderline" diabetes were categorized as having no diabetes diagnosis. We excluded women with missing data on diabetes status. For simplicity, we assume that diabetes status was reported with no misclassification.

In this analysis, our purpose was to simulate a bias analysis for which values of sensitivity and specificity were obtained through literature review (external validation data). We searched the literature for studies presenting sensitivity and specificity for self-reported compared with measured obesity among adult females in the United States (see eAppendix 2 for literature search details). We excluded estimates from published NHANES data because our purpose was to approximate an adjustment for misclassification when internal validation data were unavailable.

Based on the results of the literature review, we created triangular distributions for Se and Sp (described further in Results). We conducted the bias analysis using the same probabilistic adjustment method previously described and using the same 4 assumptions: exactly nondifferential, approximately nondifferential, differential A, and differential B misclassification. Probabilistic adjustment for misclassification was conducted over 10,000 iterations to generate a distribution of the misclassification-adjusted prevalence odds ratio (POR). Results are presented as the median POR and 95% SI and as median ROR and 95% SI. For simplicity, in this example we did not take into account in the analysis the complex sampling design of NHANES and as such, the results should not be interpreted as being representative of the United States population.

Results

We identified 5 published studies[19–23] including estimates of sensitivity and specificity that met inclusion criteria (see eTable 2 for study details). Because most estimates of sensitivity were near 0.90, we chose this as the mode of the triangular distribution and assigned a minimum and maximum of 0.85 and 0.95 to allow for uncertainty. For specificity, we chose a mode of 0.97 based on the average of all estimates and the minimum (0.94) and maximum (1.00) values of the distribution based on the highest and lowest estimates obtained through literature review. No study provided diabetes-specific estimates of sensitivity or specificity.

In our NHANES population of 8,123 women, the POR between measured obesity and diabetes was 6.06 and the misclassified POR between self-reported obesity and diabetes was 5.53 (ROR=0.91). After adjusting for misclassification, misclassification-adjusted median PORs ranged from 6.39 to 9.37 (median ROR range = 1.05 to 1.55) (Figure 1). All adjustments overestimated the magnitude of the association.

The misclassification assumption producing estimates nearest to the truth was differential misclassification A (POR = 6.39 [95% SI = 5.24 – 7.96; ROR = 1.05 [95% SI = 0.87 – 1.31]). The true misclassification type in NHANES most closely resembled differential misclassification A, which explains why this estimate was the least biased. However, our estimates of sensitivity and specificity from literature review were inaccurate. The true

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

sensitivity was 0.09 higher in cases than non-cases (we assumed 0.05 higher) and the true specificity was 0.01 lower in cases than non-cases (we assumed 0.05 *higher*) (Table 4). This explains why the final estimate was slightly biased.

We repeated the analyses using sensitivity and specificity estimates abstracted from each of the 5 studies identified in the literature review separately (see eTable 3 for sensitivity and specificity values) to determine if any one study could have provided a more accurate estimation of sensitivity and specificity. Under the most accurate assumption (differential misclassification A), median misclassification-adjusted PORs ranged from 6.15 to 7.11 (median ROR range = 1.01 – 1.17) (Figure 2). All included the true POR of 6.06 in the 95% SI.

DISCUSSION

We presented examples in which adjustment for exposure misclassification was undertaken using various assumptions about differential and nondifferential misclassification. Using simulations and data from an epidemiologic study, we found that making incorrect assumptions about exposure misclassification can produce "bias-adjusted" results that are biased, and in some cases more biased than the unadjusted estimates.

Investigators are encouraged to be cautious when presenting and interpreting results from bias analyses because results are valid only if the assumptions used in the analysis at least approximate the truth.[2] In discussions of bias analysis in the literature, more emphasis has been given to choosing distributions of sensitivity and specificity for bias analysis than to choosing the correct assumption regarding nondifferential or differential misclassification. This is not surprising, given that outcome-specific sensitivity and specificity can be difficult to find in the literature, thus providing no evidence for suspecting if misclassification might be differential or not. Published validation studies providing such information would be useful contributions to the literature, facilitating the addition of bias analyses to epidemiologic studies by providing estimates of sensitivity and specificity as evidence-based starting points. However, there will be no guarantee that the values of sensitivity and specificity from one study will be generalizable to another.

In our examples, making an incorrect assumption about nondifferential or differential misclassification produced biased results. Taking uncertainty into account in the analysis by assigning probability distributions to Se and Sp was not sufficient to make up for the incorrect assumption, with the 95% SIs often not including the true value. Widening our triangular distributions might have allowed the 95% SI to cover the true value more frequently, although widening the distribution must be balanced against the ability to be sufficiently precise to interpret results from the analysis.

In our analyses, we considered two types of nondifferential exposure misclassification: "exactly" and "approximately" nondifferential misclassification. Exactly nondifferential misclassification (cases and non-cases have the exact same values of sensitivity and specificity) is rare in reality, because even if misclassification operates through a nondifferential mechanism of systematic error, random error will likely make sensitivity and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

specificity differ between cases and non-cases. We have referred to scenarios in which sensitivity and specificity differ between cases and non-cases by chance as “approximately” nondifferential misclassification, even though numerically this situation might appear to resemble differential misclassification. This is a potentially realistic scenario for many studies, in which the mechanism is nondifferential, but sensitivity and specificity likely differ by chance. In our simulated study results, when approximately nondifferential misclassification was the truth, none of our adjustment assumptions provided unbiased estimates of the magnitude of the association.

In the simulated data, when correct assumptions about nondifferential or differential misclassification were made, the results were on average unbiased even though some individual bias-adjusted ORs deviated substantially from the truth. In these simulations, we knew the true values of sensitivity and specificity in the population and used them in the analysis. If incorrect estimates of sensitivity and specificity were used to adjust for misclassification, results could be biased even when making the correct assumption about misclassification being differential or not. In the NHANES example, our estimates of sensitivity and specificity from literature review were not always accurate. As a result, none of the adjustments produced an unbiased estimate on average. However, the simulations based on 5 literature review estimates of sensitivity and specificity produced RORs ranging from 1.01 to 1.17 under the most appropriate assumption of differential A misclassification. This suggests that, if values of sensitivity and specificity are misspecified but still reasonably close to the true values, then results near to the truth can still be obtained.

We chose NHANES as the source of data for this analysis because both self-reported and measured versions of the obesity variable were available, allowing evaluation of the potential impact on results of making correct and incorrect assumptions. Because we used a specific dataset with its accompanying limitations, results from our analysis might not be generalizable to all bias analysis results. For example, because NHANES participants are likely aware they would be weighed and measured after self-reporting their weight and height, they might have reported their weight and height more accurately than persons who do not know they will be weighted, making it difficult to find suitable estimate of sensitivity and specificity from literature review. In addition, we used unweighted NHANES data in our analysis, and the applicability of literature estimates of sensitivity and specificity to the NHANES unweighted sample is unknown. Also, we used potentially misclassified self-reported diabetes status as the outcome and did not adjust for potential confounders such as age or sex, which could have affected our results. However, all investigators will face similar issues when conducting bias analysis for exposure misclassification. Our ability to produce fairly unbiased estimates when making a correct assumption about the type of misclassification in the face of these limitations should be encouraging.

In the absence of validation data (and even when validation data are available, because results of validation studies are themselves subject to error), the rationale for choosing nondifferential versus differential misclassification in a bias analysis is often left to the investigator’s perception of how misclassification occurred in the study. This commonly consists of a qualitative description of the possible sources of bias without presentation of evidence supporting the decision.[24] This is similar to the qualitative discussion of the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

direction and magnitude of bias that quantitative bias analysis is meant to guard against. Unfortunately, this situation is difficult to avoid because there is rarely sufficient information available to determine whether nondifferential or differential misclassification is most likely for a given study design and method of exposure measurement. For differential misclassification, the magnitude of the difference in sensitivity or specificity between cases and non-cases is typically unknown. Even if a certain misclassification process is strongly suspected (for example, assuming nondifferential misclassification in a prospective cohort study in which exposure is measured before disease occurs), there is no guarantee that this type of misclassification actually occurred in the study.[25] By chance, sensitivity and specificity could have differed between cases and non-cases, producing differential misclassification instead of nondifferential misclassification, or vice versa.[26] Factors aside from chance are also important. For example, when exposure categories are combined, differential misclassification can be produced even if the measurement error or misclassification process on the original variable was nondifferential.[16,25,27] When presenting results, investigators should clearly state that the results of their analysis are valid only if their assumptions were correct. Providing evidence or a rationale to support choice of assumptions would assist the reader in evaluating the likelihood that a correct assumption was made, although it will not be possible to know this with certainty.

An important role for bias analysis in epidemiologic studies is producing ranges of plausible estimates rather than providing a single bias-adjusted effect estimate as the final result. Without knowing whether misclassification was truly differential or nondifferential in our NHANES example, we would have no evidence for choosing the results of one assumption over the others as the most likely. However, we might conclude with some confidence that exposure misclassification does not account for the observed association, given that none of the 95% SIs contained the null value (POR = 1) under any assumptions. In addition, all of the results suggested that exposure misclassification produced a bias towards the null (i.e., the true POR was larger than the misclassified POR), which is valuable information to have even if the exact magnitude of the POR is uncertain. However, it should be noted that we tested only a small number of the possible combinations of sensitivity and specificity distributions and misclassification assumptions (e.g., we did not include scenarios in which sensitivity was higher for cases but specificity was higher for non-cases). Although our analyses correctly predicted that bias was truly toward the null, it is possible that other assumptions would not have produced the same results.

In this study, we presented examples demonstrating that making inaccurate assumptions about nondifferential or differential misclassification has the potential to produce biased results when adjusting for exposure misclassification. In our examples, using an incorrect assumption created more bias than using the unadjusted estimates, highlighting the fact that results from bias analyses do not necessarily represent an improvement over unadjusted estimates, but simply provide a range of plausible estimates under various assumptions. Investigators should recognize the possibility of making incorrect assumptions during adjustment for bias and consider reporting results based on various assumptions about misclassification (i.e., allowing sensitivity and specificity to be the same and different between cases and non-cases), recognizing that the type of misclassification occurring in the analysis is unknown and the choice of assumptions can affect results of the bias analysis.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Although this strategy might not provide a single point estimate as the result, bias analysis remains a useful method for providing plausible ranges of the effect estimate in the absence of information on exposure misclassification.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Lash TL, Fink AK. Semi-automated sensitivity analysis to assess systematic errors in observational data. Epidemiology. 2003; 14(4):451–458. [PubMed: 12843771]
2. Lash, TL.; Fox, MP.; Fink, AK. Applying Quantitative Bias Analysis to Epidemiologic Data. New York: Springer; 2009.
3. Greenland, S.; Lash, TL. Bias analysis. In: Rothman, KJ.; Greenland, S.; Lash, TL., editors. Modern Epidemiology. 3. Philadelphia: Lippincott Williams & Wilkins; 2008.
4. Phillips CV. Quantifying and reporting uncertainty from systematic errors. Epidemiology. 2003; 14:459–466. [PubMed: 12843772]
5. Steenland K, Greenland S. Monte Carlo sensitivity analysis and Bayesian analysis of smoking as an unmeasured confounder in a study of silica and lung cancer. Am J Epidemiol. 2004; 160:384–392. [PubMed: 15286024]
6. Greenland S. Multiple-bias modelling for analysis of observational data (with discussion). J R Stat Soc A. 2005; 168:267–306.
7. Maldonado G. Adjusting a relative-risk estimate for study imperfections. J Epidemiol Community Health. 2008; 62:655–663. [PubMed: 18559450]
8. Fox MP. Creating a demand for bias analysis in epidemiologic research. J Epidemiol Community Health. 2009; 63:91. [PubMed: 19141660]
9. Fox MP, Lash TL, Greenland S. A method to automate probabilistic sensitivity analyses of misclassified binary variables. Int J Epidemiol. 2005; 34:1370–1376. [PubMed: 16172102]
10. Jurek AM, Maldonado G, Spector LG, Ross JA. Periconceptional maternal vitamin supplementation and childhood leukaemia: an uncertainty analysis. J Epidemiol Community Health. 2009; 63:168–172. [PubMed: 18977808]
11. Jurek AM, Lash TL, Maldonado G. Specifying exposure classification parameters for sensitivity analysis: family breast cancer history. Clin Epidemiol. 2009; 1:109–117. [PubMed: 20865092]
12. MacLehose RF, Olshan AF, Herring AH, Honein MA, Shaw GM, Romitti PA. National Birth Defects Prevention Study. Bayesian methods for correcting misclassification: an example from birth defects epidemiology. Epidemiology. 2009; 20:27–35. [PubMed: 19234399]
13. Bodnar LM, Siega-Riz AM, Simhan HN, Diesel JC, Abrams B. The impact of exposure misclassification on associations between prepregnancy BMI and adverse pregnancy outcomes. Obesity (Silver Spring). 2010; 18:2184–2190. [PubMed: 20168307]
14. Stott-Miller M, Heike CL, Kratz M, Starr JR. Increased risk of orofacial clefts associated with maternal obesity: case-control study and Monte Carlo-based bias analysis. Paediatr Perinat Epidemiol. 2010; 24:502–512. [PubMed: 20670231]
15. Orsini N, Bellocco R, Bottai M, Wolk A, Greenland S. A tool for deterministic and probabilistic sensitivity analysis of epidemiologic studies. The Stata Journal. 2008; 8:29–48.
16. Flegal KM, Keyl PM, Nieto FJ. Differential misclassification arising from nondifferential errors in exposure measurement. Am J Epidemiol. 1991; 134:1233–1244. [PubMed: 1746532]
17. Gustafson, P. Measurement Error and Misclassification in Statistics and Epidemiology: Impacts and Bayesian Adjustments. Boca Raton, FL: Chapman & Hall/CRC; 2004.
18. Centers for Disease Control and Prevention (CDC), National Center for Health Statistics (NCHS). National Health and Nutrition Examination Survey Data. Hyattsville, MD: U.S. Department of Health and Human Serves, Centers for Disease Control and Prevention; 1999–2008.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

19. Krul AJ, Daanen HA, Choi H. Self-reported and measured weight, height and body mass index (BMI) in Italy, the Netherlands and North America. Eur J Public Health. 2011; 21(4):414–419. [PubMed: 20089678]

20. Johnson WD, Bouchard C, Newton RL Jr, Ryan DH, Katzmarzyk PT. Ethnic differences in self-reported and measured obesity. Obesity. 2009; 17(3):571–577. [PubMed: 19238143]

21. Brunner Huber LR. Validity of self-reported height and weight in women of reproductive age. Matern Child Health. 2007; 11:137–144.

22. Hussain S, Taylor M, Waltermaurer E, McCauley J, Ford DE, Campbell JC, McNutt LA. Computer-administered screening of reproductive-aged women for diabetes risk in primary care settings, feasibility and acceptability of such screening, and validity of risk assessments based on self-reported weight. Preventing Chronic Disease. 2007; 4(3):A54. [PubMed: 17572958]

23. Nieto-Garcia FJ, Bush TL, Keyl PM. Body mass definitions of obesity: Sensitivity and specificity using self-reported weight and height. Epidemiology. 1990; 1(2):146–152. [PubMed: 2073502]

24. Jurek AM, Maldonado G, Greenland S, Church TR. Exposure-measurement error is frequently ignored when interpreting epidemiologic study results. Eur J Epidemiol. 2006; 21:871–876. [PubMed: 17186399]

25. Wacholder S, Dosemeci M, Lubin JH. Blind assignment of exposure does not always prevent differential misclassification. Am J Epidemiol. 1991; 134:433–437. [PubMed: 1877603]

26. Jurek AM, Greenland S, Maldonado G, Church TR. Proper interpretation of non-differential misclassification effects: expectations vs observations. Int J Epidemiol. 2005; 34:680–687. [PubMed: 15802377]

27. Rothman, KJ.; Greenland, S.; Lash, TL. Validity in epidemiologic studies. In: Rothman, KJ.; Greenland, S.; Lash, TL., editors. Modern Epidemiology. 3. Philadephia: Lippincott Williams & Wilkins; 2008.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

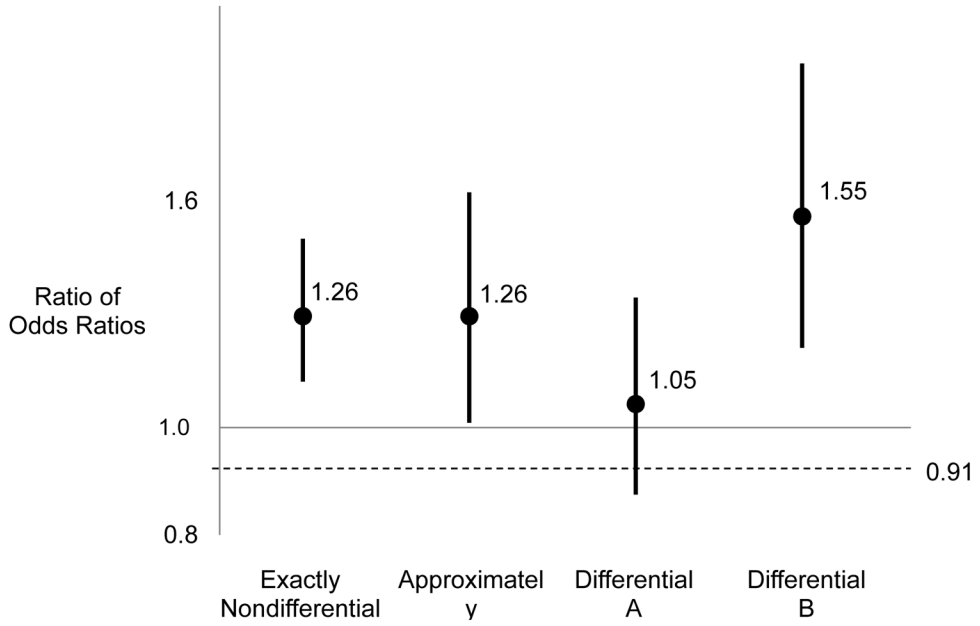

Figure 1.

Ratio of odds ratios (ROR) and 95% simulation intervals after adjusting for exposure misclassification under one of four assumptions about exposure misclassification in a study of obesity and diabetes in the National Health and Nutrition Examination Survey. The solid horizontal line indicates no bias (ROR = 1.00) and the dotted horizontal line indicates the result when using the misclassified (unadjusted) values in the analysis (ROR = 0.91).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

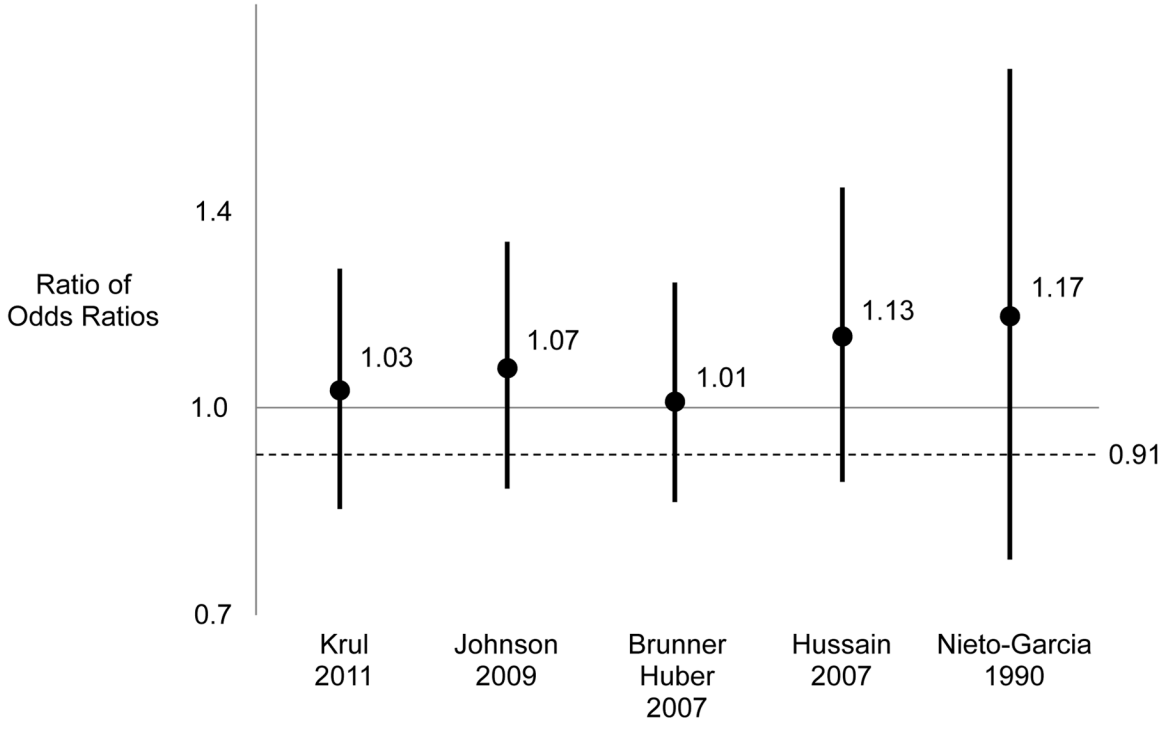

Figure 2.

Ratio of odds ratios (ROR) and 95% simulation intervals after adjusting for exposure misclassification in a study of obesity and diabetes in the National Health and Nutrition Examination Survey, using estimates of sensitivity and specificity from 5 studies, under the correct assumption of differential A misclassification. The solid horizontal line indicates no bias (ROR = 1.00) and the dotted horizontal line indicates the result when using the misclassified (unadjusted) values in the analysis (ROR = 0.91).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Contingency Tables for 2 Examples of True Associations Between Exposure and Disease.

	True Odds Ratio = 1.00 Exposure		True Odds Ratio = 1.29 Exposure	
Disease	**Yes**	**No**	**Yes**	**No**
Yes	100	900	125	875
No	900	8,100	900	8,100

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Relationships Between True (Correctly Classified) and Observed (Misclassified) Exposures.

	True Exposure		Misclassified Exposure	
Disease	**Yes**	**No**	**Yes**	**No**
Yes	A	B	$a = Se_1A + (1-Sp_1)B$	$b = (1-Se_1)A + Sp_1B$
No	C	D	$c = Se_0C + (1-Sp_0)D$	$d = (1-Se_0)C + Sp_0D$

Abbreviations: Se_i, sensitivity in cases (i = 1) and non-cases (i = 0); Sp_i, specificity in cases (i = 1) and non-cases (i = 0).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 3

Ratio of Misclassification-Adjusted Odds Ratio to True Odds Ratio over 10,000 Simulations when Making Correct and Incorrect Assumptions about Nondifferential and Differential Misclassification.

Simulation	Truth[a,b]	Adjustment Assumption[c]	True OR = 1.00	True OR = 1.29
			Median Ratio of Odds Ratio (95% Simulation Interval)[d,e]	
1	Exactly nondifferential	Exactly nondifferential[f]	1.00 (1.00–1.00)	1.00 (0.93–1.19)
2		Approximately nondifferential	1.01 (0.46–2.19)	1.01 (0.52–2.13)
3		Differential A	1.54 (0.89–3.16)	1.41 (0.84–2.90)
4		Differential B	0.43 (0.03–1.22)	0.53 (0.11–1.31)
5	Approximately nondifferential	Exactly nondifferential	1.30 (1.21–1.58)	1.24 (1.10–1.66)
6		Approximately nondifferential[f]	1.31 (0.68–2.82)	1.25 (0.69–2.62)
7		Differential A	1.84 (1.09–3.83)	1.65 (1.00–3.41)
8		Differential B	0.70 (0.14–1.74)	0.77 (0.30–1.77)
9	Differential A	Exactly nondifferential	0.56 (0.19–0.69)	0.65 (0.54–0.68)
10		Approximately nondifferential	0.56 (0.10–1.39)	0.65 (0.25–1.45)
11		Differential A[f]	1.01 (0.49–2.14)	1.00 (0.55–2.07)
12		Differential B	0.18 (0.02–0.65)	0.24 (0.02–0.77)
13	Differential B	Exactly nondifferential	1.95 (1.67–2.77)	1.77 (1.48–2.62)
14		Approximately nondifferential	1.97 (1.16–3.98)	1.78 (1.08–3.58)
15		Differential A	2.46 (1.53–4.94)	2.15 (1.36–4.31)
16		Differential B[f]	1.01 (0.40–2.23)	1.00 (0.49–2.14)

Abbreviations: OR, odds ratio.

[a] Nondifferential misclassifications: sensitivity for cases (Se_{case}) = specificity for cases (Sp_{case}) = sensitivity for controls ($Se_{non\text{-}case}$) = specificity for non-cases ($Sp_{non\text{-}case}$) = 0.90. Differential A: Se_{case} = Sp_{case} = 0.95, $Se_{non\text{-}case}$ = $Sp_{non\text{-}case}$ = 0.90. Differential B: Se_{case} = Sp_{case} = 0.90, $Se_{non\text{-}case}$ = $Sp_{non\text{-}case}$ = 0.95.

[b] Misclassified ORs for true OR = 1: exactly nondifferential = 1.00, approximately nondifferential = 1.14, differential A = 0.75, differential B = 1.42. Misclassified ORs for true OR = 1.29: exactly nondifferential = 1.12, approximately nondifferential 1.28, differential A = 0.88, differential B = 1.61.

[c] Exactly nondifferential: Se and Sp for cases and non-cases calculated from total population and are the same. Approximately nondifferential: Se and Sp for cases and non-cases drawn from the same distribution but do not necessarily have the same values; the modes of the distributions are the actual values of Se and Sp calculated from the total population. Differential A and B: Se and Sp for non-cases calculated directly from non-cases; for the correct assumption, Se and Sp for cases are the true values for cases; for incorrect assumptions, the modes of the Se and Sp distributions for cases are 0.05 higher (differential A) or lower (diferential B) than non-cases.

[d] Lower and upper bounds are the 2.5^{th} and 97.5^{th} percentiles of the ratio of odds ratio distribution generated by simulation.

[e] Percentage of simulations generating negative ORs: simulation 4 = 8% for OR = 1, <0.1% for OR = 1.29; simulation 8 = <0.1% for OR = 1; simulation 9 = 0.2% for OR = 1; simulation 10 = 0.2% for OR = 1; simulation 12 = 56% for OR = 1, 20% for OR = 1.29.

[f] Correct assumption.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 4

Contingency Tables and True Values of Sensitivity and Specificity of Obesity Classification for a Study of Self-Reported and Measured Obesity and Diabetes, Non-Pregnant Women Aged 18–49, National Health and Nutrition Examination Survey, 1999–2010.

	Self-Reported Obesity[a]		Measured Obesity[a]		Sensitivity[a,b]		Specificity[a,c]	
	Obese	Not Obese	Obese	Not Obese				
All participants Diabetes[d]	2,437	5,686	2,760	5,363	2,318/2,760	0.83	5,244/5, 363	0.98
Yes	197	89	213	73	195/213	0.92	71/73	0.97
No	2,240	5,597	2,547	5,290	2,123/2,547	0.83	5,173/5, 290	0.98

[a] Counts and proportions are unweighted.

[b] Proportion of individuals truly exposed who reported exposure.

[c] Proportion of individuals truly unexposed who reported not being exposed.

[d] Diabetes was self-reported by participants.