Published in final edited form as:

Epidemiology. 2009 May ; 20(3): 339–340. doi:10.1097/EDE.0b013e31819ec966.

Bias in full cohort and nested case-control studies?

Sholom Wacholder

Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD 20892,
Wacholds@mail.nih.gov

In this issue, Langholz and Richardson¹ and Hein et al.² address two recent articles by Deubner and colleagues^{3,4} about nested case-control studies. In one article,³ Deubner and colleagues called into question the fundamental validity of the nested case-control design. If their critique is compelling, it would raise doubts about the interpretation of hundreds of publications that report results from nested case-control studies. In another article,⁴ the same authors suggested a restriction on control selection in nested case-control studies to make cases and controls more comparable.

Fundamentally, a properly executed case-control study nested in a cohort is valid if the corresponding analysis of the full cohort is valid. The mathematics of the likelihoods are the same for both,⁵ as Langholz and Richardson¹ point out, and the same software procedures work for both. The only salient difference between the two designs is whether independent random samples or 100% samples are used in the conditional likelihood factor for each case. As Deubner and his colleagues³ note, the design and analysis of nested case-control studies are complex, but no more so than the analysis of cohorts, which must consider issues including time scales, various measures of time-dependent variables, and possible censoring.

Generally, the only disadvantages to nested case-control studies are the reduced precision and power due to sampling of controls, and the possibility of flaws in the sampling design or its implementation. Therefore, any problem with nested case-control studies must also be a problem for full cohort analysis. Demonstration of this (or an explanation for this discordance) would add to the credibility of a challenge to the validity of standard analytic methods for nested case-control studies.

Deubner and colleagues provide simulations³ to show bias in nested case-control studies with a lagged measure of exposure. Each step in their simulation seems reasonable. The authors started with an assembled cohort, for whom age and calendar time of starting and stopping work was available for all cohort members. In each simulation, the authors randomly assigned 142 of the cohort members to be cases, and took the end of their follow-up as the event or endpoint time. Controls matched to each case were selected from at-risk cohort members at the age of the case's event. A case's cumulative exposure was measured from time of entry into the cohort until event time. A control's cumulative exposure was measured from time of entry until the control reached the age at event of the index case to which the control was matched. Analysis was conducted by applied conditional logistic regression.

In fact, a subtle flaw in the design of these simulation studies makes them misleading. As Langholz and Richardson¹ point out, Deubner and colleagues mistakenly chose cases as a random sample of all cohort members; in fact, as in Table 1 of this paper by Hein et al.², the average age-at-event in cases is less than the average age-at-end-of-followup in comparable cohort members, when censoring is not informative and the exposure has no effect on risk of the event or censoring. Why are cases younger at the event? It is because the cases' age at end of follow-up has to be the minimum of i) the age of death from lung cancer, ii) the age of death from other causes, or iii) the age at other causes of censoring—whereas controls are followed

to the minimum of ii) or iii) only. The members of the cohort who did not die from lung cancer will be older than the cases at end of follow-up.

Further, controls' follow-up time in the simulations tends to begin at an older age than cases'. This is because, to be chosen, controls must be followed at the age of the event in the case, and (as is standard) follow-up time is truncated when the control reaches the age of the index case. Therefore, follow-up time in controls in the simulation will on average be less than follow-up time in the cases, randomly chosen from the cohort and with untruncated follow-up times. Similarly, all measures of exposure that depend on follow-up time (such as duration, average, and cumulative exposure) are distorted even when everyone receives the same level of exposure during follow-up. This phenomenon can be seen in the second row (and possibly the first row) of Table 2 of the paper by Deubner and colleagues³ where, in the absence of an exposure effect, the cumulative exposure of cases (proportional to follow-up time when exposure is constant) is greater than that of controls. In contrast, the results of the simulations (based on generating the cohort, rather than sampling from an existing cohort) show that cumulative exposure is equal between cases and controls when differences in a time-fixed level of exposure is constant (Table 1, rows 1 and 2).

So why does proportional hazards analysis truncate exposures for controls but not for cases? In proportional hazards analysis of full cohorts and nested case-control studies, the key calculation is the set of conditional probabilities that each case is the one who developed disease among all those in the cohort (or among the case and matched controls in the nested case-control study) under follow-up at case's age at event, given everyone's exposure through that age. Logically, any exposure in the case after the event cannot be related to risk at the time of event. Similarly, the other cohort members' exposures subsequent to the index case's age at event also should not be allowed to affect the conditional probability of the event.

I do not agree with Deubner and colleagues that lagging raises special concerns. A lagged measure of exposure with lag L bases risk at a given time point t only on exposure through time point $t - L$. Lagging is simply one way to measure exposure, and does not differ fundamentally from choosing other metrics such as average exposure, peak exposure, or cumulative exposure without lagging.¹ As long as exposure is measured only up to the time of the event, the particular choice of exposure summary cannot introduce bias in comparing cases and controls.¹

In their second paper, Deubner and colleagues (this time with Levy as the first author⁴) suggest the use of risk-set members' age at the end of follow-up as a control selection criterion. Specifically, they advocate choosing only controls whose age at end of follow-up is close to the index case's age at death, in order to avoid imbalance between cases and controls in age at start of followup or of first exposure and in age at censoring. Unfortunately, the use of risk set members' age at end of follow-up as a control selection criterion generates nonrandom samples. As Lubin and Gail⁶ state (and Levy et al⁴ quote), it is essential to choose a random sample from the risk set. Indeed, Hein et al.,² (Table 1, row 3) show that there is a bias generated from a non-random sample with controls who are younger at end of follow-up than the average in the risk set. The extra restriction proposed by Levy et al.⁴ can also cause another bias: if a time-independent exposure, one whose value is constant during follow-up, causes censoring due to death from another cause, the average exposure of cohort members in the risk set with follow-up even only slightly beyond the time of diagnosis of the case will tend to be less than the average in the risk set. Thus, the difference in exposure between cases and controls – and its estimated effect – will be exaggerated, even under the usual assumption of independent censoring. By contrast, the full cohort analysis will not have an analogous restriction and will be valid under independent censoring.

In my view, the two papers published in this issue^{1,2} and the arguments offered here provide a persuasive defense of the standard analytic approach for nested case control designs. Although perhaps compelling at first look, the arguments by Deubner and his colleagues^{3,4} about lagged exposure do not in fact undermine the standard analysis. The setup of their simulation contains an error, and their results are not confirmed by others. These authors offer no explanation of why the bias with lagged exposures would be restricted to nested case-control studies and not be present in the full cohort analysis. Their suggestion of nonrandom selection of controls could itself induce bias. Taking all things into account, their critique is not a valid criticism of this familiar and useful epidemiologic approach. Even so, such challenges to the *status quo* as offered by Deubner and colleagues are not without benefit – they push us to a better understanding of the fundamental principles that underlie our methods.

Acknowledgments

This research was supported in part by the Intramural Research Program of the NIH, National Cancer Institute, Division of Cancer Epidemiology and Genetics.

The author thanks Dr. Kyle Steenland, Emory University, for help in preparation of this manuscript.

REFERENCES

1. Langholz B, Richardson D. Are nested case-control studies biased? *Epidemiology* 2009 May;20(3): 321–329. [PubMed: 19289963]PMID: 19289963 [PubMed - indexed for MEDLINE]
2. Hein M, Deddens J, Schubauer-Berigan M. Bias from matching on age at death or censor in matched case-control studies. *Epidemiology* 2009;20(3)
3. Deubner DC, Roth HD, Levy PS. Empirical evaluation of complex epidemiologic study designs: workplace exposure and cancer. *J Occup Environ Med* 2007;49(9):953–959. [PubMed: 17848851]
4. Levy PS, Roth HD, Deubner DC. Exposure to beryllium and occurrence of lung cancer: a reexamination of findings from a nested case-control study. *J Occup Environ Med* 2007;49(1):96–101. [PubMed: 17215718]
5. Prentice RL, Breslow NE. Retrospective Studies and Failure Time Models. *Biometrika* 1978;65(1): 153–158.
6. Lubin JH, Gail MH. Biased selection of controls for case-control analyses of cohort studies. *Biometrics* 1984;40(1):63–75. [PubMed: 6375751]