

# NEGATIVE AND NON-POSITIVE EPIDEMIOLOGICAL STUDIES

**OLAV AXELSON**

Division of Occupational and Environmental Medicine  
Department of Molecular and Clinical Medicine  
Faculty of Health Sciences  
Linköping University  
Linköping, Sweden

**Abstract.** The aim of this study was to identify and discuss validity aspects on so called negative and non-positive studies. Arguments and examples are drawn from experiences in occupational health epidemiology regarding the interpretation of more or less equivocal study results. A negative study may be defined as showing a result that goes against the investigated hypothesis of an increased (or prevented) risk. Traditionally, studies with a risk estimate (relative risk or odds ratio) above, but close to unity are also referred to as negative, given a narrow confidence interval (CI) that includes unity. A risk estimate above unity with the CI including unity is non-positive, however, but an estimate below unity with upper CI bond exceeding unity might be seen as possibly negative or non-negative. A weaker “significance” than usually required should perhaps be accepted when evaluating serious hazards. In contrast to positive studies, the negative and non-positive studies tend to escape criticism in spite of questionable validity that may have obscured existing risks (or preventive effects). Even stronger arguments can be made in criticising negative and non-positive studies than positive studies, for example, regarding selection phenomena, and observational problems regarding exposure and outcome. Negative confounding should be considered although usually weak. In case-control studies, so called over-matching may obscure an existing risk as could the “healthy worker effect” in cohort studies. Small scale non-positive studies should be made available for meta-analyses and when considering studies that do not convincingly show a risk; those who are exposed should be given the “benefit of the doubt”.

**Key words:**

**Case-control, Cohort, Confounding, Exposure, Matching**

## INTRODUCTION

When epidemiologic study results are weak or inconsistent between studies, it is indeed difficult to obtain a tenable judgement on whether there is a health risk from some particular exposure at consideration. Usually, studies are preceded by some suspicion that an exposure may be hazardous (or alternatively, beneficial) to human health, with the subsequent likelihood that the study result also indicates the presence of some, but not necessarily any convincing excess risk (or prevention). Sometimes there is

also an interest to rule out a risk from a certain exposure. This is a quite demanding challenge, however, requiring a large scale study that indicates no excess risk, i.e., a so called negative study, the nature of which may be further considered here. Various aspects of the interpretation of equivocal study results have been discussed from time to time in the literature [1–3] and will be elucidated in the following, particularly with regard to experiences from occupational health epidemiology. Many of the viewpoints will have general applicability in epidemiologic research, however.

Received: January 19, 2004. Accepted: January 30, 2004.

Address reprint requests to Prof. emeritus O. Axelson, Division of Occupational and Environmental Medicine, Department of Molecular and Clinical Medicine, Faculty of Health Sciences, Linköping University, 581 85 Linköping, Sweden (e-mail: olav.axelson@imk.liu.se).

### SOME STATISTICAL ASPECTS

A truly negative study may be defined as showing a result that significantly goes against a hypothesis of risk (or prevention). Traditionally, however, studies with a risk estimate (usually taken as relative risk or odds ratio) above, but close to unity (i.e., the reference risk), are also referred to as negative, given a narrow confidence interval (CI) that includes unity. From a formal point of view, however, a risk estimate above unity with the CI including unity should be taken as non-positive. It could perhaps be proposed that even an estimate below unity with the upper CI bond exceeding unity might be seen as a possibly negative or non-negative result (the reversed view applies to preventive situations although not further discussed here).

Usually, 95% confidence intervals around the risk estimate are applied in epidemiologic studies and when excluding unity, the result is referred to as "significant". Sometimes there may even be considerations whether to use a 99% CI to reduce so called false positive results. However, there could be reasons for accepting a weaker "significance" than usually required, especially when evaluating serious risks, e.g., a 90% CI (or even less?) as corresponding to a one-tailed p-value of 0.05. Hence, when considering risks, the one-sided confidence interval should be of primary interest, at least in the early phase of research for assessing whether a health risk may be present or not. In later investigations of the effect of an adverse exposure, the upper confidence bond may also attract interest as suggesting a likely upper limit for the risk.

### VALIDITY CONSIDERATIONS

In contrast to positive studies with both the risk estimate and the lower confidence bond exceeding unity, the negative and non-positive studies tend to escape criticism. The general attitude seems to be that if no definite, or "significant", risk is shown, there is no reason to believe that the exposure at issue is hazardous. However, poor design and questionable validity of a study may obscure existing risks, and even stronger arguments can be made in criticising negative and non-positive studies than positive studies.

The reason is simply that a poorly designed and conducted study would not have sufficient discrimination capacity to reveal an existing risk. This view is not inconsistent with the fact that a biased study can also result in a false indication of an increased risk and not only in a negative or non-positive result.

Considering non-positive or negative study results, the possibility of some selection should be considered, especially a loss of exposed cases with subsequent reduction of the risk estimate. For example, deceased or disabled individuals may have been sorted out of such company files that could be used for setting up an occupational cohort study. Similarly, observational imprecision regarding exposure and outcome through non-differential misclassification would also have a reducing effect on the risk estimate as would uncontrolled negative confounding. It may be remarked also that improper measurement or non-differential misclassification of a confounder leads to poor control of confounding, either it is negative or positive [4,5].

### Assessing exposure

Regarding assessment of exposure there are concerns usually about the accuracy of the methods, and particularly the use of interviews and questionnaires are criticised. A so called job-exposure matrix, which associates various jobs with particular exposures, is often seen as a preferable instrument, but this way of assessing exposure may be less adequate so as to lead to a spurious reduction in the risk estimates in a study. For example, by comparing the performance of experts versus a job-exposure matrix, it was found that the sensitivity of a job-exposure matrix was low (23–63%) compared to judgements by experts, whereas the specificity was rather high (87–98%) [6]. Furthermore, assuming an odds ratio of 3 and an exposure prevalence of 10%, and taking the experts' classification of exposure to be completely correct, the use of a job-exposure matrix led to attenuation of the odds ratio by a factor of 1.5–2.1, and to a loss of power equivalent to a reduction in the number of subjects by a factor of 5–10. On the other hand, there is also the experience that a job-exposure matrix performed better than self-reported exposure in discriminating high-

risk subgroups in a study of lung cancer and asbestos exposure among construction workers [7]. Observations of this kind certainly suggest that risk estimates can easily be biased towards the null rather than being exaggerated, i.e., the difficulties in assessing exposure may rather favor non-positive results over any indication of a risk.

Sometimes large scale cohort studies based on register linkage are set up to confirm or refute observations made in smaller original studies. Census data on occupation may be used in such studies but usually reflect the occupational status only at a point in time (e.g., during some particular week), and are therefore inherently poor measures of the occupational exposures that may, or may not, have occurred over many years. The imprecise measure of exposure in such linkage studies may therefore attenuate an effect. The size of such register-based studies may seem impressive, however, and a non-positive result may therefore easily be taken as indicating the absence of any risk and as a refutation of a positive result shown in an earlier investigation. An example in this respect may be drawn from Swedish experiences regarding a possible cancer risk from phenoxy herbicides [8]. Hence, a large cohort study based on record linkage has not been able to reproduce the risk of soft-tissue sarcomas and lymphomas among subjects with a licence for using some but not all pesticides [9,10]. However, by just considering occupational titles the risk found in the case control studies is also very much reduced as shown in Table 1 [8,11,12]. This table also illustrates that observation bias is unlikely to explain the risks seen. Hence, those unexposed in the occupations where exposure may occur have an about normal risk and not a

lower than expected risk as would have been the result of misclassification of unexposed cases as exposed and/or a reversed misclassification for the controls [13].

### Confounding

Uncontrolled confounding is often a concern when an excess risk is found in a study but can also obscure an effect, either when the confounding risk factor tends to be more common in absence of the exposure or when there is a protective factor occurring among the exposed. However, confounding effects tend to be weaker than usually appreciated. The role of confounding can be evaluated by calculations given that there is reasonably good information about the frequency of occurrence and the magnitude of the risk or prevention exerted by the potential confounding factor. This latter condition is usually fulfilled as factors not known to exert a reasonably well-defined health effect cannot seriously be proposed to exert confounding. The reason for this is that causal relations in general could be refuted if any kind of factor or activity would be seriously taken as an alternative cause as associated with the exposure in a study. The same reasoning may be applied to any experimental research as well.

For negative and non-positive studies the concern should be about possible negative confounding masking an effect. Any criticism in this respect is rarely brought up, however, whereas for positive studies there are almost always suggestions about possible positive confounding exaggerating an effect or even being totally responsible for an elevated risk estimate. Table 2 shows the relatively weak confounding effects that can be exerted by a risk factor of varying

**Table 1.** Exposure to phenoxy acids in cases and referents by occupation in one of the soft-tissue sarcoma (STS) studies [11] and in one of the Hodgkin's (HD) and non-Hodgkin's lymphoma (NHL) studies [12]

	Agriculture/Forestry*			Other occupations		
	Ph**	Ch**	Unexp**	Ph**	Ch**	Unexp**
Soft-tissue sarcomas	13	1	17	1	10	68
Referents	5	3	39	0	5	167
Odds ratio	6.4	–	1.1	–	4.9	(1.0)
Lymphomas	35	5	37	6	15	71
Referents	23	1	114	1	7	189
Odds ratio	4.1	–	0.9	–	5.7	(1.0)

\* Considering occupations in agriculture and forestry only, the odds ratios reduce to  $(31)(172)/(47)(79) = 1.4$  for STS vs. a register linkage study [9] relative risk of 0.9 (95% CI: 0.8–1.0), and to  $(77)(197)/(138)(92) = 1.2$  for the lymphomas vs. a register linkage study [10] relative risk for HD of 1.20 (0.60–2.16) and for NHL of 1.01 (0.63–1.54).

\*\* Ph – phenoxy acids, Ch – chlorophenols, Unexp – unexposed.

**Table 2.** Example of confounding rate ratios obtained at different strengths of the confounding factor in terms of a rate ratio/relative risk, R, and degrees of association between exposure and confounding factor among exposed/unexposed ( $F_1/F_0$ )

$F_1$	$F_0$	$F_1/F_0$	Calculated confounding rate ratios by a given R			
			R = 1.5	R = 2.0	R = 4.0	R = 10.0
0.2	0.1	2.00	1.04	1.09	1.23	1.47
0.1	0.2	0.50	0.96	0.92	0.81	0.68
0.5	0.3	1.67	1.08	1.15	1.32	1.49
0.3	0.5	0.60	0.93	0.87	0.76	0.67
0.9	0.7	1.29	1.07	1.12	1.19	1.25
0.7	0.9	0.78	0.93	0.89	0.84	0.80

strength and which is associated with the exposure to various degrees as indicated in the table [14].

In case-control studies, matching is usually appreciated as a measure to control for confounding, but the problem involved in so called over-matching, which can obscure an existing risk, is not always understood. Hence, if matching is undertaken on some characteristic that is associated with the exposure, but has no bearing on any risk, the consequence will be that the controls tend to be exposed when sharing this exposure-related characteristic with the cases. Consequently, there is loss of power in the study and, dependent on proper analysis or not, the odds ratio may also be biased towards unity. Non-epidemiologists seem especially to believe that matching would have increased the validity of a case-control study, whereas this design procedure sometimes may have led to the opposite effect. Hence, the improper design may be unrecognised and a weakly positive result, i.e., a non-positive study, would be the consequence of the over-matching and can be mistaken as indicating no effect of an hazardous exposure. In cohort studies, however, matching is useful for increasing validity and there is no problem corresponding to the over-matching in case-control studies.

#### Comparability of populations and the “healthy worker effect”

In cohort studies the choice of a proper comparison population is important to prevent the so called “healthy worker effect” [15] from spuriously reducing the risk estimates. Health-related departures from the labor force may occur particularly among low socio-economic groups and more skilled jobs tend to recruit workers with different lifestyles from workers in less skilled jobs [16]. Therefore, the

general population is also likely to be less healthy than a certain group of workers, but there could be exceptions for jobs with a particularly low status. This health-related selection process makes it difficult to find proper comparison groups and explains why various worker groups often enjoy better health outcomes than expected [17,18], thus easily resulting in an underestimation of a health risk from some occupational exposure.

There is also reason to distinguish between a healthy worker survivor effect operating on a long-term basis versus a healthy worker effect in the period shortly after hire. The former may cause cumulative exposure to become associated with good health among the long-term employees and has a tendency to depress the upper end of an exposure-response curve and a weak or lacking exposure-response relationship may well be interpreted as no effect from the exposure. Regarding the early period of follow-up, pre-employment health examinations may create a strong selection for a healthy worker effect that obscures a health risk of an exposure.

In cohort studies, the healthy worker effect often results in a total mortality of about 90% or less of the expected. This healthy worker effect is usually greater in the younger age groups in a cohort and in the early phase of follow-up [19]. Cardiovascular deaths particularly tend to contribute most to the healthy worker effect, but other causes of death may also be below expected levels. Sometimes the observed number of deaths is as low as only about 50–60% of the expected, as, for example, in some studies of cardiovascular disease [20], and other non-cancer deaths [21], but such quite strong effects have also appeared for cancer [22].

When the healthy worker effect appears to be strong, the comparison of the observed number of cases with expect-

ed numbers based on national or regional rates is a questionable basis for concluding that no adverse effects are present. Usually there is no alternative reference population, however, but cohort studies of this kind showing no effect should be looked upon as essentially uninformative or non-positive rather than negative, unless perhaps when there is a large number of cases [1,3].

Cross-sectional studies may underestimate or fail to detect an existing risk due to health-related selection out of the job, i.e., those with symptoms in connection with the exposure tend to leave the job to find other work tasks. For example, a cross-sectional study of animal feed workers revealed a decreasing prevalence of most chronic respiratory symptoms with increasing years of exposure to dust and endotoxin [23], likely reflecting that a health-related selection out of the job had taken place. Nevertheless, for studying many medically less serious health problems there is no other realistic possibility than the cross-sectional approach, for example, in studies of lung or renal dysfunctions, neurobehavioural or neurophysiologic disturbances, or musculoskeletal and other non-lethal disorders.

In so called hospital-based case-control studies, recruiting the controls from among patients with other diseases than that at consideration, the healthy worker effect may lead both to upwards and downwards biased risk estimates.

Hence, if the workers under study are healthier and need less hospital care than the average population, the exposure will be underrepresented among the controls and the risk estimate biased upwards. On the other hand, if there are various diseases associated with the exposure (or some other health hazard associated with the exposure, implying a sort of confounding on the controls side), the exposure frequency among the controls may be too high in relation to the true exposure frequency in the population delivering the cases during the period of time covered by the study [24]. The result would then be a downwards biased risk estimate. These same aspects also apply to case-control studies that utilize deceased controls and to so called proportional mortality studies.

#### Non-positive studies in the aggregate and the benefit of the doubt

Small-scale, non-positive studies are less likely to be submitted or accepted for publication but should be made available for meta-analyses to assess a possible risk. An example in this respect can be found in the studies of trichloroethylene as a possible carcinogenic agent. Animal studies of trichloroethylene indicated a carcinogenic effect from this compound and subsequently also triggered epidemiologic investigations in the late 1970s [25,26]. The

**Table 3.** Meta-analysis of cohort mortality studies on trichloroethylene exposure. Expanded and updated Table 6 from Morgan et al.[32]

Cancer site	Antilla et al. [27]	Axelsson et al. [28]	Blair* et al. [31]	Morgan et al. [32]	Boice et al. [33]	Hansen et al. [34]	Total	Meta-SMR	95%CI
Liver									
Observed	5	4	15	6	9	5	44	1.39	1.01–1.87
Expected	2.2	2.8	11.5	6.1	7.0	2.0	31.6		
Prostate									
Observed	13	26	54	21	32	6	146	1.06	0.85–1.20
Expected	9.4	20.7	49.1	17.8	31.1	10.1	138.2		
Kidney									
Observed	6	6	15	8	7	3	45	1.18	0.86–1.59
Expected	6.9	5.2	9.4	6.1	7.1	3.3	38.0		
Bladder									
Observed	5	8	17	8	5	10	53	1.01	0.85–1.32
Expected	6.1	7.9	14.2	5.9	9.1	9.4	52.6		
Non-Hodgkin's lymphoma									
Observed	8	5	28	14	14	8	77	1.55	1.22–1.94
Expected	4.4	3.2	14.0	14.0	11.8	2.3	49.7		

\* Expected numbers for the Blair study are based on the internal comparison group as obtained by dividing the observed number by the relative risk; using this smaller reference population results in slightly too narrow confidence intervals (as here based on the Poisson distribution).

early results were less convincing of a cancer risk, and only by aggregating the results from three studies [27–29] including liver and biliary tract cancers as well as non-Hodgkin's lymphoma, could an IARC Working Group conclude that there was limited evidence for a carcinogenic effect from trichloroethylene in humans [30]. This view is now further supported by aggregating a number of other more recent studies, some of which may individually be seen as non-positive or even as so called negative studies and are listed in Table 3 [27,28,31–34].

Finally, when considering studies that suggest, but do not convincingly show an increased risk, such as the studies of trichloroethylene exposure, those who are exposed should be given the “benefit of the doubt” [3]. In balancing benefits against risk, one should consider who takes the risk and who gets the benefits. It may be remarked also that in occupational health, the situation is more complicated than in medical treatment, where the risk of adverse side-effects might be weighed against benefits for the same individual [35]. In contrast, occupational risks for workers may imply economic benefits for companies. Furthermore, high risks in small populations should be seen as more serious than small individual risks in large populations, the latter being scientifically interesting, but may or may not be relevant for immediate public health actions. Ethical guidelines in occupational health are important and a comprehensive discussion in this respect can be found elsewhere [36].

## REFERENCES

1. Hernberg S. “Negative” results in cohort studies. How to recognize fallacies. *Scand J Work Environ Health* 1981; 7 Suppl 4: 121–6.
2. Agnell M. *Negative studies* [Editorial]. *N Engl J Med* 1989; 321: 464–6.
3. Ahlbom A, Axelson O, Stöttrup Hansen E, Hogstedt C, Jensen UJ, Olsen J. Interpretation of “negative” studies in occupational epidemiology. *Scand J Work Environ Health* 1990; 16: 153–7.
4. Ahlbom A, Steineck G. Aspects of missclassification of confounding factors. *Am J Ind Med* 1992; 21: 107–12.
5. Plato N, Steineck G. Methodology and utility of a job-exposure matrix. *Am J Ind Med* 1993; 23: 491–502.
6. Stengel B, Pisani P, Limasset JC, Bouyer J, Berrino F, Hémon D. Retrospective evaluation of occupational exposure to organic solvents questionnaire and job exposure matrix. *Int J Epidemiol* 1993; 22 Suppl 2: S72–82.
7. Fletcher AC, Engholm G, Englund A. The risk of lung cancer from asbestos among Swedish construction workers self-reported exposure and a job exposure matrix compared. *Int J Epidemiol* 1993; 22 Suppl 2: S29–35.
8. Hardell L, Axelson O. Phenoxyherbicides and other pesticides in the etiology of cancer: some comments on Swedish experiences. In: Becker CE, Coye MJ, editors. *Cancer Prevention. Strategies in the Workplace*. Washington: Hemisphere Publishing Corporation; 1986.
9. Wiklund K, Holm L-E. Soft tissue sarcoma in Swedish agricultural and forestry workers. *J Natl Cancer Inst* 1986; 76: 229–34.
10. Wiklund K, Dich J, Holm LE. Risk of malignant lymphoma in Swedish pesticide applicators. *Br J Cancer* 1987; 56(4): 505–8.
11. Eriksson M, Hardell L, Berg NO, Möller T, Axelson O. Soft tissue sarcomas and exposure to chemical substances: a case-referent study. *Br J Ind Med* 1981; 38: 27–33.
12. Hardell L, Eriksson M, Lenner P, Lundgren E. Malignant lymphoma and exposure to chemicals, especially organic solvents, chlorophenols and phenoxy acids: a case-control study. *Br J Cancer* 1981; 43: 169–76.
13. Axelson O. A note on observational bias in case-referent studies in occupational health epidemiology. *Scand J Work Environ Health* 1980; 6: 80–2.
14. Axelson O. Aspects on confounding in occupational health epidemiology. *Scand J Work Environ Health* 1978; 4: 85–9.
15. McMichael AJ. Standardized mortality ratios and the “healthy worker effect” scratching beneath the surface. *J Occup Med* 1976; 18: 165–8.
16. Dahl E. Social inequality in health – the role of the healthy worker effect. *Soc Sci Med* 1993; 36: 1077–86.
17. Arrighi HM, Hertz-Picciotto I. The evolving concept of the healthy worker survivor effect. *Epidemiology* 1993; 5: 189–96.
18. Choi BC. Definition, sources, magnitude, effect modifiers, and strategies of reduction of the healthy worker effect. *J Occup Med* 1992; 34: 979–88.
19. Pearce NE, Checkoway H, Shy CM. Time-related factors as potential confounders and effect modifiers in studies on an occupational cohort. *Scand J Work Environ Health* 1986; 112: 97–107.
20. Ott MG, Skory LK, Holder BB, Bronson JM, Williams PR. Health evaluation of employees occupationally exposed to methylene chloride. Mortality. *Scand J Work Environ Health* 1983; 9 Suppl 1: 8–16.
21. Beall C, Delzell E, Macaluso M. Mortality patterns among women in the motor vehicle manufacturing industry. *Am J Ind Med* 1995; 28: 325–37.

22. Torchio P, Lepore AR, Corrao G, Comba P, Settini L, Belli S, et al. *Mortality study on a cohort of Italian licensed pesticide users*. *Sci Total Environ* 1994; 149: 183–91.
23. Smid T, Heederik D, Houba R, Quanjer PH. *Dust- and endotoxin-related respiratory effects in the animal feed industry*. *Am Rev Res Dis* 1992; 146: 1474–9.
24. Axelson O, Flodin U, Hardell L. *A comment on the reference series with regard to multiple exposure evaluations in a case-referent study*. *Scand J Work Environ Health* 1982; 8 Suppl 1: 15–9.
25. Axelson O, Andersson K, Hogstedt C, Holmberg B, Molina G, de Verdier A. *A cohort study on trichloroethylene exposure and cancer mortality*. *J Occup Med* 1978; 20: 194–6.
26. Tola S, Vilhunen R, Järvinen E, Korkala ML. *A cohort study of workers exposed to trichloroethylene*. *J Occup Med* 1980; 22: 737–40.
27. Anttila A, Pukkala E, Sallmen M, Hernberg S, Hemminki K. *Cancer incidence among Finnish workers exposed to halogenated hydrocarbons*. *J Occup Med* 1995; 37: 797–806.
28. Axelson O, Seldén A, Andersson K, Hogstedt C. *Updated and expanded Swedish cohort study on trichloroethylene and cancer risk*. *J Occup Med* 1994; 36: 556–62.
29. Spirtas R, Stewart PA, Lee JS, Marano DE, Forbes CD, Grauman DJ, et al. *Retrospective cohort mortality study of workers at an aircraft maintenance facility. I. Epidemiological results*. *Br J Ind Med* 1991; 48: 515–30.
30. International Agency for Research on Cancer. *Monographs on the evaluation of carcinogenic risks to humans*. Vol. 63. *Dry cleaning some chlorinated solvents and other industrial chemicals*. Lyon: International Agency for Research on Cancer; 1995.
31. Blair A, Hartge P, Stewart PA, McAdams M, Lubin J. *Mortality and cancer incidence of aircraft maintenance workers exposed to trichloroethylene and other organic solvents and chemicals: extended follow up*. *Occup Environ Med* 1998; 55: 161–71.
32. Morgan RW, Kelsh MA, Zhao K, Heringer S. *Mortality of aerospace workers exposed to trichloroethylene*. *Epidemiology* 1998; 9(4): 424–31.
33. Boice JD Jr, Marano DE, Fryzek JP, Sadler CI, McLaughlin JK. *Mortality among aircraft manufacturing workers*. *Occup Environ Med* 1999; 56: 581–97.
34. Hansen J, Raaschou-Nielsen O, Christensen JM, Johansen I, McLaughlin JK, et al. *Cancer incidence among Danish workers exposed to trichloroethylene*. *J Occup Environ Med* 2001; 43: 133–9.
35. Vineis P, Soskolne CL. *Cancer risk assessment and management. An ethical perspective*. *J Occup Med* 1993; 35: 902–8.
36. Samuels SW, editor. *The environment of the work place and human values*. *Am J Ind Med* 1986; 9: 1–113.