Influence of design characteristics on the outcome of retrospective cohort studies

G M H SWAEN, J M M MEIJERS

From the Department of Occupational Medicine, University of Limburg, PO Box 616, 6200 MD Maastricht, The Netherlands

ABSTRACT Retrospective cohort studies are increasingly being applied in occupational health. To describe and investigate further this type of study 179 retrospective cohort studies published in six scientific journals between 1975 and 1985 inclusive were reviewed. A description of the 179 reviewed articles was made and relations between investigator orientated variables, design characteristics, and the outcome of the study were investigated. Retrospective cohort studies focusing on exposures in the chemical industry appeared to yield most negative findings, which is partly explained by the relation between the affiliation of the investigator and the outcome of the study. Studies requiring a minimal latency period, an occupational reference group, and a low percentage of lost to follow up tended to have a higher chance of a positive finding. Study size, however, did not appear to be related to the outcome.

Epidemiological studies have become one of the generally recognised tools for investigating long term health effects of the occupational environment. This is particularly true for retrospective cohort studies, or non-concurrent prospective studies. A retrospective cohort study may be defined as an epidemiological study in which a group of people has been identified who have all experienced a particular event (occupational exposure) in the past and are followed up to determine the occurrence of disease or death. From a reference population an expected number of deaths for specific diagnoses may be calculated by means of indirect standardisation. The observed number of deaths, divided by the expected number, multiplied with 100 gives a standardised mortality ratio (SMR) that may be regarded as a measure of the relative risk for the exposed group to die of a particular disease compared with the reference population. Retrospective cohort studies are a powerful tool for detecting long term health effects of particular exposures, despite their observational character.¹ They are expensive to carry out and usually take a long time to complete.² They are particularly effective in cases of a specific exposure that may play an aetiological part in various diseases if the problem to be dealt with is of the type: Does exposure to substance A lead to a higher risk of cancer and if so, what type of

cancer?

Retrospective cohort studies are observational studies and have been subject to much criticism because they have no experimental design.³⁻⁶ Usually, incorrect reference populations are applied, resulting in a healthy worker effect.⁷ The classic statistical procedure, the SMR analysis, has also been subject to elaborate criticism.⁸ There is no alternative yet to retrospective cohort studies and therefore priority should be given to improving the existing methods and the policies by which these studies are conducted and the results published. It is essential to take sufficient care in collecting and interpreting the data to eliminate the possible effects of bias, confounding and chance that may lead to contradictory outcomes.^{8,9} Furthermore, the design characteristics—for example, number of people at risk and exposure definition—may influence the results of cohort studies to a great extent.

Methodological issues in epidemiological studies may be investigated in several ways. One can take theoretical considerations as a basis and deduct the empirical effects of design characteristics.^{4,10} Wang and Miettinen, for instance, have argued on theoretical grounds that applying the death rates of the general population as the reference in retrospective cohort studies is erroneous.⁴ Another example of this approach is the statement that in retrospective cohort studies the healthy worker effect will fade out after a longer period of employment or a longer period of follow up. After a longer period of employment the

initially healthy population has had enough time to develop diseases that have not been picked up at the medical examination at first employment. This theoretical approach, however, misses the potential actually to quantify the effects of design characteristics on the outcome of a study or provide evidence that such effects occur in reality.

A second approach to investigate the effects of design characteristics on the outcome of a study is by simulation. An actual conducted study may be re-analysed after artificial modification of a design characteristic. For instance, one can take a study with a long follow up in which no healthy worker effect was observed and artificially shorten the period of follow up. If, after decreasing the duration of follow up, a healthy worker effect is observed evidence is presented that a healthy worker effect occurs more frequently in retrospective cohort studies with a short period of follow up. Such an approach, for instance, was taken by Koskela *et al.*¹¹ By analysing the data of a large cohort study in various ways the investigators showed that in that particular study the cohort formation criteria had an effect on the outcome of the study. McMichael also used simulation to show how the healthy worker effect varied within different sub-groups and with the use of different design characteristics in retrospective cohort studies.¹²

A third approach is taken in this review in which 179 articles of retrospective cohort studies about occupational exposures in an 11 year period have been abstracted and analysed. Each article has been regarded as an independent entity, comprising information about the effects of investigator orientated variables (country, authors' affiliation, and financiers, for example) and design characteristics (including expected deaths, exposure definition, duration of follow up) on the outcome of the retrospective cohort study. The investigator orientated variables and design characteristics were considered as the independent variables in this review. The dependent variable consisted of the outcome of the study: a negative study was defined as a study in which the investigators concluded that it did not give enough evidence for the existence of an increased risk of dying of a particular disease. "Non-conclusive studies" were also categorised as negative studies. We describe the research methods, the investigator orientated variables, and design characteristics and report whether or not these independent variables are related to the outcome of the studies.

Certainly the present review does not cover all occupational retrospective cohort studies reported during 1975–85. To estimate the coverage we looked at the references given in the review article on occupation and cancer by Decoufle¹³ and noted that 66% of the retrospective cohort studies cited in that article were published in journals covered by this survey.

Material and methods

The following six scientific journals were searched from 1975 to 1985 inclusive.

American Journal of Epidemiology
American Journal of Public Health
British Journal of Industrial Medicine
International Journal of Epidemiology
Journal of Occupational Medicine
Scandinavian Journal of Work Environment and Health

Every article reporting on occupational retrospective cohort study was abstracted. If the outcome was incidence of cancer registered in a cancer registry the study was included on the condition that comparable follow up and analytical procedures were used as in mortality studies. In these instances the standardised incidence ratio (SIR) of the total number of cancers was taken as SMR of the total number of deaths. If more than one article was written reporting a particular study only one article was included in the review. Updates, however, were regarded as separate studies.

In appendix 1 all the variables collected in this review are displayed, with the categories that were used. All the articles were abstracted by one of us after specific agreements were made about how the items should be classified. The following guidelines were applied in the abstraction work: for the investigator orientated variables the first author of the publication was decisive. If no mention was made concerning the financial sources available to the retrospective cohort study it was assumed to be similar to the affiliation of the first author. For each study minimal and maximal latency were calculated using the cohort definition and the end date of the follow up. For instance, if all the workers employed between 1945 and 1960 were followed up until 1986 the minimal latency was 26 years and the maximal latency was 41 years. If all exposed workers employed after 1945 were followed up until 1986 the minimal latency was 0 years (since a person who started working on 30 December 1985 was still eligible for follow up) and the maximal latency was 41 years. Altogether, 179 retrospective cohort studies were abstracted.

The data collected in this review have been analysed by statistical procedures generally applied in case-control studies. Odds ratios and two sided 95% confidence limits were calculated, as described by Schlesselman.¹⁴

Results

A total of 179 retrospective cohort studies was found from 1975 to 1985 inclusive; 12 were "updates." The *British Journal of Industrial Medicine* and *Journal of*

Occupational Medicine combined published 74% of the total number of occupational cohort studies in the given period, followed by the *Scandinavian Journal of Work Environment and Health* (14%). Typical epidemiological periodicals play no substantial part in the publication of occupational cohort studies. In the *American Journal of Public Health* not one study eligible for this review was published. From 1975 to 1985 there was a consistent increase in the number of cohort studies published: in 1975 only five articles were found in the journals under investigation compared with 34 in 1985.

Of the studies, 45% were carried out in the United States, 21% in the United Kingdom, and 19% in Scandinavia; the remainder were conducted in a wide variety of countries in and outside Europe.

INVESTIGATOR ORIENTATED VARIABLES

Several variables regarding the investigator were studied in relation to the outcome of the retrospective cohort studies.

The cohort studies performed by epidemiologists employed by governmental agencies had the highest percentage of positive results. Of the studies executed by them, 76% were positive compared with only 36% of the cohort studies performed by epidemiologists employed by industry (appendix 2). Similar differences were observed regarding the financial resources available to the cohort studies. About 77% of the studies funded by governmental agencies or universities were positive compared with 53% of the studies financed by industry (appendix 2). The outcome of the study also varied by the type of occupational exposure in the industry that was under investigation (appendix 2). The only exposures in which more negative than positive outcomes were reported were those in the chemical industry. Therefore the relation between the affiliation of the principal investigator and the outcome of the cohort studies may be explained by the fact that epidemiologists employed by industry work mainly in the chemical industry, an area in which less positive results seem to be expected. The analysis was repeated after omission of all the studies conducted by epidemiologists affiliated with industry and private agencies. The tendency that proportionally most negative study outcomes were reported in the chemical industry was still noted. In fact there was no significant change in the order of substances under investigation in respect of the proportion of positive studies after the omission of studies conducted by epidemiologists employed by industry or private agencies. Nevertheless, if the analysis was restricted to studies that focused on occupational exposures in the chemical industry the difference in outcome between studies conducted by epidemiologists from industry itself and other epidemiologists was apparent. Only 32% of the

studies in the chemical industry performed by epidemiologists employed by industry were positive, compared with 59% conducted by other epidemiologists. Although the number of studies is small, this difference is statistically significant ($p < 0.05$).

Another investigator orientated variable that appeared to be related to the outcome of the cohort studies was the country in which the study was conducted (appendix 2). Restricting the analysis to the three geographical areas that contributed most retrospective cohort studies (United States, United Kingdom and Scandinavia), the highest proportion of positive studies was reported from Scandinavia (82%) and the lowest (52%) from the United States. This difference was statistically significant ($p < 0.005$). The proportion of positive outcomes of Scandinavian studies was also significantly higher than in all other studies combined ($p < 0.005$). This difference remained after omitting cohort studies conducted in the chemical industry because of their unequal distribution over the countries and their subsequent potential confounding effect.

DESIGN CHARACTERISTICS

There appeared to be no relation between the size of the exposed cohort and the study outcome (appendix 2). Of the studies with less than 5000 workers, 65% were positive, compared with 66% of the studies with 5000 or more workers. Only a weak relation was found between the total number of expected deaths, the total person-years of follow up, and the outcome of the study (appendix 2).

We investigated if the application of a minimal exposure period as an eligibility criterion affected the outcome of the retrospective cohort studies. No significant difference was observed in respect of the outcome between studies that had no minimal exposure period and studies that did (appendix 2). Among the studies in which the exposed cohort was restricted to workers who at least had experienced one year of exposure, however, there were more positive outcomes than among the studies in which no minimal latency was set. The odds ratio of a positive outcome in the first type was 1.9 higher than in the studies with no minimal latency period (95% confidence limits: 0.99–2.85).

It has often been argued that comparing occupational cohorts with the mortality rates of the general population is inappropriate^{14,15} since occupational cohorts differ from the general population in such matters as education, socioeconomic status, and health. Therefore we investigated the outcomes of retrospective cohort studies in which a specific, working comparison group was applied and cohort studies in which this was not done. In only 26 studies was a specific occupational comparison group

used. These studies did appear to have more positive outcomes (77%) than the studies with no specific comparison group (63%). The odds ratio of a positive outcome was 1.9 greater in retrospective cohort studies in which a specific comparison group was used, as compared with studies where this was not done (95% CL: 0.97–2.89).

One of the concerns regarding the interpretation of results of retrospective cohort studies has always been if an in what way cohort members lost to follow up influence the outcome of the study. To investigate this, we divided the studies reporting the percentage of lost to follow up into studies with less than 1% of the cohort lost to follow up and studies with 1% or more lost to follow up. In studies with less than 1% lost to follow up the odds ratio of a positive outcome was 2.66 times as high as in the studies with 1% or more of the cohort lost to follow up (95% CL: 1.96–3.36).

Only a weak, non-significant association was found between duration of follow up and the outcome of the study. Duration of follow up was defined as the time between the first and last year of follow up. In studies having 30 years or more of follow up the odds ratio of a positive outcome was 1.3 times as high as in the retrospective cohort studies with less than 30 years of follow up (95% CL: 0.69–1.93).

Discussion

Papers that consider the influence of methodological issues on the outcome of retrospective cohort studies usually consist of theoretical considerations or simulation examples. Here the relations between the investigator's background, design characteristics, and study outcome were analysed for 179 occupational retrospective cohort studies on the basis of structured abstracts. The outcome was defined in terms of positive or negative outcome in the eyes of the investigators. There are pitfalls, however. When correlations are found between the distribution of independent and dependent variables these may be influenced by other, confounding variables—for example, the exposure under study. It must also be kept in mind that the conclusions are drawn from a non-experimental design and therefore we cannot study whether or not a relation is a causal one.

In the present review retrospective cohort studies conducted by epidemiologists employed by industry were found to have proportionally more negative results than those conducted by governmental agencies or universities (65% versus respectively 24% and 32%). Part of the difference could be explained by the fact that industrial research workers mainly investigated exposures in the chemical industry and this may a priori lead to more negative findings. When considering only exposures in the chemical industry,

however, a tendency remained for industrially employed epidemiologists to conduct proportionally more studies with a negative outcome. One possible explanation for this finding may be that those industries that employ epidemiologists are also more aware of the potential risks involved and have put great effort into improving the occupational environment. Another hypothesis may be that the background of the investigators influences their interpretation of the results. Epidemiologists employed and financed by industries have a more dependent relation to their study objects than epidemiologists working at universities or governments. Perhaps this could be reflected by a different interpretation of health risks.

Scandinavian studies reported significantly more positive findings (82%) than studies from other countries; studies from the United States significantly less (52%), even after omitting exposures to chemicals in the chemical industry because of their confounding potentials. It follows that next to the affiliation and financial sources of the authors, the country also seems to be related to the study results. Different exposure levels at the workplace due to various preventive measures and other production methods may perhaps explain the geographical variation. It is also possible that political and economic differences shape the interpretation of the outcomes.

No obvious relation was found between design characteristics concerning the size of cohort studies and a positive or negative outcome. The total number of exposed workers, expected deaths, and total person-years of follow up did not seem to influence the results. This was a somewhat surprising result and it may indicate that power calculations should not be the sole basis for deciding whether or not it is worth while to pursue a particular study. Some other design characteristics appeared to be related to the outcome of the study.

In occupational cohort studies it is desirable to determine the vital status at the end of follow up for as many individuals under investigation as possible. Usually this is not entirely feasible. Therefore in almost every study the lost to follow up problem has to be dealt with. Interpretation of the 179 articles showed that lost to follow up percentages were related to outcome. Studies with a percentage lost to follow up lower than one had an odds ratio of a positive outcome 2.7 times as high as studies with 1% or more lost to follow up. It is not obvious that these differences are caused by the vital status of those lost to follow up; it is more plausible that the quality of a study is expressed by the percentage lost to follow up. But then again it is not clear how the relation between this variable and the outcomes must be understood.

With this survey we hope we have contributed to a better understanding of the methodological issues

regarding retrospective cohort studies in occupational epidemiology and a more objective assessment of the relevance of their results.

Requests for reprints to: Dr Swaen.

References

- 1 Friedman GD. *Primer of epidemiology*. New York: McGraw-Hill Book Company, 1980:122.
- 2 Lilienfeld AM, Lilienfeld DE. *Foundations of epidemiology*. New York: Oxford University Press, 1980:247.
- 3 Enterline PE. Pitfalls in epidemiological research, and examination of the asbestos literature. *J Occup Med* 1976;**18**:150–6.
- 4 Wang J, Miettinen O. Occupational mortality studies. *Scand J Work Environ Health* 1982;**8**:153–8.
- 5 Bell CMJ, Coleman DA. Predicted mortality patterns in cohort study populations exposed to different types of hazard: can SMR's show a dose-response? *Statistics in Medicine* 1983;**2**:363–71.
- 6 Gaffey WR. A critique of the standardized mortality ratio. *J Occup Med* 1976;**18**:157–60.
- 7 Shindell S, Weisberg RF, Giefer EE. The "healthy worker effect"—fact or artifact. *J Occup Med* 1978;**20**:807–11.
- 8 Wong O, Decoufle P. Methodological issues involving the standardized mortality ratio and proportionate mortality ratio in occupational studies. *J Occup Med* 1982;**24**:299–304.
- 9 Doll R. Occupational cancer: problem in interpreting human evidence. *Ann Occup Hyg* 1984;**28**:291–305.
- 10 Hernberg S. "Negative" results in cohort studies—how to recognize fallacies. *Scand J Work Environ Health* 1981;**7**, suppl 4: 121–6.
- 11 Koskela R, Järvinen E, Kolari P. Effect of cohort definition and follow-up length on occupational mortality rates. *Scand J Work Environ Health* 1984;**10**:311–6.
- 12 McMichael AJ. Standardized mortality ratios and the healthy worker effect: scratching beneath the surface. *J Occup Med* 1976;**18**:165–8.
- 13 Decoufle P. Occupation. In: Schottenfeld D, Fraumeni JF, ed. *Cancer epidemiology and prevention*. London: Saunders Company, 1982:318–35.
- 14 Schlesselman JJ. *Case-control studies, design conduct analysis*. Oxford: Oxford University Press, 1982.
- 15 Fox AJ, Collier PF. Low mortality rates in industrial cohort studies due to selection for work and survival in the industry. *Br J Prev Soc Med* 1976;**30**:225–30.

Appendix 1

PRESENTATION OF THE VARIABLES OF WHICH INFORMATION WAS COLLECTED FOR THE 179 RETROSPECTIVE COHORT STUDIES INCLUDED IN THE SURVEY

Dependent variable:

Positive or negative study

Independent variables, investigator orientated:

Year of publication (from 1975 to 1985 inclusive)
 Affiliation (government, university, industry, other)
 Resources (government, university, industry, other)
 Country (United States, Canada, United Kingdom, Scandinavia, rest of Europe, rest of world)

Independent variables, methodologically orientated:

Number of exposed workers (numerical)
 Definition of exposure (total industry, particular workplace, particular job, both)
 Minimal exposure period (none, < half year, < one year, longer)
 Reference group (yes, no), size if applicable
 Number of person-years of follow up (numerical)
 Minimal and maximal observed latency
 Adjustment (smoking, other variables)
 First and last year of follow up
 Number of deaths expected in the exposed group
 Percentage lost to follow up
 Number of years between end date of follow up and year in which the article was published

Appendix 2 on next page

Appendix 2

Cross tabulations of investigator orientated variables, design characteristics, and the outcomes of 179 retrospective occupational cohort studies published between 1975 and 1985.

	Positive		Negative			Positive		Negative	
	(%)	No	(%)	No		(%)	No	(%)	No
Affiliation:					No of workers:				
University	(68.4)	52	(31.6)	24	0- 999	(64.5)	40	(35.5)	22
Government	(76.1)	51	(23.9)	16	1000- 4999	(65.2)	43	(34.8)	23
Industry	(35.5)	11	(64.5)	20	5000- 9999	(73.9)	17	(26.1)	6
Private	(60.0)	3	(40.0)	2	> 10000	(53.3)	16	(46.7)	11
Total	(65.4)	117	(34.6)	62	Total	(65.4)	116	(34.6)	62
Financial resources:					Minimal exposure period:				
Government	(76.7)	56	(23.3)	17	None	(66.3)	67	(33.7)	34
Industry	(52.8)	38	(47.2)	34	< half year	(57.9)	11	(42.1)	8
University	(77.8)	21	(22.2)	6	Half year—year	(55.2)	16	(44.8)	13
Private	(28.6)	2	(71.4)	5	> year	(78.6)	22	(21.4)	6
Total	(65.4)	117	(34.6)	62	Total	(65.5)	116	(34.5)	61
Exposure under investigation:					No of expected deaths:				
Chemicals	(48.6)	36	(51.4)	38	< 100	(59.6)	28	(40.4)	19
Metals	(83.3)	30	(16.4)	6	100-499	(65.6)	42	(34.4)	22
Asbestos	(76.2)	16	(23.8)	5	500-999	(64.0)	16	(36.0)	9
Dust	(66.7)	12	(33.3)	6	> 1000	(71.4)	30	(28.6)	12
Radioactivity	(62.5)	5	(37.5)	3	Total	(65.5)	116	(34.5)	61
Mining	(85.7)	6	(14.3)	1	Percentage lost to follow up:				
Others	(80.0)	12	(20.0)	3	< 1	(77.9)	53	(22.1)	15
Total	(65.4)	117	(34.6)	62	1- 4.9	(56.7)	34	(43.3)	26
Country of investigation:					5- 9.9	(50.0)	16	(50.0)	16
United States	(51.9)	42	(48.1)	39	> 10	(87.5)	7	(12.5)	1
United Kingdom	(65.8)	25	(34.2)	13	Total	(65.5)	110	(34.5)	58
Scandinavia	(82.4)	28	(17.6)	6	Duration of follow up (years):				
Canada	(90.9)	10	(9.1)	1	0- 9	(83.3)	10	(16.7)	2
Rest of Europe	(80.0)	8	(20.0)	2	10-19	(60.0)	18	(40.0)	12
Others	(80.0)	4	(20.0)	1	20-29	(59.6)	34	(40.4)	23
Total	(65.4)	117	(34.6)	62	30-39	(66.0)	35	(34.0)	18
Journal:					40-49	(78.9)	15	(21.1)	4
<i>J Occup Med</i>	(53.7)	36	(46.3)	31	> 50	(62.5)	5	(37.5)	3
<i>Br J Ind Med</i>	(61.5)	40	(38.5)	25	Total	(65.5)	117	(34.5)	62
<i>Scand J Work Environ Health</i>	(96.0)	24	(4.0)	1					
<i>Am J Epidemiol</i>	(64.3)	9	(35.7)	5					
<i>Int J Epidemiol</i>	(100.0)	8	(0)	0					
<i>Am J Public Health</i>	(0)	0	(0)	0					
Total	(65.4)	117	(34.6)	62					