AMERICAN JOURNAL OF EPIDEMIOLOGY
Vol 106, No 3

Printed in U.S.A

RESPONSE AND FOLLOW-UP BIAS IN COHORT STUDIES

SANDER GREENLAND

Selection bias has often been considered an especially serious validity problem in case-control studies (1). A hypothetical example of this sort of bias in essentially a case-control study setting was first published by Berkson (2) in 1946. The recognition of the potential for such bias in cohort (follow-up) studies dates back no less than 20 years (3, 4) and a detailed consideration of its possible effects in several studies has been given (5). A hypothetical example of selection bias in a cohort study is presented here in an attempt to examine criteria for unbiased response and follow-up in cohort studies.

For this discussion, study loss will be taken to mean total proportional study loss. In cohort studies, this is the proportion of persons who refuse to participate in the study (nonresponse) plus the proportion of persons who enter the study only to drop out before any useful follow-up of their outcome status can be made. Study participation will be taken to mean the complement of study loss. In the best known example of heavy loss, the Framingham Study (1) experienced a 31 per cent refusal rate among originally invited individuals and a 13 per cent dropout rate among the participants. When such losses have been uneven in both the exposure and outcome categories, it is generally recognized that the validity of the statistical results derived from the study (such as risk ratios) may be affected. However, the concepts of unbiased response and follow-up (6) have, on occasion, been misinterpreted. Specifically, it is sometimes thought that non-response and losses which are evenly distributed with respect to either the exposure or the outcome categories cannot alter the risk ratios derived from the study. Unfortunately, these interpretations can be misleading, as the following examples illustrate.

LOSS EFFECTS IN A HYPOTHETICAL STUDY

As a first example, let us suppose an investigator conducted a five-year follow-up study of the risk of developing a disease, D, associated with an exposure factor, E. Suppose further that the investigator knew from previous information that the five-year incidence of D should be about 50 per 1000 in his study cohort. 250 cohort members were classified as exposed at invitation to the study. Thirteen per cent of the exposed and 14 per cent of the unexposed refused to participate, while an additional 14 per cent of the exposed and 14 per cent of the unexposed entered the study but were lost to follow-up. Hence, a total of 27 per cent of the original exposed group and 28 per cent of the original unexposed group were of unknown outcome, so that losses were similar with respect to the exposure variable. Such losses have sometimes been referred to as being unbiased with respect to exposure.

The experience of the cohort members remaining under the observation of the investigator is summarized in table 1. The investigator observed a 105 per cent greater risk of D in the exposed group than in the unexposed group, and the difference in incidence rate between the two groups was significant at the .025 level. Further, the incidence rate in the cohort closely matched that expected from previous information. Some researchers interpret such a result as being demonstrative of

Abbreviations. D, disease; E, exposure factor

From the Division of Epidemiology, University of California School of Public Health, Los Angeles, CA 90024

This work was partially supported by USPHS grant 2 Ao3 AH00526-02.

TABLE 1

*Experience of the portion of the cohort observed by the investigator**†*

	Disease *D*	Not *D*	
Exposed	16	167	183
Not Exposed	23	517	540
	39	684	723

* Relative risk = 2 05 (χ^2 = 5.38, d.f. = 1, p <.025).

† Five year incidence = 53 per 1000 (not significantly different from the expected incidence of 50 per 1000).

unbiased response and follow-up with respect to the outcome variable. Based on this and the loss pattern with respect to exposure, the investigator might have felt assured that the study was free of selection bias.

Now suppose we could see the experience of the complete original cohort of 1000 men, summarized in table 2. These numbers could not be observed by the investigator but are presented here to illustrate the possible effects of the study non-participation and drop out. For the complete cohort, we see that the risk of *D* in the exposed group was increased only 69 per cent over the unexposed group. While the relative risk in table 2 was not significantly different from the relative risk in table 1, the difference in incidence rates between the two groups in table 2 was not statistically significant. Therefore, the relative risk in table 2 could have led the investigator to different material conclusions than that in table 1. Thus, the study losses caused an exaggeration of the association of disease with exposure, even though the losses appeared unbiased to the investigator.

Table 3 shows the outcome among lost subjects. As with table 2, this could not be observed by the investigator. Again, as in table 2, the difference in incidence rate between the two exposure groups (reflected in the first χ^2 value in the table) was not statistically significant, and the overall incidence rate is not significantly different from the overall rate in table 1.

TABLE 2

*Experience of the complete cohort (numbers in parentheses were not known to the investigator)**†*

	Disease *D*	Not *D*	
Exposed	(18)	(232)	250
Not Exposed	(32)	(718)	750
	(50)	(950)	1000

* Relative risk = 1.69 (χ^2 = 3.40, not significant at the .05 level).

† Five-year incidence = 50 per 1000

TABLE 3

*Experience of subjects with outcome unknown to the investigator (upper figures in the table cells are the number of original subjects lost, lower figures are the proportion of original subjects lost)**†‡*

	Disease *D*	Not *D*	
Exposed	(2)	(65)	67
	(11)	(28)	27
Not Exposed	(9)	(201)	210
	(28)	(28)	28
	(11)	(266)	277 Total lost
	(22)	(28)	277

* Relative risk = 70 (χ^2 = 23, d f = 1, not significant)

† Five-year incidence = 40 per 1000 (not significantly different from incidence in tables 1, 2 and 4)

‡ χ^2 based on expected losses in table 4 = 1 83, d f. = 1, not significant

The distortion of relative risk seen in this example arose from the fact that less than half of the proportion of exposed disease victims were lost from the study than we would have expected if the losses had been unbiased. In this situation, it is natural to ask what the chance is that loss bias as extreme as this arose out of simple random losses. Table 4 presents the expected losses under the hypothesis of random loss given the actual total loss (277 persons) seen in table 3. The second χ^2 value in table 3 is calculated based on deviations of the interior values of table 3 from the expected values in table 4, and tests whether the deviations are statistically significant. Since this χ^2 value is nonsignificant, it is quite possible that the loss pattern arose from random losses (i.e., in the absence of any bias effect).

On the other hand, the loss pattern could have been the result of a bias effect; for example, persons knowing themselves

TABLE 4

*Expected losses based on table 2 under the hypothesis of random losses (27.7% loss per cell)**†*

	Disease D	Not D	
Exposed	(5.0)	(64.2)	69.2
Not Exposed	(8.9)	(198.9)	207.8
	(13.9)	(263.1)	277.0

* Relative risk = 1.69 (not significant).

† Five-year incidence = 50 per 1000, as in table 2

to be exposed (e.g., hypertensive persons) might have more readily remained in or returned to the study when morbid events occurred than unexposed (e.g., normotensive) persons. If this hypothesis was correct, other cohort studies of D and E might be subject to the same inapparent bias and the consequent overestimation of the association of D with E.

As a second example, let us consider what would happen if all the cell and margin numbers in tables 1–4 were 100 times larger. The total cohort would then be 100,000 persons, and the relative risks in each table would be unchanged. However, the relative risks would all be significantly different from each other and from unity, and the χ^2 based on expected losses would now be 183. In this second situation, chance effect would have to be ruled out as a source of the loss pattern, and a very strong bias effect would be necessary to explain the distortion of the relative risk.

DISCUSSION

In both examples, the effect of the study losses was to exaggerate the association of the exposure with the disease, without significantly influencing the exposure prevalence or disease incidence. A case-control study drawing its samples from the observed persons would find the same exaggerated association. Examples could be given, however, in which the losses left the exposure prevalence and disease incidence unchanged but reduced the association.

When the study losses are random (in the sense of being independent of the relationships under study), the statistical probability of loss patterns such as those illustrated is strongly dependent on the sample size and the overall extent of sample loss. The mathematical explanation of why these somewhat paradoxical loss patterns are theoretically possible is that the individual distributions of several variables do not fully determine the multivariate (joint) distribution of the variables. For any frequency table (not just 2 × 2 tables), this means that the marginal totals will not fully determine the interior cell values of the table, and the marginal losses will not fully determine the interior cell losses. In particular, the type of bias illustrated above can occur despite control of confounding factors. Statistical inferences about the properties of a multivariate distribution (such as the effect of study nonparticipation and drop out) based solely on observations about the individual marginal distributions are essentially limited to the calculation of expected values under various hypotheses (e.g., the null hypothesis) about the relationship between the variables.

Study losses will tend to bias the relative risk estimate when the exposure variable is an effect-modifier for the association of study participation (i.e., both response and follow-up) with disease. Based on this observation, a selection bias effect may be defined as a joint interaction of the variables of study participation, exposure status, and disease outcome. When a selection bias effect is present in a study situation, losses tend to produce distortions with higher probability than would be expected under the hypothesis of random loss. Other studies could be subject to the same effect, thereby lessening the power of replication in the validation of study results as long as the effect continued to be unnoticed or disregarded. This distortion produced by the effect would not be ameliorated if the study was analyzed by standard methods. Missing data techniques might aid in many, though not all, situa-

tions. They could not detect and control for a bias effect not somehow apparent in the investigator's data.

Selection bias is a theoretical possibility whenever correlates of the outcome capable of influencing study participation are existent in some individuals at the beginning of the study. These correlates may be unmeasured or even unrecognized by the investigator. The question then arises, what is the actual likelihood of an important bias effect existing in a given study situation? Answering this question involves using subject matter knowledge to derive an epidemiologic judgment of the "reasonableness" of the existence of such a bias. These judgments may, at times, diverge but many experienced epidemiologists tend to regard selection bias in cohort studies as generally unimportant. Although the arguments presented here are not meant to imply that any cohort study has ever suffered from serious "Berksonian" type bias, the intended use of study results must play an important role in determining what degree of bias is to be considered significant. A bias that may have been considered insignificant in one context (such as in research establishing the existence of an association) may be considered more significant in a later context (such as in allocation of health care resources based on the strength of the association). When study losses are heavy, unbiasedness of the results cannot be verified from general principles alone.

References

1. Gordon T, Moore FE, Shurtleff D, et al. Some methodologic problems in the long-term study of cardiovascular disease: Observations on the Framingham study. J Chron Dis 10:186–206, 1959
2. Berkson J: Limitations of the application of fourfold table analysis to hospital data. Biometrics 2:47–53, 1946
3. Berkson J. The statistical study of association between smoking and lung cancer. Proc Staff Meeting, Mayo Clin 30:319–348, 1955
4. White C, Bailar JC: Retrospective and prospective methods of studying association in medicine. Am J Pub Health 46.35–44, 1956
5. Smoking and Health. Report of the Advisory Committee to the Surgeon General of the U.S Public Health Service PHS Publ. No 1103. US-DHEW, 1964, pp 96–98, 113–116.
6. *McMahon B, Pugh TF Epidemiology: Principles* and Methods. Little, Brown & Co., Boston, 1970, pp 218–220, 231