Misclassification of Genetic Susceptibility Biomarkers: Implications for Case-Control Studies and Cross-Population Comparisons


Environmental Epidemiology Branch, Division of Cancer Etiology, National Cancer Institute, NIH, Bethesda, Maryland 20892 [N. R., N. E. C. R. B. H.]; and Department of Epidemiology [W. F. S.], The Johns Hopkins University School of Hygiene and Public Health, Baltimore, Maryland 21205

Abstract
Phenotype and genotype markers of genetic susceptibility are of increasing interest in case-control studies of cancer. It is well established that bias to the odds ratio is caused by less-than-perfect assay sensitivity and specificity and varies with risk factor prevalence. As such, the observed variation in odds ratio between studies of genetic markers and cancer risk may be real, or may be attributed, in part, to variation in assay accuracy or in risk factor prevalence (e.g., prevalence differences between racial groups). The latter can be a particular concern when the prevalence of the “at risk” polymorphism in one or more populations is very high (e.g., >95%) or very low (e.g., <15%). For example, even very high sensitivity (e.g., 98%) can produce substantial bias to the odds ratio when the risk factor prevalence is high. Under some prevalence conditions, however, assays with only moderate accuracy are sufficient and result in minimal bias to the odds ratio. Understanding misclassification in the context of marker prevalence may help to explain disparate findings in the literature and should assist investigators in selecting markers that are appropriate for future studies.

Introduction
Genetic risk factors are of increasing interest in cancer causation. Genetic markers that have come under study in epidemiological investigations include polymorphisms for enzymes which activate and detoxify xenobiotics, such as NAT2 (1), CYP1A1 (2), CYP1A2 (3), CYP2D6 (4), CYP2E1 (5), and GST1 (6), and polymorphisms in genes of uncertain function, such as HRAS-1 rare alleles (7). Although laboratory assay accuracy generally improves with time, some degree of misclassification is almost inevitable, especially during the initial application of an assay in human populations. In addition, under field conditions common to epidemiological research, factors outside of the laboratory can introduce further misclassification, such as variation in subject compliance with phenotype protocols and variation in sample collection, processing, time-to-storage, and shipment conditions. Potential sources of error for metabolic phenotype assays include drug, dietary, or other exposures which inhibit or induce enzyme activity (8–12) and which may cause individuals who have a “rapid” phenotype to be misclassified as “slow,” and vice versa. Genotype can be misclassified due to failure to recognize a variant that contributes to the genotype of interest, false priming in polymerase chain reaction-based assays (13), and erroneous recognition of a pseudogene. Finally, coding errors at any point during sample collection, laboratory analysis, or data processing are almost inevitable and accumulate in the data.

The influence of misclassification on estimates of risk has been extensively addressed in the epidemiological literature (14–19). In particular, Flegal et al. (20) have described bias to the odds ratio as a function of sensitivity and specificity of risk factor categorization, risk factor prevalence, and the magnitude of the true risk. The purpose of this article is to apply these principles to case-control studies of genetic susceptibility biomarkers and cancer.

Methods
In case-control studies, the odds ratio is derived as a measure of the strength of an association between disease status and genetic risk factor status. Errors in defining genetic susceptibility will produce an observed odds ratio that is a biased estimate of the true odds ratio. To simplify the discussion, genetic risk factor status is defined as a binary variable (susceptible versus nonsusceptible). Assay accuracy is defined by Se, the proportion of susceptibles correctly classified, and Sp, the proportion of nonsusceptibles correctly classified as such.

Using the formula derived by Flegal et al. (20) we calculate values of the observed odds ratio (OOR) as a function of the true odds ratio (TOR), the true prevalence

1 The abbreviations used are: Se, sensitivity; Sp, specificity.

Received 9/30/92.
1 Presented in part at the Molecular and Biochemical Methods in Cancer Epidemiology and Prevention—the Path Between the Laboratory and the Population conference, September 23-26, 1992, Naples, FL.
2 To whom requests for reprints should be addressed, at NIH, NCI, Epidemiology and Biostatistics Program, Executive Plaza North, Rm. 418, Bethesda, MD 20892.
Bias = \frac{\text{OOR} - \text{TOR}}{\text{TOR} - 1}

When the true odds ratio is greater than 1, bias is negative if the true odds ratio is underestimated and is positive if the true odds ratio is overestimated. We assume that misclassification is nondifferential, i.e., errors occur similarly for cases and controls. Under the range of misclassification considered in this article, the true odds ratio is always underestimated for categorical variables and bias will, therefore, be expressed as a negative value.

The true odds ratio has been set at 10 in our examples. The relative amount of bias would be more substantial at higher odds ratios and of smaller magnitude at lower odds ratios (20) for most scenarios. Although observed odds ratios are generally less than 10 in epidemiological studies that report significant positive associations, these estimates of risk are based upon exposure and biomarker measures that are likely to be reported with some error. We show that under conditions common in case-control studies, even small degrees of misclassification can reduce observed estimates of a true odds ratio of 10 to levels more commonly found in epidemiological research.

In this article, we examine how bias varies for selected assay values of Se and Sp as the prevalence of the susceptible genetic risk factor in the controls ranges from 1 to 99%. For Se and Sp, we have selected a range of values from 100 to 70%, which are consistent with our experience, and that of other investigators, in the evaluation of several phenotype and genotype assays and cancer risk (11, 12, 21-26). The calculation of Se and Sp assumes the existence of a gold standard. When perfect gold standards do not exist, which is often the case with newly developed phenotype and genotype assays, values of Se and Sp are generally underestimated.

Results

Given a Se or Sp which is less than perfect, the degree of bias to the true odds ratio will depend on the prevalence of the genetic risk factor. In Fig. 1, we first consider a scenario where the true odds ratio is 10, Sp is perfect (100%), and Se varies from 70 to 99%. Under these conditions, an assay with a Se of 90% results in a small degree of bias when the prevalence of the risk factor is low (prevalence, 5%; bias, -0.05; observed odds ratio, 9.6) but substantial bias when the prevalence is higher (prevalence, 50%; bias, -0.5; observed odds ratio, 5.5).

Fig. 1 also shows that substantial variability in assay Se produces little difference in bias when the prevalence of the risk factor is low but substantial differences in bias as prevalence increases. For example, at 10% prevalence, assay Se of 99, 90, and 70% yield observed odds ratios of 9.9 (bias, -0.01), 9.1 (bias, -0.10) and 7.75 (bias, -0.25), respectively. In contrast, at 50% prevalence, assay Se over the same values results in observed odds ratio of 9.2 (bias, -0.09), 5.5 (bias, -0.50), and 3.25 (bias, -0.75).

As previously mentioned, the magnitude of this bias is generally smaller as the true odds ratio decreases (20). Consider a scenario where the genetic risk factor prevalence is 50%, and the assay Se = 90% and Sp = 100%. As the true odds ratio ranges over 20, 10, 5, and 2, the observed odds ratios are 7.3 (bias, -0.67), 5.5 (bias, -0.50), 3.6 (bias, -0.33), and 1.83 (bias, -0.17), respectively.

In Fig. 2, we consider bias when the assay Se is perfect and the assay Sp varies from 70 to 99%. Under these conditions, less than perfect assay Sp results in substantial bias to the true odds ratio when the prevalence of the risk factor is low and smaller degrees of bias as the prevalence increases. For example, when risk factor prevalence ranges from 1, 5, 10, and 50%, and the true odds ratio is 10, an assay with a Sp of 98% results in observed odds ratio of 4.0 (bias, -0.66), 7.5 (bias, -0.28), 8.6 (bias, -0.15), and 9.8 (bias, -0.02), respectively.

Fig. 2 also shows that variability in assay Sp produces substantially different degrees of bias when the prevalence of the genetic risk factor is low and negligible differences in bias at high prevalence. For example, at a 5% prevalence of the risk factor in controls, assay Sp of 99 and 90% yields observed odds ratio of 8.5 (bias, -0.16) and 4.1 (bias, -0.66), respectively. In contrast, at a prevalence of 50%, assay Sp over the same values would result in similar observed odds ratio of 9.9 (bias, -0.01) and 9.2 (bias, -0.09).

Discussion

The prevalence of the genetic risk factor should be considered when assessing the bias caused by risk factor misclassification. Nearly perfect Sp is critical when the risk factor prevalence is very low (e.g., <15%) (27) while nearly perfect Se is important when the risk factor is very high (e.g., >85%) (20). In contrast, relatively poor assay Se produces only a modest bias to the true odds ratio at low prevalence, while relatively poor assay Sp produces little bias at high prevalence. Several applications of these observations follow.

Small differences in genetic marker misclassification may lead to discrepant findings between studies if marker prevalence is either very high or very low. For example, substantial variation has been reported for the association of the extensive/intermediate debrisoquine metabolizing phenotype (prevalence, ~92%) and risk of lung cancer. Reported odds ratios have ranged from 13.2 (28), 5.9 (29), and 4.8 (30, to 1.6 (26), 1.5 (31), and 1.4 (32). It is possible that the variation in risk estimates may be explained, in part, by small degrees of marker misclassification. For example, one source of assay inaccuracy is unrecognized intake of quinidine class drugs in study subjects, which can cause extensive metabolizers to be misclassified as poor metabolizers due to enzyme inhibition (11, 12). Table 1 portrays a hypothetical case-control study of 200 lung cancer cases and 200 controls where the true odds ratio of developing lung cancer for extensive/intermediate metabolizers is 8.6. A phenotype...
Polymorphism prevalence

Prevalence range for selected polymorphisms

Bias to a true odds ratio (OR) of 10 and genetic marker prevalence, given assay specificity = 100% and varying levels of sensitivity (Se). *Bias = (Observed OR - True OR) / True OR. Observed OR calculated with formula presented in footnote 4. Genetic biomarker prevalence values (Refs. 2-33-37).

Cancer Epidemiology, Biomarkers & Prevention

polymorphism prevalence [prevalence, 90%; prevalence estimates from Weber (28)]. This phenomenon should be excluded as a competing explanation for differences in risk before attributing such differences to race.

The slow acetylator phenotype is a risk factor for bladder cancer in populations with occupational exposure to aromatic amines (1), while the rapid acetylator phenotype may be a risk factor for colon cancer (3, 38-40). A genotype assay for NAT2 which detects 90% of slow acetylator genotypes and correctly identifies all subjects with wildtype alleles would have Se = 90% and Sp = 100% in a study of bladder cancer. At a slow acetylator genotype prevalence of 50%, this assay would produce a substantial bias of -0.50 to a true odds ratio of 10. In contrast, the identical assay, used in a study of colon cancer, would have an Se of 100% and an Sp of 90%, producing a bias of only -0.09. Thus, the same assay applied to the same population with the same prevalence of NAT2 alleles would be appropriate to use in studies of colon cancer, but would be of marginal use in studies of bladder cancer under these conditions.

Polymerase chain reaction-based techniques have recently become available (13) to analyze genomic DNA samples for CYP2D6 alleles, the gene which codes for
the debrisoquine metabolizing phenotype. Assay Sp (proportion of poor metabolizers, the protective phenotype, correctly identified) ranges from 95 (23) to 75% (25), while Se is essentially 100% (23, 25). While the debrisoquine phenotype assay described in Table 1 (Se, 99%; Sp, 100%) would bias a true odds ratio of 8.6 to 4.8, a genotype assay with an Se of 100% and Sp as low as 50% would bias the odds ratio to only 8.3 under conditions of nondifferential misclassification. These results suggest that for this case-control study, assays with only moderate specificity would be adequate, while only assays with essentially perfect sensitivity would be acceptable.

Epidemiological studies should attempt to minimize biomarker misclassification (reviewed in Ref. 41). In general, however, some degree of genetic risk factor misclassification will almost inevitably occur in population-based studies. Evaluating this phenomena in the context of marker prevalence and true risk may help to explain disparate findings in the literature and should assist investigators in selecting markers that are appropriate for future studies.

References


Misclassification of genetic susceptibility biomarkers: implications for case-control studies and cross-population comparisons.


Updated version
Access the most recent version of this article at:
http://cebp.aacrjournals.org/content/2/4/299

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.