AMERICAN JOURNAL OF EPIDEMIOLOGY
Vol. 127, No. 5
Printed in U.S.A.

ANALYSIS OF CASE-CONTROL DATA DERIVED IN PART FROM PROXY RESPONDENTS

ALEXANDER M. WALKER,[1,2] JOHAN P. VELEMA,[2] AND JAMES M. ROBINS[1]

In the conduct of case-control studies, it has been a general experience that some cases are identified only after death, and that others cannot be personally interviewed because of extreme morbidity. As a way of dealing with the dilemma of obtaining data on unavailable cases, investigators have sometimes chosen to interview a relative or friend (a proxy respondent) whose responses are based on his or her knowledge of the lifestyle, medical history, occupational history, etc., of the unavailable index subject. The same procedure may be used for controls who are unable to respond personally to the questions of the investigator.

Some methodological inquiries have been carried out to assess the agreement between index (i.e., personally interviewed) and proxy respondents. In these studies, the proxy respondents have usually been the spouses of the index respondents, and the questions have most often concerned the frequency and/or quantity with which foods were consumed (1–4). Results from these studies show a high concordance between group level aggregate measures, such as mean nutrient intake or mean frequency of consumption, estimated from index and proxy responses. On an individual level, however, dietary information obtained from proxy respondents appears to have been a relatively crude substitute for index responses, as evidenced by kappa values that rarely exceed 0.5.

It is difficult to generalize about the utility of different kinds of proxy respondents. Humble et al. (4) reported, for example, that kappa values were lower when men responded to food frequency questionnaires for their wives than when the wives did so for their husbands, suggesting that husbands may be less informative proxy respondents than their wives in dietary studies. By contrast, Cahalan (5) indicated that men know more about their wives' drinking habits than do women about their husbands' drinking habits. Wives' recall of their deceased husbands' jobs has been found to vary according to the duration of the job (6).

The problems introduced by the use of proxy respondents can be seen in data from Rogot and Reid's (7) report of spouse responses to questionnaires that were similar to ones administered several years earlier to index subjects who had died in the interim (table 1). To illustrate the effect, we have assigned a "tobacco exposure score" to each individual, based on his self-reported level of tobacco use, and we have calculated the average score across the rows of table 1, which correspond to categories of proxy-reported tobacco use. The row averages provide an index of self-reported tobacco exposure within each category of proxy response. Proxy respondents have produced a net overestimate of exposure in this case: The average exposure score computed from proxy responses is $(1{,}186 * 0 + 104 * 0.1 + 171 * 0.5 + 278 * 1.0 + 214 * 1.5)/1{,}953 = 0.36$, as opposed to 0.31 from the index responses. More seriously, the true exposure differences between the

[1] Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115. (Reprint requests to Dr. Walker at this address.)

[2] International Agency for Research on Cancer, 150 Cours Albert Thomas, 69372 Lyon Cedex 08, France.

This work was supported by the International Agency for Research on Cancer (Drs. Walker and Velema) and by a grant from the Mellon Foundation (Dr. Walker).

This report was developed as a service to collaborators in the SEARCH case-control studies network.

The authors thank Drs. S. Greenland and J. Samet who provided valuable suggestions and advice.

TABLE 1

Cigarette smoking from index and next-of-kin interviews, British and Norwegian immigrants to the United States

Proxy response	Index response					Total	Average index score
	NC* (0)†	OC (0.1)	<1 (0.5)	1 (1.0)	>1 (1.5)		
NC (0)†	1,133	13	27	7	6	1,186	0.03
OC (0.1)	49	18	31	5	1	104	0.23
<1 (0.5)	26	4	92	43	6	171	0.58
1 (1.0)	30	6	90	113	39	278	0.78
>1 (1.5)	9	1	39	73	92	214	1.08
Total	1,247	42	279	241	144	1,953	0.31

* Response abbreviations: NC, not current smokers; OC, occasional smokers; <1, 1, >1, less than one, one, more than one pack per day current smoker, respectively.

† Cigarette exposure level score given in parentheses.

proxy-derived categories are smaller than those between the index-derived categories.

Investigators have employed a variety of devices to deal with proxy information in their data sets. For example, McLaughlin et al. (8) selected two control groups, one composed of living, directly interviewable persons, the other of deceased persons, for whom proxy respondents were interviewed; the information for cases, obtained from either index or proxy respondents, was compared with the data for both control groups. Heyman et al. (9) were forced to obtain case information from proxies in a study of Alzheimer's disease, because cases with this condition could not respond directly; the investigators obtained both index and proxy interviews for controls. Presentation of results included odds ratios obtained from proxy case-control comparisons and a corresponding measure of proxy-index agreement, calculated for each variable from the controls.

Multiple interview protocols can be expensive and difficult to implement. Some investigators have therefore simply included the information obtained from proxy respondents in their data sets as equivalent to the information obtained from index respondents (10, 11). Others (12, 13) have adjusted estimates of measured associations of exposure and disease by including in their logistic models an indicator variable to distinguish proxy from index respondents. When there is one exposure of interest, estimates obtained by this latter procedure are equivalent to the maximum likelihood estimates of an odds ratio that is assumed to be constant over the strata defined by respondent status.

While each of these approaches may be defensible under particular circumstances, criteria by which to judge the appropriateness of any one approach are not well established, and the data presented rarely permit the reader to form an independent opinion on this point. In the following, we suggest a course of analysis of mixed proxy-index data sets that will provide point and interval estimates of effect whose bias is approximately known. The analytic procedure that we will recommend should also give insight into the divergence between index- and proxy-based inference in a particular analysis.

STATISTICAL CONSIDERATIONS

For the purpose of developing our arguments, we will assume in this section 1) that disease is correctly classified, 2) that proxy response is the sole source of classification error, 3) that the probabilities of misclassification are the same among proxy respondents for cases and proxy respondents for controls, and 4) that among cases, the probability of obtaining an index interview is independent of exposure status, and similarly among controls.

In the absence of misclassification, the ratio of case to control counts at each exposure level corresponds to the disease incidence rate in the underlying population from which cases and controls are drawn, divided by the ratio of sampling parameters for cases and controls. The latter may be unknown, and as a result the ratio of the incidence rates in the two exposure categories, rather than the incidence rates themselves, is estimated by dividing one of the case-control ratios by the other. This value, the odds ratio, is the usual measure derived from case-control analyses.

The net effect of exposure misclassification is to permit migration of cases and controls between exposure categories. It is not possible in the general case to predict the direction of bias that results from misclassification, even under the assumption that identical reclassification matrices apply to the cases and controls. Table 2 gives an example of a reclassification matrix in which extreme values of exposure tend to be reported as intermediate values; a process characterized by the reclassification scheme summarized in the matrix results in no bias at all in point estimates of the odds ratios when applied to the "true" data of table 2. The lack of distortion arises from the flow of heavily and lightly exposed subjects into the moderately exposed category in such a way that the average exposure intensity in the latter category is unchanged. It would be possible to construct other hypothetical reclassification matrices under which a movement of heavily exposed persons into moderate exposure categories resulted in a positive bias in the moderate exposure odds estimates, or in which an opposite flow yielded a negative bias in the high exposure odds estimates. When there is reclassification in both directions between two categories, the corresponding odds are drawn toward one another. If they are changed at all, odds in the categories of extreme risk can only be drawn by misclassification toward the values of less extreme categories. Dichotomous exposure classifications furnish a special case in which each class is an extreme, and therefore any misclassification leads to a bias toward the null in the odds ratio between the classes.

TABLE 2

A reclassification matrix that does not give rise to bias in odds ratio estimates

The matrix

Reported exposure level	True exposure level		
	Low	Moderate	High
Low	0.75	0.00	0.00
Moderate	0.25	1.00	0.25
High	0.00	0.00	0.75

True data

Disease status	Exposure		
	Low	Moderate	High
Cases	100	200	300
Controls	200	200	200
Odds ratio	1	2	3

Reported data

Disease status	Exposure		
	Low	Moderate	High
Cases	75	300	225
Controls	150	300	150
Odds ratio	1	2	3

In case-control studies, the proportion of proxy respondents (whose responses result in misclassification) is usually greater among cases than among controls, with the result that overall misclassification probabilities differ between cases and controls. Crude analysis of partially misclassified data can distort effect estimates and render tests of significance invalid. At a minimum, therefore, it would seem necessary to segregate misclassified (i.e., proxy-derived) case and control data into strata with homogeneous misclassification probabilities, leaving the index-derived data in a stratum whose point estimate of the incidence ratio is unbiased. Within the misclassified stratum, any one case-control ratio may be biased toward or away from any other, except at the extremes of risk, where the odds must be biased toward one another if they are biased at all.

We will now consider stratified analyses, beginning with the effects of nondifferential misclassification on significance tests. If the null hypothesis of no effect holds, the size of a statistical test of the null hypothesis (such as the Mantel-Haenszel test (14) or the Mantel-Haenszel extension procedure (15)) is preserved under misclassification. Even so, as an alternative to any stratified test, one could consider performing the corresponding nonstratified test in the index stratum only. Depending on the degree of misclassification in the proxy stratum, the power of the stratified test may be either smaller or greater than that of the nonstratified test performed exclusively in the index stratum. At first, the possibility of a loss of power resulting from the addition of proxy data to an analysis may seem counterintuitive. Consider, however, the extreme case of misclassification, when the probability of assignment to a given exposure category is independent of a subject's true exposure status: The proxy stratum then represents pure noise, and its inclusion necessarily diminishes the power of a test.

Greenland and Robins (16) have shown in the case of dichotomous exposure classification that misclassification rates that vary between strata (e.g., between proxy and index strata) will introduce a spurious appearance of heterogeneity in the odds ratio. When generalized to polytomous or continuous exposure measures, however, the result as demonstrated in table 2 can only be that misclassification *may* introduce spurious heterogeneity. When misclassification fails to introduce heterogeneity in point estimates, however, it also produces index and proxy strata which can be collapsed without introduction of bias. As in the example presented in table 2, this situation can hold only when the case-control ratios at all exposure levels are independent of proxy status under misclassification; the pooling of strata with identical case-control ratios produces a crude stratum whose case-control ratios are those of the component strata. Tests of heterogeneity of the odds ratio between strata are therefore valid (albeit weak) tests of the hypothesis that an analysis that collapses data over strata of respondent status is unbiased. In the case of dichotomous exposures, where misclassification necessarily implies heterogeneity, and vice versa, in the nonnull case, tests of heterogeneity between strata are tests of misclassification.

Implications for analysis

We propose the following steps in an analysis of data from case-control studies in which proxies have been used: Stratify the data by respondent status, and estimate the relative biases in the proxy stratum by dividing the proxy-derived estimates of effect at every exposure level by the index-derived estimates. Confidence intervals for the biases may be estimated either by taking the variance of the logarithm of the bias as the sum of the variances of the logarithms of the point estimates, calculated for example by Woolf's (17) procedure, or they may be estimated from a one degree of freedom heterogeneity chi-square, using a test-based procedure (18). If the relative biases are large in the substantive context of the study, or if they are incompatible with the null hypothesis of homogeneity, only the stratum-specific estimates should be presented. A global test of the null hypothesis of no association between exposure and disease, performed conditionally on respondent status, is valid in the sense of having its nominal significance level, but has the dual disadvantages of having an unknown power relative to a test limited to the index stratum, and of being unrelated to any confidence interval. If the stratum-specific odds ratios appear to be homogeneous, and if in addition the nature of the data collection mechanism makes important misclassification of proxy responses unlikely a priori, proxy and index strata may be collapsed. Even in the absence of a strong prior hypothesis about the quality of data derived from proxy respondents, if the confidence interval for the estimate of bias excludes values that would be of any

consequence in the context of the study, the proxy and index strata may be collapsed.

The logic of the analytic sequence for dealing with misclassification suggested above is "if heterogeneous, then separate, otherwise collapse." In practice, an analyst may be uncertain, even after careful review of data collection procedures and after performing appropriate tests, whether or not heterogeneity exists. A possible middle road under uncertainty is to obtain a single summary estimate over strata such as the Woolf (17), maximum likelihood (19), or Mantel-Haenszel (14) estimate of the common odds ratio. In the absence of heterogeneity induced by misclassification, all of these estimates will be consistent, and more precise than an estimate derived from the index stratum alone. They will, however, be less precise than the crude estimate, which is also consistent. If there truly is misclassification in the proxy stratum, a common estimate of a single odds ratio for a dichotomous exposure will be biased toward the null, although not so much as the estimate derivable from the proxy stratum. The crude estimate may be biased either toward or away from the null (16). The direction and magnitude of this bias depend on the actual probabilities of misclassification (which differ for different kinds of respondents), on the probabilities of exposure for cases and controls, and on the relative proportion of proxy respondents among cases and controls. Examples of this phenomenon have been presented by Greenland and Robins (16). Investigators who prefer a potential bias of known direction and estimable magnitude in the common estimate to one of uncertain direction and magnitude in the crude estimate may be willing to sacrifice the increased precision that the crude supplies if misclassification is truly absent. When data from proxy respondents are necessary to augment the information from directly interviewed subjects, and when there is residual uncertainty as to whether substantial misclassification exists, we recommend supplementing crude and stratum-specific estimates with the presentation of a common odds ratio estimate across strata defined by respondent status.

In a logistic analysis, the homologous procedure is as follows: Include in the predictor model an indicator variable which distinguishes between index and proxy respondents, and include interaction terms between this indicator variable and the exposure variables in the model. Calculate point and interval estimates for the interaction terms, which measure the bias in the proxy stratum relative to the index stratum. If any of the interaction terms is large in the context of the study, separate estimates of the corresponding exposure-disease association parameters should be presented for index and proxy respondents. A valid summary test of exposure-disease association is obtained by removing interaction terms from the model and testing the main exposure-disease effect. As indicated above in the context of stratified analyses, the main effect of respondent status should be retained at this point if the analyst wishes to ensure that the direction of any bias resulting from misclassification be of known direction (toward the null). The resulting test has, as above, the disadvantage of unknown power. If all interaction terms are unimportant, but there remains some doubt about underlying heterogeneity, an adjusted estimate of the association between exposure and disease should be derived by retaining a term for respondent status, deleting all interaction terms between respondent status and exposure, and presenting the coefficients of the exposure terms as the result of the analysis. As before, if there is a strong a priori expectation of similarity of proxy and index responses, and if there is no demonstration of heterogeneity in the data, all terms involving respondent status may be dropped from the model.

No matter which method is employed to account for respondent status, presentation of results for each respondent category, of crude results, and (if appropriate) of sum-

mary results is advisable: It provides some evidence as to the magnitude of possible distortion in the data at hand, and it offers a useful estimate for readers whose a priori beliefs in the validity of the proxy interview may differ from those of the authors.

It seems likely that different probabilities of misclassification will apply to different kinds of proxy respondents and to different interview situations. One could therefore naturally extend the proposed analysis by stratifying according to the relationship of the proxy respondent to the index subject (husband, wife, child, colleague, etc.), according to the number of years for which the proxy respondent had direct knowledge of the index subject's habits, or according to the time elapsed since the index subject's death.

EXAMPLE

To illustrate the methods described above, we have used the reclassification matrix implied in table 1 to create a hypothetical study of 400 cases and 800 controls. The marginal distribution of self-reported cigarette use in Rogot and Reid's (7) study has been taken as the distribution in the control group, and the distribution among the cases has been computed assuming odds ratios of 1.5, 2, 3, and 3.5, respectively, for occasional smokers, and smokers of less than one, one, and more than one pack of cigarettes per day, all relative to nonsmokers. For 50 per cent of the cases and 10 per cent of the controls at every exposure level, proxy rather than index respondents are assumed to have been interviewed. The constructed data are shown in table 3. The crude data show a distortion of the exposure-response function, with an interchange in the relative risk of the two lowest smoking categories.

The stratified data present a clearer picture. Taking nonsmokers as the reference stratum, estimates from proxy respondents are biased toward the null in the two highest exposure categories and away from the null for occasional smokers. Although the biases in the proxy stratum show a monotonic trend (table 4), they are not large (a posteriori, they might easily be thought to have resulted from the play of chance), and, depending on the purposes of the study,

TABLE 3

Data from a hypothetical case-control study of the relation between disease incidence and cigarette use

	Cigarette Use					
	NC*	OC	<1	1	>1	Total
			True data			
Cases	161	8	73	93	65	400
Controls	511	18	114	98	59	800
Odds ratio	1	1.4	2.0	3.0	3.5	
			Observed data			
Crude						
Cases	161	14	60	92	73	400
Controls	509	20	110	99	62	800
Odds ratio	1	2.2	1.7	2.9	3.7	
Index respondents						
Cases	80	4	36	47	33	200
Controls	460	16	103	88	53	720
Odds ratio	1	1.4	2.0	3.1	3.6	
Proxy respondents						
Cases	81	10	24	45	40	200
Controls	49	4	7	11	9	80
Odds ratio	1	1.5	2.1	2.5	2.7	

* Response abbreviations: NC, not current smokers; OC, occasional smokers; <1, 1, >1, less than one, one, more than one pack per day current smoker, respectively.

TABLE 4

Summary measures obtained from the analysis of data from a hypothetical case-control study of the relation between disease incidence and cigarette use

	Cigarette use			
	OC*	<1	1	>1
Mantel-Haenszel summary odds ratio	1.5†	2.0	2.9	3.2
Proxy stratum bias	1.05	1.03	0.81	0.75
95% confidence interval for bias				
Woolf‡	0.20–5.5	0.37–2.8	0.34–1.9	0.29–1.9
Test-based§	0.18–6.0	0.41–2.6	0.35–1.9	0.29–2.0
Logistic models				
No response status indicator				
Exposure terms	2.2	1.7	2.9	3.7
Response status indicator included				
Exposure terms	1.4	2.0	3.1	3.6
Interaction terms	1.05	1.03	0.81	0.75
95% confidence interval for interaction	0.20–5.5	0.37–2.8	0.34–1.9	0.29–1.9
Omitting interaction				
Exposure terms	1.5	2.0	2.9	3.3

* Response abbreviations: OC, occasional smokers; <1, 1, >1, less than one, one, more than one pack per day current smoker, respectively.

† All values take noncurrent smokers as the reference exposure level.

‡ Reference 17.

§ Reference 18.

they might well be considered as acceptable limits to the level of bias induced by respondent status in the full analysis. Mantel-Haenszel summary estimates of the category-specific odds ratios are found to preserve the ordering and relative magnitudes of risk found in the underlying "true" data and in each of the strata. Accounting in the analysis for respondent status reveals an exposure-response relation which would not have been observed in the crude data.

The homologous analysis, carried out through logistic regression in GLIM (20), is also summarized in table 4. Indicator variables have been used to represent smoking effects and respondent status. In the absence of an indicator variable for proxy status, the coefficients of the exposure terms, transformed to a relative risk scale, simply reproduce the results of the crude analysis. Inclusion of a respondent status term and its interactions with exposure yields exposure effect estimates that correspond to those of the index stratum and interaction effects that provide estimates of the bias in the proxy stratum. Omission of the interaction terms leads to summary exposure effect estimates similar to the Mantel-Haenszel estimates.

CAVEATS

In the development above, we have assumed that the probabilities of misclassification, given respondent status, are the same for cases and controls. If proxy respondents for cases are more or less likely to give erroneous responses than proxies for controls, a bias will result that cannot be corrected for in the analysis. A discrepancy between misclassification probabilities according to case status may arise, for example, when proxy respondents are chosen for healthy controls in an attempt to match on respondent status for deceased cases. The use of data from proxy respondents in this instance invokes the hypothesis that error in proxy responses is unaffected by the vital status of the index subject. We cannot offer any data to support or refute this hypothesis. Greenberg et al.

(21), however, have provided an example of a closely related phenomenon: They found errors in proxy respondents' reports of occupation to be related to the cause of death of index subjects.

We have also assumed that the probability that an index interview will be obtained is independent of exposure status. To see the role of exposure, consider the notation presented in table 5, in which p_{ij} is the probability of obtaining an index interview in the category of respondents indexed by the ith disease status and jth exposure status. If

$$p_{ij} = p_{i\cdot} \text{ for all } i \text{ and } j,$$

the ratio of case-control index interview probabilities (p_{1j}/p_{2j}) will be constant (and equal to $p_{1\cdot}/p_{2\cdot}$) over all exposure levels, and the resulting odds ratios between exposures will provide consistent estimates of the underlying effect parameters. In the absence of misclassification, the same would hold for the proxy stratum, in which the ratio of interview probabilities, $(1 - p_{1j})/(1 - p_{2j})$, at every exposure level equals $(1 - p_{1\cdot})/(1 - p_{2\cdot})$. This argument for lack of distortion holds of course whether it is the probability of index interview or the probability of proxy interview that is represented by p_{ij}.

If the probability of obtaining an index interview varies with exposure and with disease status in an independent fashion, that is, if

$$p_{ij} = p_{i\cdot}p_{\cdot j} \text{ for all } i \text{ and } j,$$

the ratios of interview probabilities remain constant across exposure levels in the index stratum, and the resulting odds ratios are consistent estimates of the underlying parameters. In the proxy stratum, however, the terms in $p_{\cdot j}$ in the ratios of interview probabilities do not cancel out, and the ratios of interview probabilities vary from one exposure level to another, giving rise to incorrect odds ratio estimates, even in the absence of misclassification. The proxy stratum in this case is directly analogous to the set of hospitalized patients that Berkson (22) considered in his discussion of bias in hospital-based case-control studies. If it is not the probability of obtaining an index interview but rather the probability of obtaining a proxy interview that varies independently with both exposure and disease status, estimates from the index stratum are biased, while those from the proxy stratum are distorted by misclassification. In reality, the relation of exposure status to the probability of interview is unlikely to have a simple structure under any parameterization in the nonnull case. A search for unbiased subsets of the data would then be fruitless.

Finally, if the probability of obtaining an index interview varies with both exposure and disease in any way other than the simple multiplicative relation given above, the ratios of both index and proxy interview probabilities will vary over exposure levels, with consequent distortions in the odds ratios in both strata. In studies in which proxy interviews are not obtained when the index is unavailable, this last stricture is identical to the familiar result that the probabilities of case and control selection cannot be differently dependent on exposure if one is to obtain valid odds ratio estimates.

We have also assumed that proxy responses were misclassified and that the index responses were not. In reality, both

TABLE 5

Notation for probability of interview

	Index interviews Exposure level					
	1	...	j	...	n	Total
Cases	p_{11}		p_{1j}		p_{1n}	$p_{1\cdot}$
Controls	p_{21}		p_{2j}		p_{2n}	$p_{2\cdot}$
Total	$p_{\cdot 1}$		$p_{\cdot j}$		$p_{\cdot n}$	$p_{\cdot\cdot}$
	Proxy interviews Exposure level					
	1	...	j	...	n	Total
Cases	$1 - p_{11}$		$1 - p_{1j}$		$1 - p_{1n}$	$1 - p_{1\cdot}$
Controls	$1 - p_{21}$		$1 - p_{2j}$		$1 - p_{2n}$	$1 - p_{2\cdot}$
Total	$1 - p_{\cdot 1}$		$1 - p_{\cdot j}$		$1 - p_{\cdot n}$	$1 - p_{\cdot\cdot}$

proxy and index responses will be misclassified to some extent for most factors, and it may not be clear whether the classification based on information from proxy respondents is better or worse than the classification obtained from index respondents. In some situations, a proxy respondent may be better informed than the index respondent, or his or her disease may have affected the capacity of the index subject to respond adequately, as in Heyman et al.'s study of Alzheimer's disease (9). Under such circumstances, estimation of the true population odds ratio is in fact improved by the use of proxy respondents. In practice, a researcher who detects a higher odds ratio among proxy respondents than among index respondents (as did Greenberg et al. (21)) will want to assure himself first that this difference is not due to differential misclassification between the proxies who respond for cases and those who respond for controls. The possibility that proxy interviews are more accurate than direct interviews must be entertained next. We have not considered here the analysis of data arising from proxy and direct interviews when both yield misclassified results.

The crude odds ratio which is not adjusted for respondent status can be biased in either direction away from the parameter value. Thus, in some situations, an adjusted effect estimate may be further removed from the underlying parameter value than is the unadjusted estimate. Unless the probabilities of misclassification for index and proxy respondents can be known from an information source external to the data, however, it is impossible to appreciate the bias of the unadjusted effect measure. In the absence of such external sources of information, we suggest that an analysis such as that presented here, in which controllable sources of distortion have been removed and the direction and approximate magnitude of those remaining are estimable, is to be preferred.

Implications for design

Decisions on the use of proxy respondents in a case-control study have to be made in the design stage. Once the necessity of stratification by respondent status is accepted, there is no theoretical reason not to select controls for whom proxy respondents will be required, provided that a suitable sampling frame is available, such as a list of recently deceased persons from the population at large. Gordis (23) has pointed out that since death may have resulted from diseases related to the exposures under study, the selection of dead controls may distort estimates of the prevalence of these exposures in the source population and therefore bias estimates of effect. Proceeding on an assumption that proxy responses were equivalent to those that would have been obtained from index subjects, McLaughlin et al. (24) have recently provided evidence for the anticipated distortion of exposure distributions in the cases of smoking, drinking, drug use, and some adulthood diseases. Bias arising from association between exposures under study and diseases of the reference population can be handled by selection of causes of death not related to the exposure, as has been shown to be effective in hospital-based case-control studies (25), and strongly suggested to be effective in studies employing dead controls to ascertain the prevalence of a variety of exposures (26).

We have not addressed matched designs in this review, but it should be noted that the procedures suggested for logistic regression can be translated immediately into conditional logistic regression as well. If respondent status has been a matching criterion, conditional procedures necessarily involve conditioning on respondent status; heterogeneity of exposure effect over categories of respondent status may nonetheless be examined in much the same ways suggested above.

An association between exposure and proxy status need not be directly causal to lead to distorted estimates of prevalence: McLaughlin et al. (24) have shown that the prevalence of reported alcohol use in dead controls overestimates alcohol use in the general population, even after exclusion of alcohol-induced diseases from the causes of

death admitted into the decedent series. The overestimate disappears after control for smoking, indicating that the residual association may stem from overrepresentation of drinkers among persons who die of tobacco-related causes, presumably the result of an association between smoking and drinking in the general population.

Proxy interviews have been in use in epidemiologic studies for a long time, yet there is relatively little published data on their usefulness. In part, this may be because the differences between proxy and index responses can present an annoying, even embarrassing, sidelight to studies whose purpose is more substantive. Since the use of proxy respondents is pervasive, however, we should all benefit from seeing some disaggregation of proxy and index data whenever proxy respondents are used. The circumstances under which proxy and index responses are similar and those under which they are dissimilar deserve to be well documented, so that the epidemiologic research community as a whole can begin to accumulate experience of the utility of this common data-gathering tool.

References

1. Moore MC, Moore EM, Beasley CDH, et al. Dietary-atherosclerosis study on diseased persons. J Am Diet Assoc 1970;56:13–22.
2. Kolonel LN, Hirohata T, Nomura AMY. Adequacy of survey data collected from substitute respondents. Am J Epidemiol 1977;106:476–84.
3. Marshall J, Priore R, Haughey B, et al. Spouse-subject interviews and the reliability of diet studies. Am J Epidemiol 1980;112:675–83.
4. Humble CG, Samet JM, Skipper BE. Comparison of self- and surrogate-reported dietary information. Am J Epidemiol 1984;119:86–98.
5. Cahalan D. Quantifying alcohol consumption: patterns and problems. Circulation 1981;64(suppl III):7–14.
6. Lerchen ML, Samet JM. An assessment of the validity of questionnaire responses provided by a surviving spouse. Am J Epidemiol 1986;123:481–9.
7. Rogot E, Reid DD. The validity of data from next-of-kin studies of mortality among migrants. Int J Epidemiol 1975;4:51–4.
8. McLaughlin JK, Mandel JS, Blot WJ, et al. A population-based case-control study of renal cell carcinoma. JNCI 1984;72:275–84.
9. Heyman A, Wilkinson WE, Stafford JA, et al. Alzheimer's disease: a study of epidemiological aspects. Ann Neurol 1984;15:335–41.
10. Hinds MW, Kolonel LN, Hankin JH, et al. Dietary vitamin A, carotene, vitamin C and risk of lung cancer in Hawaii. Am J Epidemiol 1984; 119:227–37.
11. Gold EB, Gordis L, Diener MD, et al. Diet and other risk factors for cancer of the pancreas. Cancer 1985;55:460–7.
12. Pickle LW, Greene MH, Ziegler RG, et al. Colorectal cancer in rural Nebraska. Cancer Res 1984;44:363–9.
13. Winn DM, Ziegler RG, Pickle LW, et al. Diet in the etiology of oral and pharyngeal cancer among women from the southern United States. Cancer Res 1984;44:1216–22.
14. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. JNCI 1959;22:719–48.
15. Mantel N. Chi-square tests with one degree of freedom: extensions of the Mantel-Haenszel procedure. J Am Stat Assoc 1963;58:690–700.
16. Greenland S, Robins JM. Confounding and misclassification. Am J Epidemiol 1985;122:495–506.
17. Woolf B. On estimating the relation between blood group and disease. Ann Hum Genet 1955;19:251–3.
18. Miettinen OS. Estimability and estimation in case-referent studies. Am J Epidemiol 1976; 103:226–35.
19. Gart JJ. Point and interval estimation of the common odds ratio in the combination of 2 × 2 tables with fixed marginals. Biometrika 1970; 57:471–5.
20. Baker RJ, Nelder JA. The GLIM system release and generalised linear interactive modeling. London: Royal Statistical Society, 1978.
21. Greenberg ER, Rosner B, Hennekens C, et al. An investigation of bias in a study of nuclear shipyard workers. Am J Epidemiol 1985;121:301–8.
22. Berkson J. Limitations of the application of four-fold table analysis to hospital data. Biometrics Bull 1946;2:47–53.
23. Gordis L. Should dead cases be matched to dead controls? Am J Epidemiol 1982;115:1–5.
24. McLaughlin JK, Blot WJ, Mehl ES, et al. Problems in the use of dead controls in case-control studies. I. General results. Am J Epidemiol 1985;121:131–9.
25. Silverman DT, Hoover RN, Swanson GM. Artificial sweeteners and lower urinary tract cancer: hospital vs. population controls. Am J Epidemiol 1983;117:326–34.
26. McLaughlin JK, Blot WJ, Mehl ES, et al. Problems in the use of dead controls in case-control studies. II. Effect of excluding certain causes of death. Am J Epidemiol 1985;122:485–94.