



# Springer-Verlag

## **CANCER PRECURSORS: Epidemiology, Detection, and Prevention**

**EDUARDO L. FRANCO**

McGill University  
Montreal, Canada

**THOMAS E. ROHAN**

Albert Einstein College of Medicine  
Bronx, New York

This book provides an overview of the progress made in the last few years on the epidemiology, detection methods, and preventive strategies for cancer precursors. Contributors to the 25 chapters are among the world's most knowledgeable scientists in the areas of molecular pathology, epidemiology, and control of cancer precursors. Their reviews of the topic are accessible to a large professional base that includes basic cancer researchers, clinical oncologists, pathologists, molecular biologists, epidemiologists, nurses, health professionals working on policy implications, and graduate students in cancer-related fields. Those who specialize on specific organ sites will find valuable information in the section covering site-specific precancers. Additional sections include chapters on timely topics, such as mechanisms of carcinogenesis, epidemiologic approaches to studying intermediate endpoints, the impact of measurement error, methods for processing biological specimens for molecular epidemiology studies, and the morphological classification of cancer precursors. The public health professional will find valuable overviews on screening and prevention strategies and the role of evidence-based medicine in judging the value of such strategies for national and international policy guidelines on cancer control.

**January 2002**

**430 pages / Hardcover**

**Price: US\$ 89.00**

**ISBN 0-387-95188-1**

**<http://www.springer-ny.com/detail.tpl?isbn=0387951881>**

**Springer-Verlag New York, Inc.**

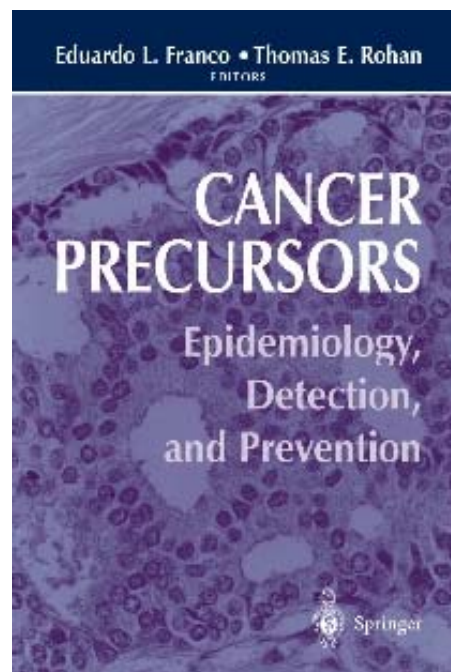
**175 Fifth Avenue**

**New York, NY 10010**

**USA**

**Call toll free: 1-800-SPRINGER [1-800-777-4643]**

**E-Mail: [service@springer-ny.com](mailto:service@springer-ny.com)**



### **TABLE OF CONTENTS:**

- Foreword (J. Fraumeni)
- Preface (E. Franco, T. Rohan)
- Introduction (T. Rohan, E. Franco).
- I. Biological Basis of Carcinogenesis:
  - 1. Mechanisms and Biological Markers of Carcinogenesis (R. Santella),
  - 2. Histological Aspects of Precancer (M. Cohen);
- II. Issues Related to Measurement and Interpretation of Cancer Precursors:
  - 3. Collection, Processing, and Analysis of Preneoplastic Specimens (N. Caporaso, J. Vaught),
  - 4. Intermediate Markers in Cancer Research (A. Schatzkin),
  - 5. Assessing Epidemiologic Relations and the Role of Measurement Error (E. Franco, T. Rohan);
- III. Site-Specific Precancerous Conditions:
  - 6. Upper Aero-digestive Tract (R. Sankaranarayanan, T. Somanathan),
  - 7. Esophagus (T. Vaughan),
  - 8. Stomach (P. Correa),
  - 9. Colon and Rectum (J. Baron),
  - 10. Anus (J. Palefsky),
  - 11. Liver (M. Sherman, I. Wanless),
  - 12. Pancreas (R. Hruban, C. Iacobusio-Donahue, M. Goggins),
  - 13. Lung (P. Marcus, W. Travis),
  - 14. Skin (R. Gallagher, T. Lee),
  - 15. Breast (T. Rohan, R. Kandel),
  - 16. Cervix (E. Franco, A. Ferenczy),
  - 17. Endometrium (E. Franco, A. Ferenczy),
  - 18. Ovary (J. Schildkraut),
  - 19. Vulva and Vagina (M. Madeleine, J. Daling, H. Tamimi),
  - 20. Prostate (J. Marshall, D. Wood),
  - 21. Bladder and Urinary Tract (K. Pohar, C. Cordon-Cardo);
- IV. Control of Cancer Precursors:
  - 22. Screening (A. Miller),
  - 23. Chemoprevention (G. Kelloff, C. Sigman),
  - 24. Evidence-based Policy Recommendations on Screening and Prevention (E. Franco, T. Rohan);
  - 25. Prospects (T. Rohan, E. Franco).

## Reprint of Chapter 5: Assessing epidemiologic relations and the role of measurement error

Eduardo L. Franco and Thomas E. Rohan

### Introduction

Understanding the role of cancer precursors in the natural history of cancer requires careful scrutiny of plausible epidemiologic relations involving remote exposures, intermediate endpoints, correlates of these variables, and, if available for study, the occurrence of cancer per se. To assess the pertinence of etiologic hypotheses in such investigations, one must specify *a priori* models that depict the relations among variables and their directionality, in an attempt to reproduce the causal pathway leading to disease. These variables include remote exposure variables - typically measured by questionnaire, such as diet, lifestyle, sexual activity - and genetic susceptibility traits, cancer precursor lesions or their surrogates (for instance, a cytologic test such as the Pap smear or a biomarker), and the occurrence of clinically evident cancer. Verification of the appropriateness of such models is based on some time-tested statistical analysis tools used in the context of epidemiologic study designs, such as case-control and cohort investigations.

The epidemiologic assessment of etiologic models in cancer causation relies on the verification that estimates of relative risk (RR) for the various exposure-outcome relations being probed -- which can be estimated directly in cohort studies as the incidence rate ratio or indirectly via the odds ratio (OR) in case-control studies -- change in some expected fashion upon adjustment by other relevant variables in the model. This provides an empirical demonstration of the putative relation specified in the hypothesized etiological model under consideration.

This chapter illustrates the theoretical basis for some common types of etiologic relations encountered in epidemiologic studies of cancer precursors and the expected outcomes from statistical data analysis that provide empirical evidence in support of the assumed relations (see chapter by Schatzkin for models involving intermediate endpoints that are relevant in controlled prevention trials). In practice, however, proper interpretation of the outcomes of these analyses must consider the role of misclassification of the remote exposures, genetic susceptibility markers, covariates, and also the cancer precursor itself. A few examples are considered in this chapter to illustrate the impact of misclassification, some of which are based on the literature on the role of human papillomavirus (HPV) infection in cervical carcinogenesis.

### Component relations in etiologic models in cancer epidemiology

Much has been written concerning disease causation and diagrammatic representation of causes in epidemiology. The interested reader should refer to previous work in this area [1-3]. Figure 1 shows some simple theoretical models that are typically entertained in cancer epidemiology studies (independence, confounding, and interaction), whereas others (causal pathway, correlates of outcome) are more rarely contemplated, despite their plausibility in natural history studies. V1 and V2 designate any two variables that are believed to mediate risk of an outcome (O) in the natural history of the disease. V1 and V2 could denote candidate exposure variables or intermediate endpoints. 'O' could denote a cancer outcome or even a cancer precursor, depending on the situation. The fact that these designations are interchangeable merely implies that the assumed relations do not necessarily represent entire etiologic models but, rather, components of more complex sets of relations that occur in the natural history of different types of cancer. The important aspect here is to have a range of expected statistical effects of such component relations in empirical research, that is, in case-control and cohort studies (and occasionally also cross-sectional studies) that collect information on the pertinent variables.

Independence of effects: The first model in figure 1 represents a set of relations that is probably the most commonly entertained scenario in cancer epidemiology: that of independence of causal effects. It simply states that V1 leads (causes) O but V2 also causes O independently. V1 and V2 could be two independent intermediate endpoints (for instance two distinct lesions originated from separate etiologic pathways) leading to cancer as non-sufficient, non-necessary precursors of O. Alternatively, V1 and V2 could be two exposures independently affecting the risk of a cancer precursor O, for instance, two separate dietary constituents increasing the risk of adenoma formation in the colon.

Confounding: Assessment of confounding bias and attempts to correct for it are among the most important tasks in the analysis of cancer epidemiology studies. Confounding is best defined as the distortion of the epidemiologic effect for a given exposure of interest by those of one or more extraneous factors (the confounding variables) which are associated with the disease and with the exposure. In consequence, the effect of the exposure of interest may be overestimated or underestimated because of the mixing of effects for the exposure and confounders [2]. In the simple example shown in figure 1, V1 is the true cause or precursor of O but V1 also leads to V2, which is only statistically related to O (hence the dotted arrow) secondarily via its relation with V1. It is said that the association between V2 and O is confounded by that of V1-O. For instance, V1 could be tobacco smoking and O

could represent the onset of oral leukoplakias, an oral cancer precursor (see chapter by Sankaranarayanan) and V2 could represent coffee drinking as a lifestyle variable collected by interview via the same questionnaire that elicited information on smoking (V1). In this case, a V1-V2 association exists because smokers tend to consume more coffee than non-smokers. Alternatively, V2 could designate a biochemical or serological marker of V1, the true precursor event leading to O. For instance, persistent HPV infection (V1) is the precursor event leading to cervical intraepithelial neoplasia (CIN) (O) (see chapter by Franco) but it also elicits an immune response that can be assessed serologically via detection of anti-HPV capsid antibodies. In the two examples, the associations for coffee drinking and a positive anti-HPV serology in each case are “confounded” by their respective models’ genuine causal variables, that is, smoking and cervical HPV infection.

The two examples of factors given above (coffee and HPV serology) are assumed not to play a causal or mediating role in the genesis of their respective outcomes, that is, their presumed associations with risk are entirely confounded by those of their respective confounders, smoking and HPV infection. In practice, however, an exposure could have its own genuine relation with risk and also be confounded. For instance, if coffee drinking had on its own a direct oral carcinogenic effect -- perhaps as an irritant or because of a hypothetical genotoxic constituent -- the estimated association would be the mixing of two effects: one for the direct causal effect of coffee drinking and another secondary to the confounded association with smoking.

Interaction: Cancer epidemiology investigations devote much energy also to assessing interaction among risk factors. As shown in figure 1, V1 and V2 interact to influence the overall risk of O in such a way that their joint effect on risk will be greater (if the interaction is synergistic) or lower (if the interaction is antagonistic) than that estimated by considering their independent effects added together (the scale in which the interaction is investigated is a key issue when investigating interaction). V1 and V2 could be two exposures such as tobacco and alcohol and O could denote oral dysplasias, also an oral cancer precursor. Alcohol drinking may enhance (promote) the carcinogenic effect of tobacco by a number of mechanisms, such as by increasing the permeability of oral cells to tobacco carcinogens, by inducing cell proliferation, and/or by leading to nutritional deficiencies that increase the oral epithelium’s susceptibility to the action of tobacco [4,5]. Interaction is also the mechanism to be entertained in gene-environment studies, in which a genetic susceptibility trait (V2) could affect the relation between a carcinogenic exposure (V1) and a cancer precursor O. Typical examples include the role of genetic polymorphisms in the major histocompatibility complex [6] or in the p53 tumour suppressor gene [7,8] potentially interacting with HPV infection to affect the overall risk of CIN.

Causal pathway: The causal pathway model in figure 1 is particularly relevant to studies that collect information on cancer precursors or their surrogates and also on remote as well as ‘downstream’ outcome variables. It states that V1 increases risk of V2, an intermediate endpoint or early event that with time (and possibly with the influence of additional cofactors) culminates with an increased risk of O. One of the most thoroughly studied examples is that of the genesis of CIN: high risk sexual activity (high number of sexual partners and an early age at first intercourse) leads to a higher risk of HPV infection, which in turn causes CIN (also described in the chapters by Schatzkin and Franco) [9,10]. Another example could be the model of a genetic polymorphism in V1 that leads to an increased risk of a given precancerous lesion (V2), which in turn places the individual at a higher risk of malignancy (O). For instance, a mutated APC gene confers a higher risk of adenomatous polyposis a precursor event that increases risk of colorectal carcinoma (see chapter by Baron). Although the model simply states three variables one can also envisage a more complex multi-step scenario in which several events or variables are in a linear arrangement leading to cancer development. The basic empirical assumptions (see below) also hold in such more complex causal pathway models.

Correlates of outcome: The last two types of epidemiologic relations described in figure 1 are frequently neglected in practice. They refer to situations in which the disease outcome leads to (i.e., causes) one or more conditions or states that may be mistaken as intermediate outcomes and not true correlates of disease. The first of the two theoretical examples shows V1 leading directly to O which in turn increases the probability of V2, a single correlate (consequence) of the outcome, whereas the second example shows O directly leading to two correlates of outcome, V1 and V2. Under the category of correlates of outcome one may consider scenarios in which disease development may lead to certain surrogate markers that occur as a consequence of the final outcome being analyzed. In the CIN example described above one might be interested in focusing on the relation between sexual activity (V1) as the exposure and HPV infection as the outcome (O), but data might also have been collected on occurrence of CIN (V2) in the same women. Another plausible situation that may occur in cancer precursor studies is the analysis of multiple histopathological and genetic characteristics of the outcome. Certain genetic instability markers may follow (and thus would be designated as correlates or consequences) rather than precede (which would imply that they are precursors) the development of the neoplastic lesion and should thus be placed ‘downstream’ from O in a causal pathway diagram.

### **Empirical demonstration of different types of epidemiologic relations**

The process of analyzing the epidemiologic data collected in observational investigations of the etiology of cancer precursors and cancer requires that the pertinence of the above relations be verified in specific statistical

analyses using suitable regression models or by constructing contingency tables with the Mantel-Haenszel technique. The investigator should have a *a priori* expectation that some of the relations depicted in figure 1 may be influencing the joint distribution of observations across all variables on which information was collected in the study. As part of the analytical strategy such relations can be verified by computing crude (unadjusted) estimates of epidemiologic effect (the RR) for the associations between each of the suspected etiologic variables or key events, that is, V1 or V2, and the outcome of interest in a given situation (O). A second set of analyses, adjusting each of the associations for the remaining covariate (i.e., V1-O for V2 and V2-O for V1) will aid the investigator in deciding whether or not there is empirical evidence in support of the suspected relation. In some circumstances, computing the crude estimates followed by the adjusted ones and then comparing the two sets for an indication of change is also called mediation analysis (see chapter by Schatzkin).

Table 1 shows the expected results from these two sets of analyses under the assumption that the underlying relations follow each of the theoretical etiologic models described in figure 1. These results were obtained by simulation using single-cycle, one- or two-state Markov transition models specifying a range of plausible statistical association scenarios that are likely to occur in practice, all of which assume that the relations under study are not absolute and that V1, V2, and O are binary variables (that is, exposed/non-exposed, positive/negative, or attribute present/absent). A ‘++’ indicates that the relation is identified as present of moderate or high magnitude (as estimated by the RR) and statistically precise (as judged by the width of the 95% confidence interval) depending on the size of the study. A ‘+’ indicates that the relation is appreciably lower in strength than the one indicated by ‘++’ and a minus sign ‘-’ indicates that the association is non-identifiable, that is, an unity RR. Modeling was done with the EpiMod1M software utility [11]

Independence of effects: In the first situation in table 1 it can be seen that the magnitude of the crude associations V1-O and V2-O are unaffected upon mutual adjustment. Were these results seen in practice, the observer should conclude that it is unlikely that the pathway leading to O from V1 is mediated by V2 or vice versa. For instance, if V1 and V2 were two dietary variables associated with risk of colonic adenomas (O), then in the absence of other information, one would be tempted to conclude that they are possibly affecting the same initiating or promoting step in the development of lesions.

Confounding: The confounded relation (table 1) specifies that V1 is the confounder (the actual cause of O) and V2 is the variable with no causal effect of its own but whose association with O is confounded by that of V1-O. The stronger statistical association for V1-O than for V2-O is the first indication that this could be the case, since V2-O is a spurious, indirect relation. Adjusting V1-O for V2 has no effect on the magnitude of the RR but the V2-O association disappears after adjustment for V1, which provides empirical evidence for the hypothesis of a confounded relation with a zero effect. In the examples discussed above, the crude association between coffee drinking (V2) and oral leukoplakias (O) would not persist upon adjustment for smoking (V1), the true risk factor. Similarly, the anti-HPV serological response described above in the context of a CIN study will probably be a good statistical marker of the presence of lesions but adjustment of such an association for persistent cervical HPV infection -- should this variable be available and properly measured (see below “Effect of measurement error”) -- will make it disappear.

The confounding scenario assumed in table 1 is one of a null effect for V2. It is conceivable, however, that V2 could have its own mediating effect on the genesis of O. In that case, the adjusted RR will not necessarily disappear but reveal a net effect that represents the average V2-O association among those exposed and not exposed to V1.

Interaction: If V1 and V2 interact to produce a modified joint effect on the risk of O the results will be consistent with the range shown in table 1, which would not be helpful in interpreting the nature of the relation. However, mutual adjustment is not the appropriate strategy for assessing interaction because the resulting adjusted RRs are simply summary measures of the net effect of these variables and would hide the critical evidence that the strength of the association for either V1-O or V2-O varies according to the categories of the other variable. The appropriate strategy, should interaction be suspected, would be to conduct a stratified analysis that would reveal the association of V1-O in each of the categories of V2 and vice-versa. This could also be accomplished in a regression model that included a cross-product term for V1xV2 in addition to the main effects for V1 and V2. Any of these approaches would reveal effect modification and provide empirical evidence in support of interaction. In the example of tobacco (V1) and alcohol (V2) as risk factors for oral dysplasias (O) the latter analytical approaches (i.e., stratification or fitting a cross-product term) would be equally suitable. On the other hand, in a case-control study collecting information on genetic polymorphisms it is possible that the investigators might opt to conduct the genetic assays only among cases because of the expense of carrying them out among all subjects. In such gene-environment studies, case-only analyses of the association between an environmental exposure and a genotype treat case subjects without the susceptibility genotype as the control group and the case subjects who were not exposed to the environmental factor serve as the referent group. This type of study design (the case-only design) assumes that the exposure and genetic attribute occur independently in the population and that the disease is rare. The OR thus obtained is an adequate estimate of the multiplicative interaction effect that would have been measured if the genetic information were available for all subjects [12-14].

**Causal pathway:** The simple pathway model in figure 1 specifies that V1 is the remote variable mediating risk of O via V2, the intermediate endpoint. In consequence, a stronger association has to be expected for V2-O than for V1-O, but the latter may be moderate or even large, depending on the strength of the V1-V2 and V2-O relations. More important, however, is the fact that adjustment of V1-O for V2 will necessarily make the association disappear, whereas adjustment of that of V2-O for V1 has no effect (table 1). In the CIN example described above, HPV infection (V2) is in the causal pathway between sexual behavior (V1) and cervical lesion outcome (O) [9]. Empirical evidence supporting this model was obtained by Schiffman et al. [15], who showed that the OR for sexual activity is considerably attenuated upon adjustment for HPV infection (see chapter by Schatzkin).

It is noteworthy that the fact that the V1-O association was “adjusted away” by V2 may be erroneously interpreted as being caused by confounding, a scenario that is suspected far more frequently by cancer epidemiologists when conducting data analysis. This underscores the importance of considering plausible biological mechanisms that could influence the natural history of the target disease so that an appropriate *a priori* hypothesis of a causal pathway may be formulated beforehand.

**Correlates of outcome:** Finally, analysis of the relations in the two situations depicted in figure 1 where correlates of outcome are to be suspected may generate some unexpected results if the underlying causal relations are not considered. In the first of the two examples, the observer may decide that adjustment of the V1-O association for V2 is justifiable perhaps because of a mistaken assumption that V2 would be an intermediate endpoint preceding the outcome (O) in the natural history of the disease (as in the causal pathway described above). The resulting adjusted estimate would show an attenuated association and could lead to the erroneous interpretation that V2 would be in the causal pathway from V1 to O, that is, as a mediator of the effect of V1. A suitable example would be the model of sexual activity as a remote risk factor (V1) leading to HPV infection (considered here as O, the dependent variable in the analysis) and ultimately to CIN (V2). Adjustment of the association between the sexual activity variable and HPV infection for the presence of CIN would lead to spurious results, since CIN is a direct correlate of HPV infection, taken here as the outcome. Likewise, the second example of correlates of outcome would also lead to unpredictable results that could lead to an erroneous interpretation of the underlying biological model. In this more extreme case, the analysis implies that the observer inverted the directionality of the causal relation and undertook an inappropriate adjustment.

In practice, the empirical validation of the types of expected etiological relations described in figure 1 would require accommodating the role of covariates in more complex relations that could involve joint confounding and other potential sources of biases. Nevertheless, the basic premise of how the stated relations would be verified empirically through statistical analysis should follow the range of expectations mentioned above. Many consumers of information from epidemiologic studies tend to consider adjusted RR estimates in cancer etiology studies as inherently bias-free. As shown above, covariate adjustment has a clear role in mitigating the effects of confounded relations and in revealing mediation in causal pathways but one should refrain from generalizing this strategy indiscriminately without considering the underlying biological models.

### **Effect of measurement error**

The etiologic models depicted in figure 1 represent relations in the source populations, that is, structural relations that occur among the variables of interest before a study sample is taken and exposures, endpoints, and outcome are measured in an investigation. Therefore, such relations are unaffected by random and systematic errors that occur in empirical research. A major caveat in the interpretation of the data analysis results described in the preceding section is the effect of measurement error. The effects of exposure and covariate misclassification are a long-standing interest of statistical methodologists [16-18]. Misclassification bias is among the most serious problems hampering the validity of much of the epidemiologic research on risk factors for cancer and other chronic diseases. Those attempting to assess the effects of dietary or environmental factors on cancer risk are well aware of the difficulties involved. Epidemiologic common sense has it that improper ascertainment of an exposure will bias its RR estimate generally towards the null, if the misclassification is random and non-differential with respect to the outcome (being a case of the disease or not). If the measurement error is not random or non-differential with respect to the outcome the direction and degree of the bias are difficult to predict *a priori*.

The following paragraphs describe typical situations that are affected by measurement error of intermediate endpoints or outcomes in epidemiologic studies of cancer precursors. The examples that are described to illustrate specific scenarios of misclassification were produced by simulation of hypothetical cohorts of size 100,000 using single-cycle, one- or two-state Markov models based on a wide range of plausible parameters for transitions. Correction for the presumed misclassification is sometimes feasible if knowledge exists about the performance of techniques used to measure exposure or outcome variables.

**Effect on the prevalence of the endpoint:** Measurement error in the endpoint under study can have a marked biasing effect on the presumed prevalence of the condition [9]. Table 2 shows the effect of varying test sensitivity and specificity on the presumed prevalence of cancer precursors or intermediate endpoints in three hypothetical populations with prevalences that vary from 0.1% to 10%. As shown, when the endpoint is rare the presumed prevalence can be a gross overestimation of the true prevalence. The bias always results in an overestimation of the

prevalence and is more influenced by a low test specificity than by low sensitivity. Lowering sensitivity has only a moderate biasing effect on the presumed prevalence, which is only evident for conditions that are relatively common.

It is noteworthy that although there is a 10-fold difference in true prevalences between the respective situations shown in table 2 the perceived magnitude of the difference is much lower, particularly if the test produces too many false positive results. For instance, at 10% false positives (i.e., 90% specificity) the estimated prevalences will differ by less than twofold, regardless of test sensitivity. This indicates, for example, that a molecular epidemiologic survey may fail to detect important differences in prevalence or incidence of two distinct endpoints that are substantially different in their occurrence in a given population.

When both the sensitivity and specificity of the test used to detect the endpoint are known it is possible to correct the resulting rate to eliminate the bias. The formula [19] is as follows:

$$P_c = (P_u + W - 1) / (S + W - 1)$$

where  $P_c$  and  $P_u$  are the corrected and uncorrected prevalence rates, respectively,  $S$  denotes test sensitivity and  $W$  denotes test specificity.

**Effect on the association between exposure and endpoint or outcome:** Figure 2 illustrates the biasing effect of measurement error on the estimates of association between an exposure biomarker and its cancer precursor outcome in a cohort study. The two graphs show separately the effects of misclassification of the biomarker (figure 2, top) and of the outcome (figure 2, bottom). Three hypothetical underlying causal relations between these two variables are depicted in each graph based on the strength of their statistical associations: RRs of 2, 10, and 50. Baseline biomarker prevalence is assumed as 20% and the cumulative risk of the outcome is constrained to be around 2.5%, levels that are consistent with those of natural history studies of HPV infection (taken as the biomarker measured at baseline) and high-grade CIN (taken as outcome, measured by short-term cytological follow-up) in many Western populations [20].

Under conditions of perfect outcome classification it is possible to observe that increasing misclassification of baseline biomarker status leads to biased estimates of RRs towards unity (figure 2, top). At 10% misclassification, that is, testing for the biomarker with 90% sensitivity and 90% specificity, the original RRs of 2, 10, and 100 are mistakenly estimated as being 1.7, 5.8, and 15.0, respectively. At 30% misclassification the bias is so severe that the measured RRs are 1.3, 2.3, and 3.3, respectively, values that would be indistinguishable statistically in most epidemiologic studies.

Although the biasing effect of biomarker exposure misclassification may seem appreciable, it is not as damaging to the validity of a study as misclassification of the cancer precursor outcome (figure 2, bottom). Assuming perfect ascertainment of the baseline biomarker, even minor levels of outcome misclassification diminish considerably the statistical strength of the underlying associations. With only 5% misclassification, that is, a level equivalent to an impressive cytological performance of 95% sensitivity and 95% specificity to detect high-grade CIN, the RRs of 2, 10 and 50 are mismeasured as 1.3, 2.4 and 2.7, estimates that cannot be distinguished statistically in practice.

It can be generalized that, for the conditions described above, outcome misclassification has a greater impact on study validity than exposure misclassification. Using the HPV-CIN example, had HPV infection been the outcome instead of the biomarker exposure, as is the case in cross-sectional surveys of determinants of HPV infection, the same conclusion would apply, i.e., the effect of misclassification of viral status would be more important than if HPV infection had been assumed as the exposure variable.

In practice, study validity is further aggravated by concomitant misclassification of exposure as well as outcome, which is a real concern in cohort studies since the latter frequently use precancer lesions as endpoints. On the other hand, case-control studies of invasive cancer are far less likely to be affected by outcome misclassification, but are prone to more severe exposure misclassification that can be differential or non-differential [21].

If the diagnostic performance of the biomarker assay is known (perhaps from a substudy done for validation purposes) separately for cases and non-cases in a case-control or cross-sectional study it is possible to correct for the biasing effect of differential and non-differential exposure misclassification using a simple formula [19]:

$$OR = \frac{(W_1 n_1 - b)(S_2 n_2 - c)}{(W_2 n_2 - d)(S_1 n_1 - a)}$$

where  $S$  = sensitivity,  $W$  = specificity,  $n$  is the number of subjects, and the subscripts 1 and 2 indicate cases or controls (or non-cases), respectively. The frequencies  $a$ ,  $b$ ,  $c$ , and  $d$  represent the cells of the 2x2 table in which the frequencies of exposure and outcome are cross-classified as follows:  $a$  = biomarker positive cases,  $b$  = biomarker negative cases,  $c$  = biomarker positive controls, and  $d$  = biomarker negative controls.

In the absence of data on the validity of an exposure assessment instrument or for a biomarker assay one cannot resort to the above formula to correct for the effects of misclassification. If misclassification is known to exist because of test variability -- for instance, because of inter-rater disagreement, inter-laboratory variation, or similar reasons -- one can use latent-class statistical techniques [22-24] that consider the collective information on the multiple results assessing the exposure or biomarker. This permits the computation of RR estimates that are less

biased than if considering the exposure information from a single assessment with error.

A simple illustration of how misclassification bias can be minimized by exploring the inter-laboratory assay variability for a genetic susceptibility marker was published recently [25]. An association between codon-72 p53 polymorphism and risk of HPV-induced cervical cancer (*Arg/Arg* genotype equated to greater risk) has been found [7,8] but this finding has not been replicated by others [26-31]. One possible reason may have been measurement error in the genotyping assay used by most laboratories that have attempted to reproduce the association. This was shown in a case-control study that assessed the impact of inter-laboratory variation in p53 genotyping on the magnitude of the proposed association [25]. This study analyzed codon-72 polymorphism blindly in three different laboratories. The disagreement among laboratories was substantial with Kappa coefficients in the 0.49-0.63 range. When disagreement between labs was allowed, the OR for the *Arg/Arg* genotype, compared to other forms, was 1.5 (95% CI: 0.5-3.9). In contrast, the OR increased to 8.0 (95%CI: 2.3-28.5) after exclusion of discordant genotypes, which presumably restricted the analysis to those subjects in whom measurement error was minimal or nonexistent. This strategy provided empirical evidence for the hypothesis that homozygous codon-72 p53-*Arg* may confer a higher susceptibility to HPV-associated cervical tumorigenesis.

Misclassification of biomarker as a covariate: As stated above in the discussion on empirical verification of the causal pathway model, the etiologic role of an intermediate endpoint can be verified by mediation analysis, that is, by using it as a covariate for adjusting the relation between a remote exposure and the cancer precursor outcome. If the intermediate endpoint is truly in the causal pathway between exposure and outcome the association between the latter two variables should disappear upon adjustment for the endpoint. Table 3 shows how this assessment may be hampered by measurement error in the intermediate endpoint. Three hypothetical scenarios are considered that specify different strengths for the relations between exposure and endpoint and between endpoint and outcome (V1-V2, and V2-0, respectively, using the causal pathway scheme in figure 1). Given the assumptions of strength of association for the latter 'connecting relations' and baseline risk the resulting crude RR between exposure and outcome varies from 1.7 to 17.1. As shown, misclassification of the endpoint as a covariate severely affects the ability to "adjust away" the indirect exposure-outcome association. At moderate levels of misclassification (10% false negatives and false positives) there remains a substantial portion of the original crude effect. In the second scenario shown in table 3 (at 90% sensitivity and specificity) the adjusted RR of 3.1 may be indistinguishable in practice from the crude RR of 4, depending on the size of the study. Upon realizing that the adjusted RR is not materially different from the crude one the observer may erroneously interpret this as an indication of an independent effect for the remote exposure via a separate mechanism that does not involve the intermediate endpoint. With greater misclassification the efficacy of the adjustment strategy is virtually non-existent. It can also be seen that test specificity influences the degree of misclassification bias more prominently than sensitivity.

As an illustration using the HPV-CIN model, lifetime number of sexual partners is a predictor of risk for acquiring HPV infection, an intermediate endpoint that influences risk of CIN. In this model the statistical association between sexual activity and CIN should only be present in analyses that do not control for properly ascertained HPV status [9]. This has been empirically demonstrated in studies using polymerase chain reaction techniques to detect HPV infection [15,32,33]. However, earlier molecular epidemiology studies using first generation hybridization assays to detect HPV infection were prone to misclassification and erroneously revealed that number of sexual partners was independently associated with risk of cervical cancer, even after adjustment for HPV infection [34,35].

Misclassification of an intermediate endpoint producing spurious interaction: Another deleterious effect of measurement error in the causal pathway model described above results in the erroneous perception of interaction between the remote exposure and the intermediate endpoint on the risk of the cancer precursor. This situation will occur if an observer is attempting to verify interaction assuming the third model scenario in figure 1 when in reality the correct underlying model is that of a causal pathway. As part of the empirical verification process the investigator will probably have calculated adjusted estimates for V1-0 given V2 and for V2-0 given V1 and then proceeded to compute stratum specific RRs for each of these two relations to probe for effect modification which would have given empirical proof of interaction between V1 and V2 on the risk of O.

The potential findings from such an analysis can be best illustrated by considering the first situation in table 3 at 90% sensitivity and specificity for the testing of V2, the unsuspected intermediate endpoint. In addition to the estimates for V1-0 which are shown in the table (RR for V1-0: crude = 1.74, V2-adjusted = 1.58) the results for V2-0 are as follows: crude RR = 3.19, V1-adjusted RR = 3.08. Taken together and before any supplemental analysis are done these findings are consistent with independence of effects (with V2 being the stronger determinant of the two) accompanied by mild mutual confounding; but the results are certainly not consistent with the unsuspected, true causal pathway that 'drives' those relations. The analysis of a randomly chosen sample of 500 cases and 500 controls (to simulate a moderately large case-control study) from the source cohort that was used to illustrate these associations reproduces these findings using ORs to estimate the RRs. The next step is the stratification of the two associations to verify effect modification. The stratum specific ORs for V1-0 are 3.24 (95%CI: 1.7-6.2) for V2-positive and 1.10 (95%CI: 0.8-1.5) for V2-negative individuals, with strong evidence of heterogeneity between these two estimates ( $P = 0.002$ ). The ORs for V2-0 after stratification by levels of V1 are 5.15 (95%CI: 3.0-8.8) for V1-positive and 1.74 (95%CI: 1.1-2.8) for V1-negative, and equivalent evidence of heterogeneity between estimates. The

conclusion that would be reached by the analyses of this case-control study would point to interaction between V1 and V2 on the risk of O, clearly an erroneous interpretation because the underlying model specified a causal pathway relation scheme. In fact, without the 10% misclassification of the intermediate endpoint V2 the stratum specific RRs for V1-O are both equal to 1.0, whereas those for V2-O are both equal to 10.0, the actual magnitude of the relation specified in the first scenario of table 3.

Measurement error affecting the assessment of a necessary cause: On the basis of epidemiologic theory, it can be expected that the RR between risk factor and disease will tend to infinity (because of a denominator of zero due to the non-existence of non-exposed cases) if the disease cannot arise from routes other than that initiated by the risk factor, that is, if the latter is a necessary cause for the disease. On the other hand, if a cause is not necessary the risk of disease among those not exposed to the risk factor will not be negligible, and thus, the magnitude of the RR will be lower, all other conditions being held constant. The difference in the magnitude of RRs between these two situations is a key element in judging whether the results from traditional epidemiologic studies can be used to distinguish a necessary from a non-necessary cause. However, measurement error in ascertaining an exposure or intermediate endpoint will also hamper our ability to assess empirically whether or not such an endpoint is a necessary event in carcinogenesis.

Of the known causes and determinants of cancer, none is considered necessary or sufficient but recently it has been suggested that HPV infection may be the first determinant of a human cancer that may have characteristics of a necessary cause [10,36,37]. If confirmed, this would have obvious implications for the primary and secondary prevention of cervical cancer. However, misclassification of cumulative exposure to HPV alone makes it impossible to use the magnitude of the RR estimates for the association between HPV and CIN to differentiate between the necessary and non-necessary cause assumptions [11]. Figure 3 illustrates the problem by showing how RRs vary in response to changes in sensitivity (abscissa) and specificity (the 3 curves in each set), separately, and for two scenarios of HPV exposure prevalence: 10% and 50% [11]. Cumulative risk of CIN among those exposed to HPV is fixed at 50%, resulting in a RR of 100, if HPV exposure ascertainment is free of error and we assume a non-necessary causal relation (that is, the risk of CIN among the unexposed is non-negligible at 0.5%). At an HPV prevalence of 10%, there is substantial overlapping of curves (Fig. 3, top), with relations becoming of comparable magnitude at sensitivity levels of 90% and lower (i.e., false negative rates of 10% and higher). A nearly complete loss of the ability to distinguish between causality assumptions occurs at the higher HPV prevalence of 50% (Fig. 3, bottom). Figure 3 also shows that the lower the exposure prevalence, the less important the effect of losses in sensitivity in reducing RR estimates. Conversely, specificity takes a more important role at the lower prevalence levels and has an almost negligible effect on the magnitude of the HPV-CIN relation at the relatively high 50% HPV exposure.

Minimizing the influence of measurement error: It is obvious from the preceding demonstrations that epidemiologists and their laboratory collaborators would do well to devote energy and resources towards the improvement of questionnaire instruments to obtain information on exposure to risk factors and developing more specific and sensitive assays to detect intermediate endpoints and other biomarkers of cancer risk. In addition, characterizing or quantifying the measurement characteristics of biomarkers and endpoints enables subsequent correction for measurement error. However, even with nearly perfect study instruments there is another aspect of measurement error that remains critical: that of an incomplete ascertainment of the true exposure and endpoint status due to fluctuations in these variables over time. This is an issue that is particularly important in studies of cancer precursors because of the dynamic changes in early events of the natural history of neoplasia that precede the onset of irreversible lesions [11,38]. Unfortunately, most epidemiologic studies of cancer and precancer are based on only one measurement of exposure to remote determinants, intermediate endpoints, and lesion outcomes. Case-control and cohort investigations thus fail to capture the complete picture of the history of transition events in carcinogenesis.

Statistical modelling by logistic and proportional hazards regression methods enhances the ability to probe associations in epidemiologic datasets, by allowing control of confounding, assessment of interaction among variables, and stratification by design and matching variables and by time between onset of exposure and outcome. However, behind the added level of insight that multivariate modelling brings to epidemiologic data analysis, the basic 2x2 table correlating exposure and outcome remains the fundamental unit of information used to generate epidemiologic evidence for or against the role of a risk factor. Unfortunately, this central 2x2 table is usually based on a single-specimen assessment of exposure, which combines the sampling and testing errors typical of one testing opportunity with those resulting from temporal fluctuations in detectability of biomarkers during the natural history of cancer.

Therefore, to understand the role of and mechanism for such dynamic changes in the natural history of cancer one must conduct studies that collect data repeatedly on risk factors, intermediate endpoints, and lesion outcomes on multiple occasions during follow-up. A longitudinal, repeated-measurement cohort study is required to increase the accuracy and reduce bias in the assessment of cumulative exposure and outcome history. The longitudinal nature of such studies poses new challenges in study conduct, data management [39-41] and analysis [42-44], but greatly enhances our ability to learn about key events leading to cancer development.



## Conclusion

As discussed in this chapter, epidemiologic studies coupled with careful, hypothesis-driven data analysis [45,46] can be useful in providing empirical evidence for the role of intermediate endpoints and cancer precursors in the genesis of cancer and for susceptibility traits and exposures in the development of cancer precursors. The specification of *a priori* models that depict the relations among variables and their directionality requires an eclectic consideration of all aspects of disease biology, histopathology, and genetics within the time-tested epidemiologic framework for assessing causal mechanisms of disease [47]. However, even the most multidisciplinary molecular epidemiology studies must contend with the impact of measurement error in remote exposures, genetic susceptibility markers, intermediate endpoints, and precursor lesions. Although careful consideration of the possible effects of misclassification of key variables helps our understanding of the nature and degree of the ensuing biases on the measures of association a preventive approach to minimizing misclassification is a preferred solution. This involves not only the development and validation of better survey instruments and laboratory assays but also designing epidemiologic studies that can properly measure the dynamic changes in the early events in the natural history of cancer.

## References

1. Rothman KJ. Causes. Am J Epidemiol. 1976; 104: 587-92.
2. Rothman KJ, Greenland S. Modern Epidemiology. 2<sup>nd</sup> Edition. Lipincott-Raven, Philadelphia, 1998, pp. 7-28, 115-134.
3. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. Epidemiology. 1999; 10:37-48.
4. Cann CI, Fried MP, Rothman KJ. Epidemiology of squamous cell cancer of the head and neck. Otolaryngol Clin North Am. 1985; 18: 367-88.
5. Schlecht NF, Franco EL, Pintos J, Negassa A, Kowalski LP, Oliveira BV, Curado MP. Interaction between tobacco and alcohol consumption and the risk of cancers of the upper aero-digestive tract in Brazil. Am J Epidemiol. 1999; 150: 1129-37.
6. Apple RJ, Erlich HA, Klitz W, Manos MM, Becker TM, Wheeler CM. HLA DR-DQ associations with cervical carcinoma show papillomavirus-type specificity. Nat Genet 1994; 6:157-62.
7. Storey A, Thomas M, Kalita A, Harwood C, Gardiol D, Mantovani F, Breuer J, Leigh IM, Matlashewski G, Banks L. Role of a p53 polymorphism in the development of human Papillomavirus-associated cancer. Nature 1998; 393: 229-234.
8. Zehbe I, Voglino G, Wilander E, Genta F, Tommasino M. Codon 72 polymorphism of p53 and its association with cervical cancer. Lancet 1999; 354, 218-219.
9. Franco EL. The sexually transmitted disease model for cervical cancer: incoherent epidemiologic findings and the role of misclassification of human papillomavirus infection. Epidemiology. 1991; 2: 98-106.
10. Franco EL: Cancer causes revisited: human Papillomavirus and cervical neoplasia. J Natl Cancer Inst 1995; 87: 779-80.
11. Franco EL, Rohan TE, Villa LL: Epidemiologic evidence and human papillomavirus infection as a necessary cause of cervical cancer. J Natl Cancer Inst 1999; 91: 506-11.
12. Begg CB, Zhang ZF. Statistical analysis of molecular epidemiology studies employing case-series. Cancer Epidemiol Biomarkers Prev 1994; 3: 173-5.
13. Khoury MJ, Flanders WD. Nontraditional epidemiologic approaches in the analysis of gene-environment interaction: case-control studies with no controls. Am J Epidemiol 1996; 144: 207-13.
14. Goldstein AM, Andrieu N. Detection of interaction involving identified genes: available study designs. J Natl Cancer Inst Monogr. 1999; 26:49-54.
15. Schiffman MH, Bauer HM, Hoover RN, Glass AG, Cadell DM, Rush BB, et al. Epidemiologic evidence showing that human papillomavirus infection causes most cervical intraepithelial neoplasia. J Natl Cancer Inst 1993; 85: 958-64.
16. Greenland S. The effect of misclassification in the presence of covariates. Am J Epidemiol. 1980; 112: 564-9.
17. Greenland S, Robins JM. Confounding and misclassification. Am J Epidemiol. 1985; 122: 495-506.

18. Armstrong BG. Effect of measurement error on epidemiological studies of environmental and occupational exposures. *Occup Environ Med* 1998; 55: 651-6.
19. Franco EL: Measurement errors in epidemiological studies of human papillomavirus and cervical cancer. In: Muñoz N, Bosch FX, Shah KV, Meheus A (eds): *The epidemiology of human papillomavirus and cervical cancer*. Oxford: Oxford University Press, 1992, pp. 181-197.
20. Franco EL. Statistical issues in studies of human papillomavirus infection and cervical cancer. In: Franco EL, Monsonego J (eds.) *New developments in cervical cancer screening and prevention*. Blackwell, London, 1997 (pp. 39-50).
21. Wynder EL. Investigator bias and interviewer bias: the problem of reporting systematic error in epidemiology. *J Clin Epidemiol* 1994; 47:825-7.
22. Liu XH, Liang KY. Adjustment for non-differential misclassification error in the generalized linear model. *Stat Med* 1991; 10:1197-211.
23. Bashir SA, Duffy SW. The correction of risk estimates for measurement error. *Ann Epidemiol* 1997; 7:154-64.
24. Emsley CL, Gao S, Hall KS, Hendrie HC. Estimating odds ratios adjusting for misclassification in Alzheimer's disease risk factor assessment. *Stat Med* 2000; 19: 1523-30.
25. Makni H, Franco EL, Kaiano J, Villa LL, Labrecque S, Dudley R, Storey A, Matlashewski G. P53 polymorphism in codon 72 and risk of human papillomavirus-induced cervical cancer: effect of inter-laboratory variation. *Int J Cancer* 2000; 87: 528-33.
26. Helland A, Langerod A, Johnsen H, Olsen AO, Skovlund E, Borresen-Dale AL. p53 polymorphism and risk of cervical cancer. *Nature* 1998;396:530-531.
27. Josefsson AM, Magnusson PK, Ylitalo N, Quarforth-Tubbin P, Ponten J, Adami HO, Gyllensten U. p53 polymorphism and risk of cervical cancer. *Nature* 1998; 396: 531.
28. Hildesheim A, Schiffman M, Brinton LA, Fraumeni Jr JF, Herrero R, Bratti MC, Schwartz P, Mortel R, Barnes W, Greenberg M, McGowan L, Scott DR, Martin M, Herrera JE, Carrington M. p53 polymorphism and risk of cervical cancer. *Nature* 1998; 396: 531-2.
29. Lanham S, Campbell I, Watt P, Gornall R. p53 polymorphism and risk of cervical cancer. *Lancet* 1998; 352: 1631.
30. Klaes R, Ridder R, Schaefer U, Benner A, Von Knebel Doeberitz M. No evidence of p53 allele-specific predisposition in human papillomavirus-associated cervical cancer. *J Mol Med* 1999; 77: 299-302.
31. Minaguchi T, Kanamori Y, Matsushima M, Yoshikawa H, Taketani Y, Nakamura Y. No evidence of correlation between polymorphism at codon 72 of p53 and risk of cervical cancer in Japanese patients with papillomavirus 16/18 infection. *Cancer Res* 1998; 58: 4585-4586.
32. Munoz N, Bosch FX, de Sanjose S, Vergara A, del Moral A, Munoz MT, Tafur L, Gili M, Izarzugaza I, Viladiu P, et al. Risk factors for cervical intraepithelial neoplasia grade III/carcinoma in situ in Spain and Colombia. *Cancer Epidemiol Biomarkers Prev.* 1993; 2: 423-31.
33. Kjaer SK, van den Brule AJ, Bock JE, Poll PA, Engholm G, Sherman ME, Walboomers JM, Meijer CJ. Human papillomavirus--the most significant risk determinant of cervical intraepithelial neoplasia. *Int J Cancer* 1996; 65: 601-6.
34. Reeves WC, Brinton LA, Garcia M, Brenes MM, Herrero R, Gaitan E, Tenorio F, Britton RC, Rawls WE. Human papillomavirus infection and cervical cancer in Latin America. *N Engl J Med* 1989; 320: 1437-41.
35. Donnan SP, Wong FW, Ho SC, Lau EM, Takashi K, Esteve J. Reproductive and sexual risk factors and human papilloma virus infection in cervical cancer among Hong Kong chinese. *Int J Epidemiol* 1989; 18: 32-36.
36. Walboomers JMM, Meijer CJLM: Do HPV-negative cervical carcinomas exist? *J Pathol* 1997; 181: 253.
37. Walboomers JM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah KV, Snijders PJ, Peto J, Meijer CJ, Munoz N. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol.* 1999; 189:12-9.
38. Schiffman MH, Schatzkin A. Test reliability is critically important to molecular epidemiology: an example from studies of human papillomavirus infection and cervical neoplasia. *Cancer Res* 1994; 54: S1944-7.

39. Duffy SW, Rohan TE, McLaughlin JR. Design and analysis considerations in a cohort study involving repeated measurement of both exposure and outcome: the association between genital papillomavirus infection and risk of cervical intraepithelial neoplasia. *Stat Med* 1994; 13: 379-90.
40. Rothman N, Stewart WF, Schulte PA. Incorporating biomarkers into cancer epidemiology: a matrix of biomarker and study design categories. *Cancer Epidemiol Biomarkers Prev* 1995; 4: 301-11.
41. Franco E, Villa L, Rohan T, Ferenczy A, Petzl-Erler M, Matlashewski G. Design and methods of the Ludwig-McGill longitudinal study of the natural history of human papillomavirus infection and cervical neoplasia in Brazil. *Panam J Pub Hlth* 1999; 6: 223-33.
42. Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* 1986; 42: 121-130.
43. Zeger SL, Liang KY. An overview of methods for the analysis of longitudinal data. *Stat Med.* 1992; 11: 1825-39.
44. Ho GY, Burk RD, Klein S, Kadish AS, Chang CJ, Palan P, Basu J, Tachezy R, Lewis R, Romney S. Persistent genital human papillomavirus infection as a risk factor for persistent cervical dysplasia. *J Natl Cancer Inst* 1995; 87:1365-71.
45. Greenland S. Modeling and variable selection in epidemiologic analysis. *Am J Public Health* 1989; 79: 340-9.
46. Maldonado G, Greenland S. Simulation study of confounder-selection strategies. *Am J Epidemiol* 1993; 138: 923-36.
47. Schulte PA, Rothman N, Schottenfeld D. Design considerations in molecular epidemiology. In: Schulte PA, Perera FP (eds.) *Molecular epidemiology: Principles and practices*. Academic Press, San Diego, 1993, pp. 159-198.

Table 1. Possible findings for the statistical associations between V1 and O and between V2 and O under the different models of the natural history of cancer precursors described in figure 1.

Etiologic model	Assumed relations	Association V1 – O		Association V2 – O	
		Crude	Adjusted for V2	Crude	Adjusted for V1
Independence	$V1 \rightarrow O \leftarrow V2$	++	++	++	++
Confounding	$V2 \leftarrow V1 \rightarrow O$	++	++	+	- <sup>1</sup>
Interaction <sup>2</sup>	$V1 + V2 \rightarrow O$	+	++ / + / -	+	++ / + / -
Causal pathway	$V1 \rightarrow V2 \rightarrow O$	+	-	++	++
Correlates of outcome <sup>3</sup>	$V1 \rightarrow O \rightarrow V2$	++	+	++	+
	$V1 \leftarrow O \rightarrow V2$	++	++ / +	++	++ / +

<sup>1</sup> A negative sign indicates that the association disappears, i.e., a RR = 1.

<sup>2</sup> Stratification is more informative than adjustment and should thus be the preferred approach.

<sup>3</sup> A change from ++ to + indicates that in some combinations the adjusted estimate is substantially lower than the crude one.

Table 2. Presumed prevalence (%) of a cancer precursor endpoint resulting from testing for a surrogate marker under different levels of test sensitivity and specificity in three hypothetical populations with different prevalence rates of the endpoint.

True prevalence <sup>1</sup>	Specificity	Presumed prevalence (%) at different sensitivity levels		
		99%	90%	70%
0.1%	99%	1.1	1.1	1.1
	90%	10.1	10.1	10.1
	70%	30.1	30.1	30.0
1%	99%	2.0	1.9	1.7
	90%	10.9	10.8	10.6
	70%	30.7	30.6	30.4
10%	99%	10.8	9.9	7.9
	90%	18.9	18.0	16.0
	70%	36.9	36.0	34.0

<sup>1</sup> Assumed free of error.

Table 3. Effect of misclassification of an intermediate endpoint on the adjusted relative risk estimates for the relation<sup>1</sup> between a remote exposure and a cancer precursor outcome that is mediated by the intermediate endpoint.

Underlying RR (exposure-endpoint)	Underlying RR (endpoint-outcome)	Resulting Crude RR (exposure-outcome)	Specificity in detecting endpoint	Adjusted RR (exposure-outcome) <sup>2</sup> by sensitivity level		
				99%	90%	70%
10	10	1.74	99%	1.16	1.22	1.36
			90%	1.55	1.58	1.65
			70%	1.68	1.70	1.72
10	50	3.96	99%	1.47	1.68	2.18
			90%	2.98	3.13	3.45
			70%	3.62	3.70	3.84
50	50	17.11	99%	1.82	3.06	5.99
			90%	6.40	7.78	10.83
			70%	11.63	12.81	15.04

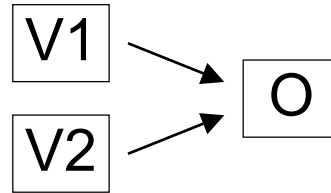
<sup>1</sup> Model assumptions: cumulative risk of endpoint among non-exposed = 1%, cumulative risk of outcome among those with endpoint = 1%, prevalence of exposure = 30%.

<sup>2</sup> RR adjusted for the intermediate endpoint measured by a hypothetical assay with stated levels of sensitivity and specificity.

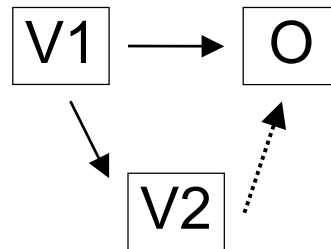
**Figure 1:**

Simple theoretical models specifying some epidemiologic relations of interest in studies involving cancer precursors. V1 and V2 designate any two variables that mediate risk of an outcome (O) in the natural history of disease. V1 and V2 could denote exposure variables or intermediate endpoints. 'O' could denote a cancer outcome or a cancer precursor, depending on the situation. Solid arrows denote causal relations. Dotted arrows denote secondary, non-causal relations.

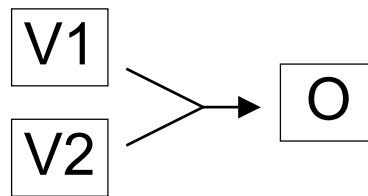
**Independence of effects:**



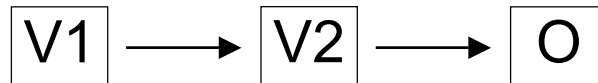
**Confounding:**



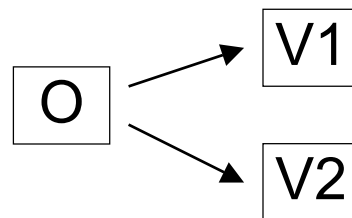
**Interaction:**



**Causal pathway:**

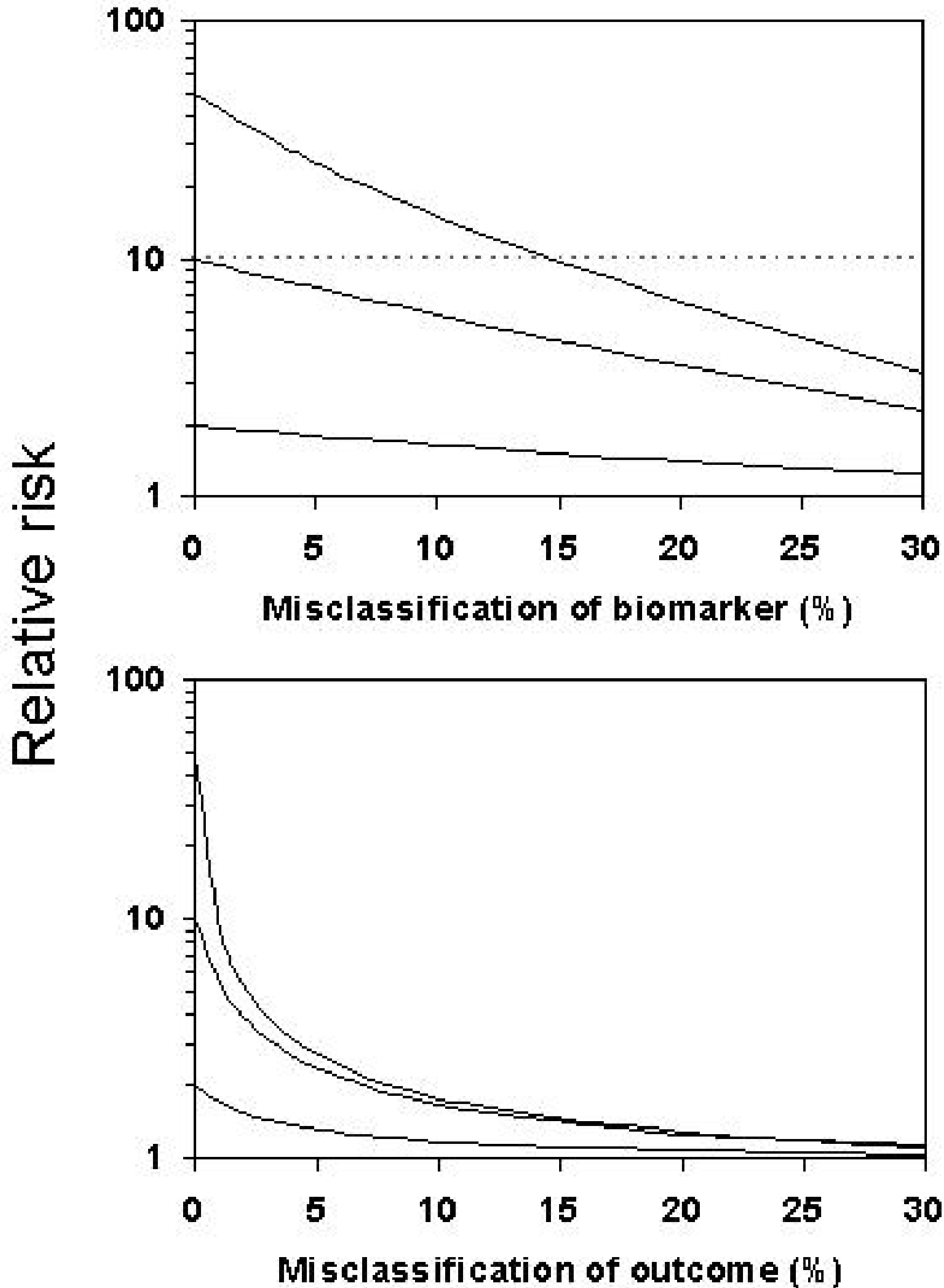


**Correlates of outcome:**



**Figure 2:**

Biassing effect of misclassification of an exposure biomarker (top graph) or of its resulting cancer precursor outcome (bottom graph) on the RR estimate for the association between the biomarker and the outcome in a hypothetical cohort study. Baseline prevalence of the biomarker is assumed as 20% and the cumulative risk of the outcome as between 2% and 3%. Three hypothetical relations are assumed based on the strength of the statistical associations (RRs of 2, 10, and 50) for the underlying biomarker-outcome relation.



**Figure 3:**

RR estimates for the human papillomavirus-cervical intraepithelial neoplasia (HPV-CIN) association as a function of the misclassification in HPV exposure, assuming a 50% cumulative risk of CIN among those exposed to HPV and 0.5% or 0% among those unexposed, for non-necessary and necessary cause assumptions, respectively. Continuous lines: cohorts based on the necessary-cause assumption for the HPV-CIN association; broken lines: cohorts based on the assumption that HPV is not a necessary cause of CIN. The three curves in each set represent different specificity levels: 99% (top), 90% (middle), and 80% (bottom). The two graphs differ in the assumed prevalence of HPV exposure: 10% (top), 50% (bottom). Adapted from reference 11, with permission.

