Scand J Work Environ Health 13 (1987) 62–66

Hospital versus population referents in two case-referent studies

by Staffan E Norell, MD, PhD, Anders Ahlbom, PhD[1]

NORELL SE, AHLBOM A. Hospital versus population referents in two case-referent studies. *Scand J Work Environ Health* 13 (1987) 62–66. In case-referent studies, the choice between hospital and population referents involves the consideration of nonresponse, information bias, and selection bias. A case-referent study of astrocytoma and one of pancreatic cancer used both hospital and population referents. In the present report the two series of referents were compared with regard to potential sources of systematic errors. There were differences in reported exposure between the hospital and population referents. These differences seemed to be due to differences in exposure experience ("selection bias") rather than to differences in the quality of reported information ("information bias"). This conclusion was based on two findings. First, in the comparison between the hospital and population referents, similar differences were found in the reported exposure for "hard" questionnaire items as for "soft" questionnaire items. Second, in a comparison between reported radiographs and radiographs according to medical records, the same differences were found for the two series of referents, although there was considerable underreporting of such exposure in both series. Nonresponse rates were somewhat higher among the population referents than among the hospital referents.

Key terms: information bias, nonresponse rate, selection bias.

The cases included in a case-referent study sometimes represent all cases of the disease occurring in a population and fulfilling certain criteria with regard to age, sex, etc, of the subjects. Routines for registering certain diseases, eg, cancers and cardiovascular diseases, have increased the possibilities of locating all cases diagnosed in a population. For some of these diseases, virtually all cases of the disease will be diagnosed and perhaps also the patients hospitalized. Furthermore, hospitals may (as in Sweden) serve a geographically or otherwise defined population.

When the cases included in a study represent all cases of the disease occurring in a population during a certain period, the referents should reflect the distribution of the exposure(s) under study in the source population. This prerequisite may be achieved if referents are selected as a representative sample from that population. The alternative of selecting a hospital reference group with one or several diseases would seem less attractive from this point of view. In any particular situation it may be difficult to judge the extent to which such a group reflects the distribution of exposure in the population generating the cases. However, healthy population referents may be far from ideal from other points of view. The following two potential problems are commonly mentioned in the literature: (i) nonresponse rates may be high among population referents and (ii) healthy subjects may devote less time and effort to remembering past exposure than sick patients and thus they tend to underreport their exposure (information bias). It has been suggested that these sources of error could be reduced by the selection of hospital referents among patients with some other disease(s) (5, 6).

Although these problems have been discussed in several papers (2, 3, 4, 7, 8, 9, 10, 11, 12, 15, 16, 17, 18, 19), there are not many empirical data to illustrate the magnitude of the potential problems involved. In this paper we present data from two case-referent studies in which both hospital and population referents were used. Our purpose is to compare, within each study, the two series of referents with regard to response rates and responses to different questions about exposure. Evidence of differences with regard to selection bias, as well as information bias, is discussed.

Subjects and methods

Data were obtained from two case-referent studies, one on astrocytoma and one on pancreatic cancer (1, 13). Included were incident cases, diagnosed during the study period at certain hospital departments serving geographically defined (catchment) areas with practically no private or other alternative hospital facilities. The hospital referents were patients hospitalized for meningioma, aneurysm, or pituitary adenoma (the astrocytoma study) and inguinal hernia (the pancreatic cancer study). They were selected as stratified (age and sex) samples from patients in the same hospital departments during the same study period and were in the same age ranges as the subjects who made up the cases (20–75 and 40–79 years, respectively). In both studies, population referents were selected from the parish registers, which list all inhabitants of the parishes by date of birth. For each subject serving as a

[1] Department of Epidemiology, The National Institute of Environmental Medicine, Stockholm, Sweden.

Reprint requests to: Dr S Norell, Department of Epidemiology, The National Institute of Environmental Medicine, Box 60208, S-104 01 Stockholm, Sweden.

case, a person of the same sex with the closest date of birth was selected.

Each person diagnosed as a case and each hospital referent received a questionnaire before they were discharged from the hospital. When a case was identified, a questionnaire was mailed to the corresponding population referent. The questionnaires covered a wide range of exposures, including tobacco, alcohol, dietary and occupational factors, and housing and other environmental factors. Subjects who did not return a questionnaire were reminded to do so by letter and eventually by telephone. Those who submitted a questionnaire were contacted by telephone by an interviewer to clarify or complete specific items whenever necessary.

The hospital and population referents were compared with regard to response rate and reasons for nonresponse. The two series of referents were also compared for each item (exposure variable) in the questionnaire. For each exposure an odds ratio was calculated, comparing hospital referents with population referents, as follows;

	Notation	
	Exposed	Unexposed
Cases	a	b
Hospital referents	c	d
Population referents	c'	d'

For each item an exposure odds ratio (OR) was calculated, comparing hospital referents with population referents:

$$OR = (c \cdot d')/(d \cdot c').$$

The rate ratio estimated from the comparison with the hospital referents (RR_h) and the rate ratio estimated from the comparison with the population referents (RR_p) were estimated as:

$$RR_h = (a \cdot d)/(b \cdot c) \text{ and } RR_p = (a \cdot d')/(b \cdot c').$$

Therefore,

$$OR = RR_h/RR_p,$$

which means that OR may be interpreted as the ratio of the rate ratios estimated from the two reference series. For example, if OR = 0.5, any RR_p would be half of the RR_h, and, if OR = 2.0, any RR_p would be twice the RR_h. An OR above unity indicates that the hospital referents more often reported that they had been exposed than the population referents, and vice versa. In the calculations of the OR, the Mantel-Haenszel principle was used to adjust for differences between the reference groups with regard to age, sex, and catchment area (14). Test-based confidence limits for the OR were calculated (14).

The distribution of odds ratios for hospital versus population referents in the two studies was examined. This distribution was also examined separately for major categories of exposure. Differences in reported exposure between hospital and population referents may be due to true differences in exposure or to differences with regard to information bias.

Differences with regard to information bias are likely to increase for "soft" items, for which memory or judgment are important in answering the questions, but decrease or disappear when simple hard facts are asked for. Two independent interviewers were asked to classify each item in the two questionnaires as "soft," "intermediate," or "hard." The odds ratios for the hospital versus population referents were compared for items classified by both interviewers as "soft" and "hard," respectively. The accuracy of the information on alcohol consumption may vary considerably, and the two series of referents were compared with regard to items related to consumption of alcoholic beverages. For specific radiographic exposure it was possible to compare reported exposures during the past five years to that recorded in the medical information system of Stockholm County.

Results

Nonresponse

The total nonresponse rates were about 15 % among the referents selected from population registers. The corresponding rates were somewhat lower for the hospital referents, about 5 – 10 % (table 1).

Distribution of odds ratios

Almost 75 % of the odds ratios in the astrocytoma study were below unity. This result was due to differences between the reference groups in reported exposure to factors related to occupation, traveling, and leisure time (table 2). In the study of pancreatic cancer, the odds ratios were more often above unity, the reference groups being more similar with respect to occupational factors than with respect to other factors (table 2).

Hard versus soft items

In the questionnaire used in the astrocytoma study, there were 17 items classified by both interviewers as "hard" and 31 items classified as "soft." The odds ratios for the hospital versus population referents were below unity for 15 and 25 of these items, respectively. In the questionnaire used in the study of pancreatic cancer, 16 items were classified as "hard" and 51 as "soft." The odds ratios were below unity for 8 and 25 of these items, respectively.

Alcoholic beverages

Different aspects of alcohol consumption, as reported by the hospital and population referents, were examined. The results shown in table 3, with odds ratios unusally close to unity, indicate that there were no systematic differences between the two groups in reported consumption of alcoholic beverages.

Table 1. Nonresponse among the hospital and population referents in the two case-referent studies.

	Astrocytoma study		Pancreatic cancer study	
	Hospital referents (N = 207) (%)	Population referents (N = 111) (%)	Hospital referents (N = 179) (%)	Population referents (N = 162) (%)
Nonrespondents				
Refused	3.4	14.4	7.8	12.3
Dead	1.4	0	0.6	0
Not located	0	2.7	0.6	2.5
Total	4.8	17.1	9.0	14.8
Respondents to questionnaire[a]	95.2	82.9	91.0	85.2
Total identified	100	100	100	100

[a] Includes questionnaire filled out by spouse for 2.4 % of the hospital referents and 1.0 % of the population referents in the astrocytoma study and for 0.9 % of the hospital referents and 0.6 % of the population referents in the pancreas cancer study.

Table 2. Distribution of the odds ratios (OR) for the hospital versus population referents by category of exposure.

	OR < 1.0[a] (N)	OR ≥ 1.0 (N)
Astrocytoma study		
Occupational factors	39	3
Traveling, leisure time	21	3
Other (including diet, alcohol, and tobacco)	31	30
Pancreatic cancer study		
Occupational factors	20	22
Traveling, leisure time	11	16
Other (including diet, alcohol, and tobacco)	36	65
Total	67	103

[a] The population referents gave the highest exposure rate.

Table 3. Odds ratios (OR)[a] and 90 % confidence limits (90 % CL) for the hospital versus population referents according to reported consumption of alcoholic beverages.

Alcoholic beverage	Astrocytoma study		Pancreatic cancer study	
	OR	90% CL	OR	90% CL
Beer	1.0	0.5–2.0	1.4	0.7–2.8
Wine,				
Frequency	0.4	0.2–0.8	0.9	0.5–1.5
Volume	1.0	0.6–1.6	1.1	0.6–1.8
Spirits				
Frequency	1.1	0.6–2.0	1.3	0.8–2.1
Volume	1.0	0.6–1.8	0.7	0.4–1.2

[a] An OR of < 1 means that the population referents gave the highest exposure rate.

Reported versus recorded exposure

Table 4 shows specific radiographic exposure during the past five years as reported on the questionnaire in relation to the records from the medical information system. The odds ratios for the hospital versus population referents were calculated on the assumption that (i) only questionnaire data had been available, (ii) only those with recorded exposure had been exposed, and (iii) subjects with reported or recorded exposure had been exposed. The crude odds ratios were 0.5, 0.6, and 0.6, respectively, for radiographs in the A category (upper limb, including shoulder), and 1.1, 1.0 and 1.2, respectively, for radiographs in the B category (lower limb, including hip).

Table 4. Specific radiographic exposure (category A = upper limb, including shoulder; category B = lower limb, including hip) during the past five years, as reported on the questionnaire and as recorded in the medical information system of Stockholm County (pancreatic cancer study).

Medical information system	Questionnaire			
	Hospital referents		Population referents	
	Exposed	Unexposed	Exposed	Unexposed
Category A				
Exposed (N)	10	8	17	8
Unexposed (N)	6	111	6	84
Category B				
Exposed (N)	11	13	13	8
Unexposed (N)	8	103	2	92

Discussion

There was evidence of systematic differences between the rate ratio estimated from the hospital referents and that estimated from the population referents, particularly in the astrocytoma study. For exposure related to occupation, traveling, and leisure time no less than 60 out of 66 odds ratios were below unity (table 2). The direction of the difference shows that the population referents more often reported that they had been exposed than the hospital referents. This direction is the opposite of that expected from a difference with regard to information bias. (See the Introduction.) The difference was only evident for items related to occupation (39 of 42, OR < 1) and traveling and leisure time (21 of 24, OR < 1) and not for other items, including diet (31 of 61, OR < 1), something which is hard to explain on the basis of information bias. Furthermore, the difference between the reference groups was no less

for hard items (15 of 17, OR < 1), for which information bias is less likely to have any influence, than for soft items (25 of 31, OR < 1). With these observations being taken into consideration, the difference between the reference groups in the astrocytoma study is not likely to have been due to differences with regard to information bias. However, patients with certain benign diseases (our hospital referents) may have been less active in occupations, traveling, and leisure time than healthy population referents; or subjects who are less active in occupations, traveling, etc, may be more likely to be hospitalized. Each of these possibilities (or both) could explain the difference in exposure experience between the hospital and the population referents in the observed direction.

In the pancreatic cancer study the odds ratios were somewhat more often above than below unity. This result was largely due to differences with regard to dietary factors (table 2). The hospital referents reported, more often than the population referents, a high consumption of several food items related to the intake of fat and calories (which may be related to overweight and perhaps also to inguinal hernia, the disease of the hospital referents). The proportion of odds ratios that was below unity for hard (8 of 16) and soft (25 of 51) items was similar and did not suggest a difference with regard to information bias.

Several studies have shown that alcohol consumption is likely to be underestimated in interviews and questionnaires. A low accuracy of exposure information would leave considerable space for differences with regard to information bias between groups. However, the comparisons between the hospital and population referents indicated in both studies that there were no systematic differences in the reported consumption of alcoholic beverages. Although the accuracy of exposure information may be low, there was no evidence of differences with regard to information bias between the hospital and population referents.

All radiographic examinations carried out in Stockholm County are recorded in a computerized medical information system. Exceptions are the relatively small proportion of all such examinations carried out in private practice (almost exclusively outpatients). When subjects living in the County have a radiographic examination while traveling outside the County, the examination is not recorded in the information system. In the study of pancreatic cancer, information on specific radiographic examinations recorded in the medical information system was obtained for all the hospital and population referents who had lived in the County for at least five years before the study. The record for the last five years before a subject filled out the questionnaire was compared to the information on specific radiographic examinations during the past five years given in the questionnaire. Subjects with recorded exposure would often have been classified as unexposed on the basis of their questionnaire responses. Although such false negatives were somewhat more common among the hospital than among the population referents, there was no great difference between the two groups in this respect. The odds ratios estimated from the questionnaire data did not differ much from the odds ratios calculated on the assumption that only subjects with recorded exposure had been exposed or that subjects with reported or recorded exposure had been exposed. For the exposure for which the two reference groups differed most, the crude odds ratio based on the questionnaire data was 0.5 versus 0.6 when the data from the medical information system were taken into account. This finding suggests that the difference between the reference groups was due to differences in exposure experience and not to differences with regard to information bias.

In the two case-referent studies, the cases were selected in such a way as to represent all cases in a geographically defined population. Referents selected as a representative sample of that population should minimize selection bias, but perhaps not nonresponse rates and information bias. Comparisons between the hospital and population referents in the two studies indicated that there were differences in reported exposure between the two groups, particularly in the astrocytoma study. There was, however, no evidence of important differences with regard to information bias between these two groups. The differences in reported exposure between the two groups were more likely to reflect true differences in exposure. Nonresponse rates were somewhat higher among the population referents than among the hospital referents.

Acknowledgments

This study was supported by grants from the Swedish Medical Research Council (project no 6616).

We are indebted to our collaborators in each of the two case-referent studies, in particular Ms I Lindberg-Navier, Department of Social Medicine, Huddinge University Hospital, and Mr R Olin, Preventive Unit for Occupational Medicine, Royal Institute of Technology.

References

1. Ahlbom A, Lindberg-Navier I, Norell S, Olin R, Spännare B. Non-occupational risk indicators for astrocytoma in adults. Am J Epidemiol 124 (1986) 334–337.
2. Axelson O. The "case-control" study: Valid selection of subjects. J Chronic Dis 38 (1985) 553–555.
3. Clarke EA, Stavraky KM. Response. J Chronic Dis 37 (1984) 339–340.
4. Cole P. The evolving case-control study. J Chronic Dis 32 (1979) 15–27.
5. Cole P. Introduction. In: Breslow NE, Day NE, ed. Statistical methods in cancer research. Volume 1 (The analysis of case-control studies). International Agency for Research on Cancer, Lyon 1980. (IARC scientific publication no 32).
6. Cornfield J, Haenszel W. Some aspects of retrospective studies. J Chronic Dis 11 (1960) 523–534.
7. Feinstein AR. Experimental requirements and scientif-

ic principles in case-control studies. J Chronic Dis 38 (1985) 127–133. (Editorial).
8. Feinstein AR. The case-control study: Valid selection of subjects. J Chronic Dis 38 (1985) 551–552.
9. Ibrahim MA, Spitzer WO. The case control study: The problem and the prospect. J Chronic Dis 32 (1979) 139–144.
10. Miettinen OS. The "case-control" study: Valid selection of subjects. J Chronic Dis 38 (1985) 549–550.
11. Miettinen OS. Author's response. J Chronic Dis 38 (1985) 557–558.
12. Miller AB. Hospital or population controls, an unanswered question. J Chronic Dis 37 (1984) 337–338.
13. Norell S, Ahlbom A, Erwald R, Jacobson G, Lindber-Navier I, Olin R, Törnberg B, Wiechel K-L. Diet and pancreatic cancer: A case-control study. Am J Epidemiol 124 (1986) 894–902.
14. Rothman KJ, Boice JD. Epidemiologic analysis with a programmable calculator. US Government Printing Office, Washington, DC 1979. (DHEW publication no (NIH) 79-1649).
15. Sackett DL. Bias in analytic research. J Chronic Dis 32 (1979) 51–63.
16. Schlesselman JJ. Valid selection of subjects in case-control studies. J Chronic Dis 38 (1985) 549–550.
17. Shuster JJ, Cook B. Hospital or population controls: A discussion. J Chronic Dis 36 (1983) 315–316.
18. Spitzer WO. Ideas and words: Two dimensions for debates on case controlling. J Chronic Dis 38 (1985) 541–542.
19. Stavraky KM, Clarke EA. Hospital or population controls? An unanswered question. J Chronic Dis 36 (1983) 301–307.

Received for publication: 26 May 1986