CONFOUNDING: ESSENCE AND DETECTION[1]

OLLI S. MIETTINEN AND E. FRANCIS COOK

Miettinen, O. S. (Harvard School of Public Health, Boston, MA 02115) and E. F. Cook. Confounding: essence and detection. *Am J Epidemiol* **1981; 114:593-603.**

Confounding is examined from first principles. In follow-up studies a confounder is a predictor of diagnosing the illness—by being either a risk indicator or a determinant of diagnostic errors; in addition, it shows different distributions between the exposed and nonexposed series. In case-referent studies confounding can arise in two ways. A priori confounders are correlates of exposure in the joint source population of cases and reference subjects; also, they are determinants of diagnosing the illness or have different selection implications between cases and referents. In addition, factors bearing on the accuracy of exposure information are confounders if distributed differently between cases and referents. Criteria based singularly on relationships in the data can be misleading. Similarly, a change in the estimate and even a change in the parameter as a result of control is not a criterion rooted in first principles of confounding and can lead to a false conclusion.

biometry; epidemiologic methods; follow-up studies

In epidemiologic research directed to the effect of a particular exposure on the risk of developing a particular illness, a central problem is the need to consider extraneous factors that might be explanatory, partially or totally, of the magnitude of the estimate of the effect. An understanding of the nature of such factors, referred to as "confounders" or "confounding factors," is thus essential to study design and data analysis—and finally to the interpretation of the resulting estimates.

Epidemiologic literature on confounding, still very scanty, appears to echo a few basic principles. In the detection of confounding it is commonplace to use the data-based criterion that control of the extraneous factor changes the estimate of effect (1−4). An alternative to this is that the extraneous factor be associated, in the data, with both the exposure and the illness (1−9). Insofar as there is any a priori aspect to the prevailing criteria for confounders, it is that confounders must be risk indicators for the illness (i.e., determinants, causal or noncausal, of the health state at issue). With this constraint, confounders are taken to be a subset of known risk indicators for the illness in question—that subset which meets the data-based criterion used (4, 8).

In none of the literature that we are familiar with is any distinction made between follow-up and prevalence studies on the one hand and case-referent (case-control) studies on the other.

In this paper we outline our current insights into the nature and manifestations

of confounding, different in some important ways from those previously presented. To a large extent these new views have arisen inductively from the examination of particular problems, and this inductive approach will be followed in this presentation as well in that the proposed principles are introduced by the use of particular examples.

CONFOUNDING IN FOLLOW-UP STUDIES

Example 1. One of the findings' of the University Group Diabetes Program—a randomized clinical trial—was an excess of coronary heart disease mortality among users of tolbutamide (a drug used for the reduction of blood glucose level) when compared with placebo users (10). In the context of considering possible causal interpretations of this observation it was pointed out that, by chance or otherwise, the tolbutamide series might be of higher risk in terms of its distribution by familiar coronary heart disease risk indicators than the group given the placebo, so that "one would expect to see more cardiovascular deaths" in this series (11). This concern led to a comparison of the distributions of risk indicators between the two series and, where differences were found, an exploration of the magnitude of their confounding impact in terms of estimates derived by the use of stratification and multivariate analysis (12). Had there been errors of diagnosing coronary heart disease deaths dependent on a covariate (such as physician identity) and had this factor had different distributions between the compared series, this factor, too, would have had to be controlled (8).

Example 2. Suppose that in Example 1 the tolbutamide series was older (on account of chance) than the placebo series. Suppose further that, contrary to expectation, age failed to show any association with coronary heart disease mortality in the data (within either treatment group) and that, as a result, the usual statistical control of age did not change the estimate of the tolbutamide effect. Nevertheless, part of the excess coronary heart disease mortality among tolbutamide users would have been attributable to the age difference: despite the lack of a relationship between coronary heart disease mortality and age within the study experience, one would expect the placebo series to have shown a higher mortality had its follow-up started at a later age, comparable in distribution to that in the tolbutamide series. (For elaboration of the subtleties here, see Appendix 1.)

These examples illustrate the nature of confounders in follow-up studies in general. As is well known, they are *predictors (determinants) of diagnosing the illness*—by being either risk indicators or determinants of diagnostic errors—in the type of setting represented by the study. It is worthy of emphasis that it is this predictive tendency (an a priori property) that is relevant rather than its manifestation in the data at hand (cf. Example 2). It is equally well known that a determinant of the (empirical) outcome is a confounder only if *the exposed and nonexposed subjects in the study show different distributions by this factor, for whatever reason (13).*

What is not well known is that this outlook on confounding can come in conflict with the one based on the consequence of control.

Example 3. Consider a follow-up study with identical distributions by gender among the exposed and nonexposed, so that there can be no confounding by this characteristic according to the above principles. Suppose the data, with gender ignored, are as presented in panel A of table 1. Presuming no confounding by gender, the estimates for illness odds for the exposed and nonexposed may be computed as 104/96 and 96/104, respectively, leading to the odds ratio estimate of $(104/96)/(96/104) = 1.2$. Consistent with this, the data, when stratified by gender (despite the identity of distributions by

this characteristic), could be as shown in panel B of table 1; the estimates for both strata equal 5.2. Thus, even though gender should not be a confounder, its control changes (dramatically) the value of the estimate (and the parameter) of effect. It is of added note that the risk difference estimate remains unchanged upon stratification in this example, while the stratum-specific risk ratio estimates straddle the crude one. Thus, an added dilemma arising from the change-in-estimate (or parameter) criterion is whether confounding is to be defined differently for different parameters.

In the face of this conflict the conventional view is, we believe, that confounding is present whenever the parameter-value that corresponds to the control of the covariate differs from the ("crude") value without such control (14)—a view that makes confounding dependent on the type of parameter considered (cf. Example 3). We believe that the change-in-estimate (or parameter) criterion for confounding is not a derivative of tenable first principles, and that it can lead to false conclusions, as will be discussed in a later section. As for the above example, it is our view that it does not illustrate confounding but modification, specifically the peculiarity and subtlety of the odds ratio (OR) parameter relative to risk difference (RD) and risk ratio (RR) in the context of modification. Consider the situation in which the exposed and nonexposed have identical distributions by the covariate, so that confounding is absent by the criteria involving its relationship to exposure and illness. If the covariate is not a modifier of RD in the usual sense, i.e., if the value of RD is the same at all categories of the covariate, then the crude RD equals the value specific to each category of the covariate (cf. Example 3). An analogous statement applies to RR, but *not* to OR (cf. Example 3). An interpretation of this in terms of modification requires expansion of the

TABLE 1

Hypothetical follow-up data for Example 3, together with estimates of illness-odds ratio (OR), illustrating a peculiarity of this parameter— that its unconfounded overall value can fall outside the range of its stratum-specific values

A) Unstratified data (unconfounded by gender*)

		Exposure	
		+	−
Illness	+	104	96
	−	96	104
		200	200
Proportion male*		50%	50%
$\widehat{OR}$		1.2	

B) Data stratified by gender

		Male		Female	
		Exp+	Exp−	Exp+	Exp−
Illness	+	99	95	5	1
	−	1	5	95	99
		100	100	100	100
$\widehat{OR}$		5.2		5.2	

* Equal distribution implies no confounding by gender.

concept of modification: a parameter may be modified according to category of the covariate (prevailing view (8)) and/or according to the *distribution* of the covariate. In these terms it may be said that if, in the absence of confounding by the covariate, RD or RR are not modified by category of the covariate, they are not modified by its distribution either; by contrast, OR unmodified by category can be modified by distribution (cf. Example 3). When, in the absence of confounding, RD and/or RR are modified by category of the covariate, they, too, are modified by distribution. Thus, with the overall gender-distribution in Example 3, the unconfounded RR is (104/200)/(96/200) = 1.08, whereas with distribution shifting toward the female gender it would approach the female-specific value of (5/100)/(1/100) = 5.0. In seeking to internalize this it is important to bear in mind that RR and OR must be thought of as functions of the *average* risks for the exposed (R_1) and

nonexposed (R_0), respectively, and not as averages of individual RR or OR values. For, conditionally on a most profound determinant of risk, both risks take on the values of unity or zero only, so that RR is either 1, 0, infinity or undefined, while OR is infinity, zero, or undefined.

The above criteria for confounding in follow-up studies require some refinement. First, a confounder is a determinant of the (empirical) outcome specifically on the null hypothesis or, equivalently, among the nonexposed (since risk indicator status among the exposed could result from modification by the potential confounder) (8). Furthermore, the relatedness of the confounder to both exposure and to the (empirical) health outcome at issue must hold conditionally on all other factors being controlled (8, 15). Finally a set of factors which singly satisfy these criteria, including the conditionality aspect, may in the aggregate result in no confounding: the individual confounding effects may cancel each other, i.e., the risk function involving all the covariates (1) may be identically distributed between the exposed and nonexposed (16).

CONFOUNDING IN CASE-REFERENT STUDIES

Example 4. Consider again Example 1 but now with a hypothetical case-referent study within the University Group Diabetes Program cohort, using all the cases and a simple random sample of the noncases. If it were known only that predictors of (diagnosing) coronary heart disease had no tendency to become associated with tolbutamide use (because of the randomization), it would be proper to infer absence of confounding. On the other hand, if it were known that, despite the randomization, certain determinants of (diagnosing) coronary heart disease death showed different distributions between the tolbutamide and placebo cohorts, these factors would be confounders, just as in Example 1. Throughout,

the issue is association with exposure in the joint source population from which both cases and reference subjects arose, rather than among the cases and/or referents themselves.

Example 5. Consider a case-referent study on the effect of non-O blood group on the risk of developing coronary heart disease. Suppose the cases are "representative" of all cases, and thus show a predominance of males (reflecting the gender effect on the risk of coronary heart disease) while the reference series, being "representative" of the source population of cases (the candidate population), contains equal numbers of males and females. This difference in gender distribution is of consequence only insofar as gender has some connection with blood group distribution. In point of fact, it is known in a priori terms that there is no such relationship, and thus a reference series which is 50 per cent male is interchangeable with one that shows a predominance of males, so that gender is not to be regarded as a confounder in this example (17).

Example 6. Consider a case-referent study on the effect of smoking on the risk of lung cancer among uranium miners. The study protocol would of course stipulate that the case series derive from the occupational population at issue. Suppose that the reference series represents another occupational group. If there is no difference in smoking habits between the occupations involved, i.e., in the exposure patterns of the respective source populations of cases and reference subjects, then the reference series is interchangeable with one drawn from the source population of the cases, so that occupation is not a confounder; otherwise, it is.

These examples illustrate a first—singularly a priori—type of confounding in case-referent studies. The factors at issue are correlates of the (empirical/knowable) exposure in the *joint source* of cases and referents. (The source of cases is

the same as that of the referents in Examples 4 and 5, but in Example 6 the joint source is the union of the source populations for cases and referents, respectively.) In addition, these factors satisfy the illness-relationship of confounders in one (or both) of two ways: they are predictors of diagnosing the illness and/or they have different selection-implications between cases and referents (18).

Example 7. Expanding on Example 5, parental blood group is a determinant of each person's blood group in the source population and thus satisfies the exposure-relationship of an a priori confounder in this example. It does not, however, satisfy either one of the above criteria for the illness-relationship of such a confounder: parental blood group is not a determinant of coronary heart disease or its diagnosis (on the null hypothesis), and it had no implications for subject selection. But suppose that the compared series nevertheless show different proportions of parents with non-O blood group. This observed difference between the compared series might lead one to be concerned with the possibility of parental blood group being a confounder after all. However, as long as one is prepared to see a difference in blood group distribution between cases of coronary heart disease and reference subjects, one must also be prepared to see different distributions for all correlates of blood group, such as parental blood group. Controlling for such correlates would result in undue reduction in the variation of the exposure and thereby obscure the manifestation of its effect. For example, in the domain where all parents show blood group O, both of the compared series will show a 100 per cent prevalence of blood group O, supplying no information on the effect of interest.

Example 8. Consider again the case-referent study described in Example 6. Presumably the case series would show a higher prevalence of various correlates of cigarette smoking such as history of smoking among family members and habitual match-carrying. Such differences would not tend to occur on the null hypothesis, as they are simply a manifestation of the role of cigarette smoking as a cause of lung cancer. Thus, the observation of a difference between the compared series in terms of the distribution of such correlates is no basis for inferring confounding—as long as these characteristics are not indicators of the risk of the illness (or determinants of its diagnosis) independent of the exposure and have no differential implications on the selection of cases and reference subjects.

These latter two examples illustrate the principle, generally unappreciated, that a correlate of exposure (in the joint source population) is not an actual confounder by virtue of a difference in its distribution between cases and reference subjects, i.e., that the illness-relationship of an a priori confounder cannot be viewed in terms of comparing the distributions of cases and referents.

Example 9. Consider again Example 6, but suppose now that the reference series had been matched to the cases with regard to match-carrying. Matching on this factor would have the tendency of making the smoking rate among the reference subjects similar to that of the cases—as a consequence of leading to identity of the distribution of the index and reference series according to this factor. This would be distortive of the value of the effect measure (were match-carrying not properly allowed for in the analysis); in other words, matching on the history of match-carrying would make this correlate of exposure an actual confounder.

This example does not bring up any further principle. It only serves to illustrate two (seemingly novel) principles already considered. First, it shows that matching can satisfy the illness relationship so as to make a source-correlate of exposure a confounder in case-referent

studies—because a matching factor has different selection implications for cases and reference subjects (it has implications on the selection of reference subjects only). Second, it underscores the principle that relative distributions of cases and referents do not constitute a criterion for confounding by a source-correlate of exposure in case-referent studies.

Example 10. Consider a case-referent study of coronary heart disease in relation to exercise and smoking, with the history of exposure elucidated by interview. Suppose some of the cases are deceased at the time of potential interview, so that the information is obtained from the next of kin, while all reference subjects are alive and interviewed directly (19, 20); and suppose, too, that the cases are interviewed by the investigator himself while the reference subjects are interviewed by his assistants (19, 20). In the analysis of the data—or at least in the inference stage—there is a need to make allowance for the differences between the index and reference series in terms of the types of informant and interviewer, i.e., to treat these information factors within the study as confounders (19, 20).

This example indicates that in case-referent studies there is a second type of confounding, one that is not based on correlates of (knowable) exposure in the joint source population of cases and referents (Examples 4–6) but on study procedures that bear on the accuracy of information on exposure. Such accuracy-related aspects of exposure-ascertainment are confounders if they are distributed differently between the index (case) and reference series.

ANALOGIES AND DIFFERENCES BETWEEN
FOLLOW-UP AND CASE-REFERENT STUDIES

In the above, criteria for confounding have been explored separately for follow-up and case-referent studies. For each type of study the criteria involve two components. One refers to the covariate's relationship to (observing) the illness and the other to its relationship to (observing) the exposure. The particulars of these relationships as criteria for confounding, as presented, may seem to be inconsistent between the two types of study. Thus, a comparative review may be called for.

The illness-related criterion for confounding in a follow-up study—that the covariate be a predictor of observing the illness—refers to a general tendency, a relationship in the abstract, and not to a relationship in the data at hand. It is a matter of complete analogy, then, that in case-referent studies the same abstract relatedness satisfies the illness criterion for an a priori confounder, while a difference between the distributions of the cases and noncases within the study does not.

That there is an alternative outcome-related a priori criterion in case-referent studies only—the covariate having differential selection implications between cases and noncases—reflects the fact that selectiveness by a covariate differing according to the health outcome can be a design feature in case-referent studies only. In follow-up studies the selection, while it may be related to determinants of outcome, is inherently independent of the outcome itself, as it is performed before the outcome is known.

A confounder's relationship to the exposure in a follow-up study is completely a matter of the data at hand: a confounder must show a difference in distribution between the exposed and nonexposed subjects within the study. Thus, in a follow-up study a confounder is associated with exposure in the population from which all cases and noncases observed in the study arise. This is exactly the exposure-related criterion that was given for a priori confounding in case-referent studies. The reason why the application of this criterion in a case-referent study necessarily involves a priori knowledge is

that, in contrast to a follow-up study, the joint source population of cases and non-cases is not studied in toto.

Matching in a follow-up study makes the covariate uncorrelated with exposure in the total study cohort of exposed and nonexposed subjects enrolled in the study, i.e., in the population from which the ultimate cases and noncases in the study arise. Thus, in a follow-up study matching assures that the exposure-criterion for confounding will not be satisfied. In a case-referent study, by contrast, it does not have the implication of making the covariate uncorrelated with exposure in the base population but constitutes, instead, a criterion of differential selection according to the health outcome among the members of the source population. It is a consequence of this that in case-referent studies matching gives no assurance about absence of confounding but is, instead, conducive to it.

The second type of confounding in case-referent studies—by outcome-associated factors of exposure-ascertainment—has no express counterpart in follow-up studies, because, in the latter, exposure status is ascertained before the outcome is known. On the other hand, there is an analogy in the sense that in both types of study a determinant of errors in the comparison criterion is a confounder if it has different distributions between the index and reference series (cf. Examples 1 and 10).

PROBLEMS WITH THE CHANGE-IN-ESTIMATE CRITERION

As was already noted, a commonly employed criterion for confounding in data-analysis is that control of the extraneous factor—by stratification or by a multivariate technic—changes the estimate of effect. Both methods of control are (and as a result the criterion itself is) totally dependent on the data at hand.

One mechanism by which such a criterion can lead to a false conclusion about confounding, related to the choice of the effect parameter, was already discussed (see Example 3.).

A second mechanism—rather peculiar to case-referent studies—for malfunctioning of this criterion has to do with variation among the stratum-specific estimates of the effect parameter of interest: control of the covariate can result—or fail to result—in a change in the estimate by virtue of a change in the relative "weightings" of the different stratum-specific estimates implied by the control—relative to the "weightings" inherent in the crude estimate.

Example 11. As an illustration of this, consider a case-referent study on the use of a particular drug, D_1, as a potential cause of agranulocytosis, a blood dyscrasia mostly caused by drugs, with concern about (the union of) known causal factors, D_2, as a possible confounding factor. Suppose the data, with D_2 ignored, are as presented in panel A of table 2, with the corresponding estimate for incidence-density ratio equal to 3.0 (21). If exposure to D_1 is not correlated with exposure to D_2 in the joint source population of cases and referents, then D_2 is not a confounder (in the a priori sense). Consistent with this, the data, when stratified by exposure to D_2, could be as shown in panel B of table 2, with estimates of 1.5 and 9.0 for the domains of nonidiopathic (D_2+) and idiopathic (D_2-) agranulocytosis, respectively. The maximum likelihood estimate for the overall incidence density ratio (ignoring modification, as usual) is 7.5, differing substantially from the crude estimate and, thus, giving a false indication of confounding. (Where control changes the estimate/parameter on account of appreciable modification, it may be unadvisable to consider an overall measure of effect, even if unconfounded, in contrast to measures specific to levels/categories of the covariate—an issue that falls outside the concerns in this paper.)

TABLE 2

Hypothetical case-referent data for Example 11, together with estimates for the incidence density ratio (IDR), illustrating that, particularly in case-referent studies, modification can cause malfunction of the change-in-estimate criterion of confounding

A) Unstratified data (unconfounded by exposure to other causes, D_2*)

	D_1		
	+	−	Total
Cases	25	75	100
Referents	10	90	100

$$\widehat{IDR} = 3.0$$

B) Data stratified by exposure to other causes, D_2 (a modifier of IDR)

	D_2+			D_2-		
	D_1			D_1		
	+	−	Total	+	−	Total
Cases	10	60	70	15	15	30
Referents	1	9	10	9	81	90

$$\widehat{IDR} = 1.5 \qquad \widehat{IDR} = 9.0$$

Overall $\widehat{IDR}_{ML}$† = 7.5

* D_1 and D_2 uncorrelated in the joint source of cases and referents.

† $\widehat{IDR}_{ML}$ = maximum likelihood estimate for IDR.

Yet another mechanism by which the change-in-estimate criterion can result in a false conclusion about confounding is based on chance elements in the association in the data between the covariate and the illness in follow-up studies or the covariate and the exposure in case-referent studies. (This problem is, of course, shared by the alternative criterion when based singularly on associations in the data in place of firm a priori knowledge.) It persists even when only "statistically significant" associations are used as the criteria (1, 3, 8, 13). In particular, "nonsignificance" constitutes little assurance about the absence of the associations at issue.

EPILOGUE

In the above, confounding has been examined with reference to data-analysis and inference rather than with a view to study design (to say nothing about Nature at large). The reason for this is that confounding in its ultimate essence is a problem with a particular estimate—a question of whether the magnitude of the estimate at hand could be *explained* in terms of some *extraneous* factor. This view is, we believe, shared by epidemiologists in general. In this sense, *confounding bias* in an estimate, as has been discussed in the above, may result from confounding in the study (or source) population per se, from the covariate's relationship to the way the population was sampled and/or from errors of observation associated with the covariate.

By "extraneous factor" is meant something other than the exposure or the illness—a characteristic of the study subjects or of the process of securing information on them.

"Explanation" by such a factor cannot refer to anything mechanistic—such as change in the value of the estimate as a result of control. Instead, "explanation" has to do with *understanding*, and this is inherently dependent on *a priori knowledge*. We have seen that each criterion of confounding involves a priori knowledge of the covariate's relationship to observing the illness and/or the exposure—knowledge that has to do with Nature per se, the setting in which the study is conducted, and/or the study protocol and related issues. (For subtleties on this, see Appendix 1.)

Not all explanation by a covariate is considered confounding. Thus, an unusually strong association between coronary heart disease risk and smoking may be explicable in terms of the study having been restricted to the domain of young people (22), and this is an explanation in terms of modification (8) instead of confounding: the distribution of modifiers within the study implies the target (unconfounded) value for the measure of effect, while confounding serves to explain the estimate's deviation from that target. Even the explanation of a discrepancy be-

tween such a target value and its estimate (in terms of an extraneous factor) is not necessarily confounding. For example, the empirical association between coronary heart disease risk and smoking may be much weaker than the target (theoretical) association on account of the study having been confined to a domain characterized by substantial misclassification as to the illness and/or the exposure. These two examples underscore a special feature of explanation in terms of confounding: it implies a need and desire to replace the "crude" estimate of effect by one that has been *adjusted* for the covariate at issue—so as to better represent the manifest aggregate effect in the type of setting actually studied (defined in terms of modifiers and determinants of accuracy of information).

This connection between confounding and adjustment helps explain why the criterion of change in estimate as a result of "control" of the covariate is so widely used as a criterion for confounding. However, correct application of the latter criterion requires that it be based on a comparison of the actual crude estimate with a particular kind of adjusted estimate—one that does not lose sight of confounding being a completely ad hoc problem in the sense just described (23). The observed number of exposed cases (O) and its null value (E) have this focus—their difference being the (estimated) number of cases attributable to the exposure in the particular framework of the study. Both measures (O and E) refer to the index series, since exposed cases occur in the index series only (the reference series serving the estimation of E). These measures, together with the size (S_1) of the index series, are sufficient for the expression of all desired measures of effect in the ad hoc sense that is of concern here. Thus, in a follow-up study, the estimate for rate difference is $\hat{RD} = (O - \hat{E})/S_1$; for rate ratio it is $O/\hat{E}$, and for illness odds ratio it is $[O/(S_1 - O)]/[\hat{E}/(S_1 - \hat{E})]$. In a crude es-

timate, $\hat{E}$ is computed on the basis of the crude overall observation in the reference series, and for the purpose of detecting confounding this crude result must be compared with the adjusted result which employs the same O but an adjusted $\hat{E}$—one which takes account of potential lack of direct comparability of the reference series in terms of the covariate at issue.

Example 12. As an illustration of this, consider again the "data" in table 1. The observed number of exposed cases (O) is 104. The crude estimate of the corresponding null-expected number ($\hat{E}$) is, from Panel A, $(96/200)200 = 96$. The corresponding adjusted $\hat{E}$ is the sum of the gender-specific values: $\hat{E} = (95/100)100 + (1/100)100 = 96$. Thus, the crude and the adjusted estimates of the odds ratio are the same, equal to $[104/(200-104)]/[96/(200-96)] = 1.2$, indicating no confounding (cf. Example 3).

Example 13. As an illustration of this principle in the context of a case-referent study, consider again the "data" in table 2. The observed number of exposed cases is 25. The crude estimate of the corresponding null-expected value is $(23)75(10)/90 = 8.33$—not $(10/100)100 = 10.00$. The $\hat{E}$ adjusted for other causes is $60(1)/9 + 15(9)81$, which equals the crude $\hat{E}$ of 8.33. Consequently, the rate ratio estimate, $O/\hat{E}$, is $25/8.33 = 3.0$ in both crude and adjusted terms (internally standardized terms, with the exposed in the source population as the standard (24)). This lack of change from the crude estimate to the adjusted estimate, computed properly for the detection of confounding, again indicates lack of confounding (cf. Example 11).

In these terms, the general topic of confounding seems to be gaining coherence. Confounding is a question of whether the crude estimate of effect—based on the observed and (estimated) expected number of exposed cases—is an adequate estimate of the overall effect under the conditions

of the study (regarding modifiers and factors bearing on accuracy of information); more specifically, confounding means that an adjustment of (the estimate of) the expected number for lack of comparability of the reference series is needed. This outlook disposes of any conflict between the two types of criteria for confounding—one based on the covariate's relationships to exposure and illness, and the other on the change in estimate (or parameter) as a result of control of (adjustment for) the covariate. Secondarily, it becomes clear that confounding does not have different criteria according to what parameter is being estimated.

The thought patterns involved in the detection of confounding were shown to be different between follow-up (and prevalence) studies on one hand and case-referent studies on the other. Yet, in the discussion in an earlier section they were already found to be mutually coherent, and the outlook in this section reinforces that conclusion.

REFERENCES

1. Miettinen OS. Stratification by a multivariate confounder score. Am J Epidemiol 1976;104:609-20.
2. Rothman KJ. A pictorial representation of confounding in epidemiologic studies. J Chronic Dis 1975;28:101-8.
3. Dales LG, Ury HK. An improper use of statistical significance testing in studying covariables. Int J Epidemiol 1978;7:373-5.
4. Miettinen OS. Reply by Dr. Miettinen. (Letter.) Am J Epidemiol 1977;106:191-3.
5. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. JNCI 1959;22:719-48.
6. Fleiss JL. Statistical methods for rates and proportions. New York: John Wiley & Sons, 1973.
7. Prentice R. Use of the logistic model in retrospective studies. Biometrics 1976;32:599-606.
8. Miettinen OS. Confounding and effect-modification. Am J Epidemiol 1974;100:350-3.
9. Bishop YMM, Fienberg SE, Holland PW. Discrete multivariate analysis. Cambridge, MIT Press, 1975.
10. Meinert CL, Knatterud GL, Prout TE, et al. A study of the effects of hypoglycemic agents on vascular complications in patients with adult-onset diabetes: II. Mortality results, UGDP. Diabetics 1970;19 (suppl 2):789-830.
11. Schor S. The UGDP—A statistician looks at the mortality findings. JAMA 1971;217:1671-5.

12. Cornfield J. The UGDP—A further statistical analysis of the mortality findings. JAMA 1971;217:1676-87.
13. Greenland S, Neutra R. Control of confounding in the assessment of medical technology. Int J Epidemiol 1980;9:361-7.
14. Yanagawa T. Designing case-control studies. Environ Health Perspect 1979;32:143-56.
15. Fisher L, Patil K. Matching and unrelatedness. Am J Epidemiol 1974;100:347-9.
16. Rothman KJ. Epidemiologic methods in clinical trials. Cancer 1977;39:1771-5.
17. Jick H, Slone D, Westerholm B, et al. Venous thromboembolic disease and ABO blood type. Lancet 1969;1:539-49.
18. Day NE, Byar DP, Green SB. Overadjustment in case-control studies. Am J Epidemiol 1980;112:696-706.
19. Matroos A, Magnus K, Strackee J. Fatal and nonfatal coronary attacks in relation to smoking in some Dutch communities. Am J Epidemiol 1979;109:145-51.
20. Magnus K, Matroos A, Strackee J. Walking, cycling, or gardening, with or without seasonal interruption, in relation to acute coronary events. Am J Epidemiol 1979;110:724-33.
21. Miettinen OS. Estimability and estimation in case-referent studies. Am J Epidemiol 1976;103:226-35.
22. Slone D, Shapiro S, Rosenberg L, et al. Relation of cigarette smoking to myocardial infarction in young women. N Engl J Med 1978;298:1273-6.
23. Miettinen OS. Components of the crude risk ratio. Am J Epidemiol 1972;96:168-72.
24. Miettinen OS. Standardization of risk ratios. Am J Epidemiol 1972;96:383-8.

APPENDIX 1

Prior information versus data

In the context of a confounder's relationship to the health outcome, the relative roles of prior information on one hand and the data from the study at issue on the other involve subtlety beyond what is evident from Example 2.

Consider the "null outcomes" in the study population, say in the subjects in a clinical trial. For the nonexposed, they are the actual outcomes, whereas for the exposed they are the hypothetical outcomes that would have materialized had the exposure had no effect. If it were known that in the study population, combining the exposed and nonexposed, a covariate is associated with this null outcome and also with the exposure, it would be concluded that the covariate is a confounder—regardless of whether the

covariate tends to have these associations in general.

Example A. Extending Example 2, suppose that the null outcomes regarding coronary mortality were known for the tolbutamide series just as they actually became known for the placebo series. Suppose further that, in the two series combined, there was no relationship between the null outcome and age. It would be concluded that, even with different age-distributions between the two series, age would not be a confounder in the study. That the null outcome would tend to be related to age in general would be irrelevant in the face of this knowledge about the study population at issue. On the other hand, knowing only that the actual outcomes are unrelated to age does not mean that the null outcomes are, owing to the possibility that the effect of tolbutamide is modified by age (8).

Example B. Extending Example 2 further, suppose the result from a random-number generator was recorded for each patient, and that this (truly) random variate turned out to be associated with the null outcome and also with the exposure. This would have meant that the random number was a confounder, even though the generator has no value in predicting coronary mortality.

Thinking of confounding in terms of the necessary analysis-of-covariance allowance for the covariate under the simplest possible model, the question is what coefficient should be used for the covariate. For example, in the context of no confounding the coefficient should be set to zero. Let

$\hat{\beta}$ = coefficient obtained by fitting in the nonexposed cohort (e.g., the placebo series in a clinical trial), and

$\hat{\beta}^*$ = coefficient that would be obtained by fitting in the total cohort, with null outcomes for the exposed as well as the nonexposed.

β = coefficient that represents "the true value" in general (such as $\beta = 0$ for the result from a random-number generator).

The ideal adjustment involves $\hat{\beta}^*$ rather than $\hat{\beta}$ or β, as has been discussed in the above. The practical question is how to infer $\hat{\beta}^*$ in the face of the data and whatever may be known about β apart from the data at hand. We have no general answer to this question.