14. LUNG CANCER, ROASTER WORK AND SMOKING

14.1 Introduction

The purpose of the case referent study introduced in the previous chapter was, by using a limited number of individuals, to obtain further information which is rarely available for all the individuals in the cohort. Smoking is a factor which sometimes 'disturbs' the picture in a study in occupational medicine. It is rarely possible to collect information concerning the smoking habits of all the 3 915 workers. The 'disturbance' caused by the smoking factor can appear in two different ways: either differing smoking habits seemingly increase or decrease differences between two populations ('confounding'), or the smoking can have a (biological) role in 'modifying' the effect of a certain exposure (usually strengthening it). Both these aspects will be dealt with in this chapter.

14.2 Smoking habits among lung cancer cases and referents

The smoking habits of the cases and their birth-year matched dead referents were surveyed by means of interviews with close relatives. For further details see Pershagen et al. (1981). The smoking data obtained in this way are given in Table 14.1.

Assuming that the distribution of the smoking habits is the same in category 'quality unknown' as among other smokers, we obtain the picture displayed in Fig. 14.1.

On the basis of these smoking data, it is obvious that:

- There is a greater proportion of non-smokers among the referents as compared with the cases.
- The smoking referents seem to have smoked more on average than the smoking cases.

It was shown in the previous chapter, that the number of etiologic cases due to roaster work was 19.4 among the 76 lung cancer cases. Of course, we can apply exactly the same procedure for the estimation of etiologic cases due to smoking habits. We then arrive at the results given in Table 14.2.

Table 14.1

Smoking habits among lung cancer cases and referents. The amount of smoking is measured in number of cigarettes/day or equivalent

Smoking	Cases	First ref.	Second ref.
Never smoked	8	25	28
<5 cig./day	18	12	6
5–19 cig./day	31	16	17
20–39 cig./day	17	20	15
≥40 cig./day	1	2	1
Quantity unknown	1	1	9
Total	76	76	76

We know, however, that there were a certain number of individuals who had been working at the roaster departments and were smokers. We also know that there is not necessarily only one cause behind a given lung cancer case. Therefore, it is not logical to claim that there are 53 + 20 = 73 etiologic cases. The effects which smoking and roaster work might have on the lung cancer risk cannot be calculated separately in this way.

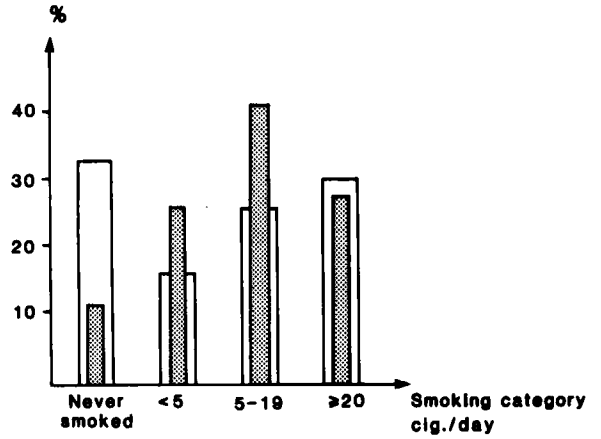

Fig. 14.1. Smoking habits among lung cancer cases and referents. ■ Cases; □ Referents.

Table 14.2

Separate effects of smoking habits and roaster work

Exposure	Category	Relative risk		Etiologic fraction		Etiologic cases
		RR	95% C.L	EF_{exp}	EF_{pop}	
Smoking	0	1.0	—	—		
	0+	4.55	(2.2; 9.7)	78%	69% ± 21%	53 ± 16
Roaster work	0	1.0	—	—		
	0+	2.43	(1.4; 4.3)	59%	26% ± 13%	20 ± 10

14.3 'Confounding'

It is indeed well known that smoking influences the lung cancer risk. However, it could also be true that smoking habits differ between those who have worked at the roasters and the other workers. If this is so, it must be taken into account when estimating the possible relationship between roaster work and lung cancer risk. The smoking habits constitute a so called 'confounding factor'.

If the actual individuals are classified with regard to both smoking habits and work at the roaster departments, we get the data given in Table 14.3.

It is obvious from Table 14.3, that smoking has been less common among roaster workers (17/38 = 45%) than among other workers (83/114 = 73%). This implies that if the smoking factor is neglected in the statistical analysis, it will in fact act as a negative confounder, i.e. it will 'dilute' the association between roaster work and lung cancer occurrence.

The confounding due to the smoking factor is also reflected by the discrepancy between the crude relative risk CRR = 2.43 and the standardized mortality ratio SMR = 3.39. Therefore, the ratio between CRR and SMR is sometimes called 'the confounding ratio' and is used as a measure of the degree of confounding.

The conditional relative lung cancer risks connected with work at the roaster departments are 2.46 and 3.63 for non-smokers and smokers respectively. This can be seen as a first indication that there is an interaction between smoking and roaster work with regard to the lung cancer risk.

Table 14.3

Lung cancer cases and referents, classified with regard to smoking and roaster exposure

Smoking	Category	Roaster exposure		RR
		0+	0	
0	Cases	5	3	
	Referents	21	31	2.46
0+	Cases	29	39	
	Referents	17	83	3.63
Total	Cases	34	42	CRR = 2.43
	Referents	38	114	SMR = 3.39

14.4 Upper limit of the effect due to confounding

We already know that there is a considerable difference in the number of lung cancer occurrences between persons who have worked at the roasters and the other Rönnskär workers. There is a similar difference between Rönnskär workers and Swedish men in general. What portion of such a difference could be explained by another causal factor of some kind? An upper limit for this possible effect was given by Cornfield et al. (1959). Their result is of great interest, especially in situations where individual data are lacking on the possible confounding factor(s).

Let us consider a certain period of time. We assume that a basic lung cancer incidence $I_0 \neq 0$ exists for individuals who are not smokers and who are not exposed to any special hazardous factor. Further, suppose that smoking gives the relative risk $RR_s > 1$ so that the smokers have the lung cancer incidence $RR_s \cdot I_0$. Analogously, we assume that work at the roasters is connected with the relative lung cancer risk RRe. For the moment, we also assume that the relative risks will act multiplicatively. The proportion of smokers among those who have not worked at the roasters is denoted by P_0 and the proportion of smokers among the roaster workers is denoted by P_1. The expected proportions of lung cancer cases in the four sub-categories of the PAR are given in Table 14.4.

Let us now calculate an odds ratio with regard to lung cancer and roaster work, neglecting the individuals' smoking habits. This odds ratio will then be an estimate of an incidence ratio, which has the structure

$$T_1/T_0 = RR_e[P_1(RR_s - 1) + 1]/[P_0(RR_s - 1) + 1]$$
$$= RR_e[P_1 + 1/(RR_s - 1)]/[P_0 + 1/(RR_s - 1)].$$

Table 14.4

Expected proportions of lung cancer cases in various sub-groups of the PAR

Smoking habits	Roaster work	
	Non-exp.	Exposed
Non-smoker	$(1 - P_0)I_0$	$(1 - P_1)I_0 \cdot RR_e$
Smoker	$P_0I_0 \cdot RR_s$	$P_1I_0 \cdot RR_s \cdot RR_e$
Total	T_0	T_1

It is easy to see that

$$\lim_{RR_s \to \infty} T_1/T_0 = RR_e \cdot P_1/P_0$$

which means that

$$T_1/T_0 \leq RR_e \cdot P_1/P_0 = P_1/P_0 + (RR_e - 1) \cdot P_1/P_0.$$

It is of special interest to study the case when $RR_e = 1$, which corresponds to the situation where the lung cancer risk is completely unaffected by roaster work. In this situation, the incidence ratio can at most be P_1/P_0. Consequently, P_1/P_0 is the upper limit of the possible distortion due to smoking habits on the relative lung cancer risk, calculated with regard to roaster exposure. Suppose for example that we arrive at $RR_e = 2.3$. If this value was totally 'explained' by differences in smoking habits, then smoking must have been at least three times more common among roaster workers than among other workers.

Fig. 14.2a) demonstrates the confounding effect as a function of the relative risk (RR_s) connected with smoking. Suppose, for example, that in a comparison between Rönnskär workers and Swedish men, 50% of the latter are smokers ($P_0 = 0.5$). If $RR_s = 1$ (i.e. smoking is non-hazardous), this factor cannot of course be a confounder and the estimate of the relative risk will be the same, irrespective of the value of P_1. However, if $RR_s = 5$ and $P_1 = 0.75$ (i.e. the proportion of smokers was 75% among the Rönnskär workers, cf 50% for Swedish men), a relative risk for the Rönnskär workers of the magnitude $RR_e = 1.3$ could be completely 'explained' by the difference in smoking habits. If, say, $RR_s = 50$ and all the Rönnskär workers were smokers (which means $P_1 = 1$), this could 'explain' a relative risk $RR_e = 2$.

Even if it would appear that the Rönnskär workers in the thirties and forties were smokers to a larger extent than Swedish males in general, this fact could not explain the large excess risks found in Chapter 7 (see for example Figure 7.3).

Thus, it is true that confounding due to different smoking habits can at most give the value P_1/P_0 of the relative risk. If, say, $P_0 = 0.3$ and $P_1 = 0.7$, this would give a maximum value of $RR_e = 2.33$, due to the differences in smoking habits. Cornfield (1951) has also shown that a confounding factor can partly or totally explain an observed relative (excess) risk, only if the relative risk connected with the confounding factor and the ratio P_1/P_0 are both greater than the CRR value. These results are of great interest, especially in situations where actual data are not available concerning the confounding factor but where certain general judgements can be made about what is realistic or not.

In this study, work at the roaster departments was connected with a relative lung cancer risk CRR = 2.43. This value is an underestimate, since it does not take into account the fact that there were fewer smokers among the

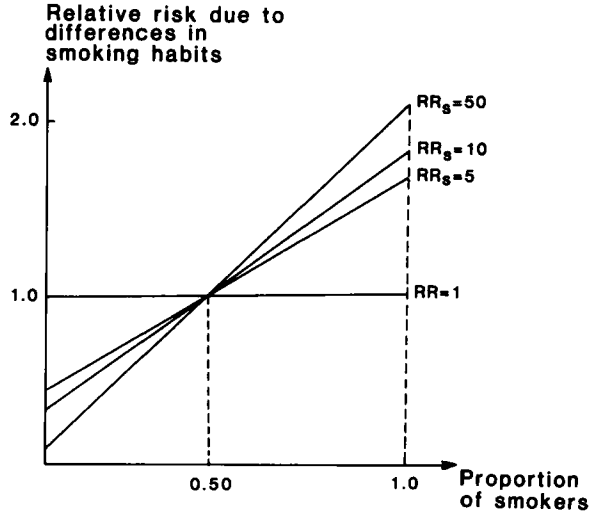

Fig. 14.2. a) Relative risk connected with a special exposure, as a result of confounding with, say, smoking. RR_s is the relative risk due to smoking and P_1 = the proportion of smokers in the exposed group, assuming that the proportion of smokers in the non exposed group is $P_0 = 0.5$;

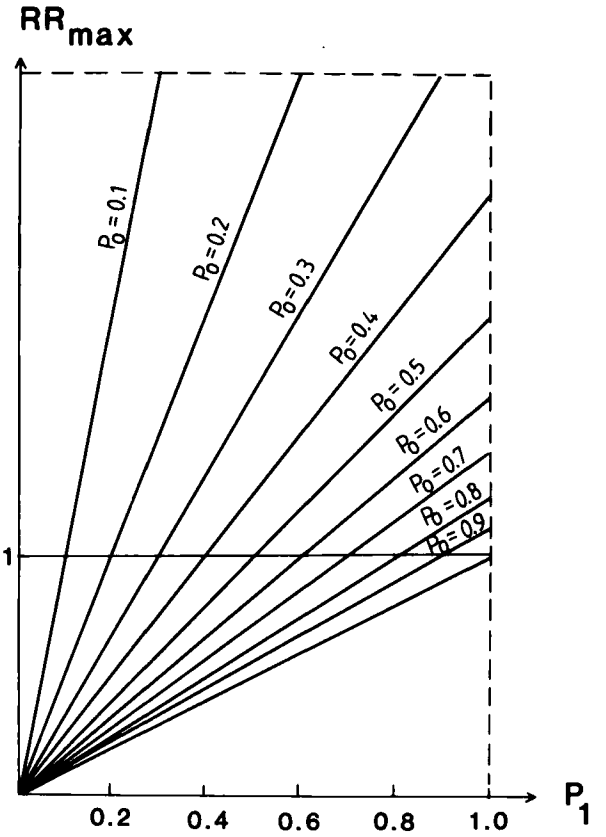

Fig. 14.2. b) RR_{max} as a function of P_1 and P_0.

workers at the roaster departments than among the other Rönnskär workers. In fact, $P_1/P_0 = (0.45)/(0.73) = 0.62$. Suppose that $RR_s = 5$ is the relative risk connected with smoking and that I_0 is the baseline lung cancer incidence. The total lung cancer incidence among workers who have

not worked at the roaster departments is then

$$I_{non\text{-}exp} = 0.73 \cdot 5 \cdot I_0 + 0.27 \cdot I_0 = 3.92 \cdot I_0.$$

If RR_e denotes the relative risk connected with roaster work, the total incidence at the roaster departments can be written

$$I_{exp} = RR_e[0.45 \cdot 5 \cdot I_0 + 0.55 \cdot I_0] = RR_e \cdot 2.80 \cdot I_0.$$

If we now enter $RR_e = 1$, this gives the incidence ratio $I_{exp}/I_{non\text{-}exp} = 2.80/3.92 = 0.71$. This value corresponds to the situation where the roaster work has no association with the lung cancer risk. Consequently, 0.71 can be used as a basis (instead of 1.0) for an adjustment of the CRR value. It is worth noting that

$$2.43/0.71 = 3.42 \approx SMR$$

14.5 Analysis of joint effects

When estimating the effects of several operating factors simultaneously, it would be suitable to base the analysis on a theoretical model, specifying how the various factors interact. A fundamental problem is, however, that such a model can be made in many different ways and it is not always self-evident as to which specifications are the most realistic.

We can apply the notation principles introduced previously. The incidence increase due to smoking for a worker who has not been exposed to the roaster departments can thus be written $RR_s \cdot I_0 - I_0 = I_0(RR_s - 1)$. Analogously, the incidence increase due to roaster work for a non-smoking worker can be written $I_0(RR_e - 1)$. We could then assume that the lung cancer incidence increase for workers who smoke and who work at the roaster departments is $I_0 + I_0(RR_e - 1) + I_0(RR_s - 1)$. This means that we have adopted an additive model.

Another approach is to assume that if smoking causes a RR_s-fold increase in lung cancer incidence and roaster work an RR_e-fold increase, then these two risk factors, acting simultaneously, would cause an $RR_s \cdot RR_e$-fold increase. This is in accordance with a multiplicative model. It is worth noting that this model is additive in the logarithms of the relative risks and is thus the basis of the log-linear approach presented in Chapter 12.

If the number of actual lung cancer cases among the smoking roaster workers is larger than was expected from the theoretical model, we say that the two risk factors (smoking and roaster work) are synergistic. This means that when they act together the effects are strengthened. On the other hand, if the number of actual lung cancer cases is lower than was expected from the model, then so called antagonism occurs. It is important to notice that the concepts of synergism and antagonism are well-defined only on the basis of a specified, theoretical model. It could very well happen that when analysed by means of an additive model, a certain data set gives the conclusion that there is an obvious synergism. If, on the other hand, the same set of data is then analysed on the basis of a multiplicative model, it is possible to reach the conclusion that there is no synergism at all.

Table 14.5

Classification of cases and referents with regard to roaster exposure and smoking

Smoking	No	No	Yes	Yes
Roaster work	No	Yes	No	Yes
Cases	3	5	39	29
Referents	31	21	83	17
RR	1.0	2.46	4.86	17.63
Lower 95% C.L	—	0.54	1.5	5.5
Upper 95% C.L	—	11.2	15.4	56.0

Breslow and Day (1980) present a long discussion on the choice of model. On the basis of empirical data, the multiplicative model has often appeared to be more realistic than the additive one, especially with regard to cancer diseases.

When studying the joint effects of smoking and roaster work, the natural basis is the data displayed in Table 14.5. When calculating relative risks, the group of non-smoking, non-roaster workers has been used as the reference point. As the group of non-smoking workers is small, the relative risk estimates have very wide confidence limits.

The relative lung cancer risk, connected with roaster work, was 2.46 among non-smokers and (17.63/4.86 = 3.63) among smokers. This indicates a certain interaction effect. The relative risk connected with smoking was 4.86 and 17.63/2.46 = 7.17 among non-roaster workers and roaster workers respectively. The joint effect of smoking and roaster work is illustrated in Figure 14.3.

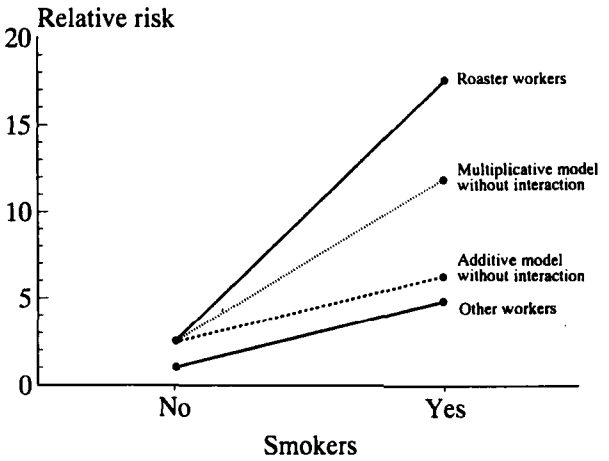

Fig. 14.3. Estimated joint effects of smoking and roaster exposure and the corresponding expected effects on the basis of models without interaction.

Based on the multiplicative model, the relative lung cancer risk with simultaneous exposure to smoking and roaster work is 11.96. When applying the additive model, the relative risk estimate is only 6.32. In the population in question there is apparently a synergistic effect, over and above the multiplicative model.

A relative risk can be estimated for each combination of risk factors (Table 14.5). Consequently, it is also possible to estimate the number of etiologic cases for each sub-category. The total number of etiologic cases is 61 and may be traditionally formulated as: 'assuming the estimated associations to be causal, there were 61 cases among the 76 lung cancer cases, which would not have occurred if the persons in question had not smoked and/or had not worked at the roaster departments'.

14.6 Analysis of further exposure information

It is obvious that the results of a statistical analysis can only reveal information contained in the data—not more. Until now, the interest has been focused on the possible effect that the factor 'roaster work' might have had on the lung cancer occurrence. In accordance with the explanation (b) in section 13.1, this factor might be a rough indicator of something more specific, but, to take this further, it would be necessary to obtain more detailed information on the individuals themselves.

We know that work at the roasters means exposure to a series of more or less potentially hazardous substances. We also know that during the decades studied here, the environment for the roaster workers has been altered. As shown in Table 1.1, the environment at the roaster departments has contained dust, arsenic, sulphur dioxide and nickel. Arsenic, especially, is an important component and was present in large quantities during the thirties and forties. Therefore, we can deepen the analysis by investigating the possible importance of arsenic on the lung cancer risk taking into account that the amount of arsenic has varied during the period in question. Hopefully, this will give a more sensitive risk indicator than that provided simply by 'exposure to roaster work'. It must be noted that arsenic has not only been present at the roaster departments, but also at certain other work sites in the smelter.

Table 14.6

Cases and referents classified with regard to arsenic exposure and smoking

Age (years)	Arsenic work Smoking	No No	Yes No	No Yes	Yes Yes
-59	Case	0	3	15	9
	Ref	2	4	16	3
60-69	Case	0	2	7	23
	Ref	6	14	22	26
70-	Case	1	2	2	12
	Ref	6	20	11	22
Total	Case	1	7	24	44
	Ref	14	38	49	51
SRR		1.0	3.0	4.9	14.6

Pershagen et al. (1981) investigated the effects of arsenic and sulphur dioxide on the 76 lung cancer patients and their matched referents. The sulphur dioxide exposure did not demonstrate any clear association with lung cancer occurrence, unlike the arsenic exposure. The main results are displayed in Table 14.6. The relative lung cancer risk associated with arsenic exposure is approximately 3 and the relative lung cancer risk connected with smoking is approximately 5. Further analysis and discussion can be found in the original report (Pershagen et al. 1981).

14.7 Conclusions thus far

It is obvious from the above results, that both arsenic exposure and smoking influence the lung cancer risk. The pattern is in accordance with the multiplicative model. A calculation here of the number of etiologic cases arrives at 65, which is somewhat more than the 61 cases explained by the two variables, smoking and roaster work.