Passive Smoking

Can observed associations with health be artefacts of measurement error?

Author : P.N. Lee
Date : 11.10.83.

In many situations in epidemiology, error in measurement of exposure variables has no effect on the direction of an association, serving only to make the magnitude of an observed association less marked than the association that actually exists. This is not always so, however, and the purpose of this note is to highlight some situations relevant to the passive smoking story in which, as a result of error in measurement of exposure variables, an apparent association can occur when no real association exists at all.

The first situation relates to where one is standardising for a confounding variable that is strongly related to risk of disease and is also related to passive smoking, and where the confounding variable is not accurately measured. An example of this sort of situation is the relationship between respiratory symptoms or pulmonary function in young children and the smoking habits of the mother. One knows that lower social class related factors are very important determinants of these conditions and that lower social class is smoking related. One also knows that social class must be an inaccurately determined measure of the true aetiological agent which social class is surrogate for.

Let us illustrate what can happen by a simplified numerical example in which (hopefully) reasonable looking assumptions are made.

Let us assume the following:

(a) we study a population of 1000 children in which half the mothers smoke,

(b) maternal smoking has no effect on risk of the condition we are studying,

(c) there is an aetiological agent which multiplies risk by 5,

(d) half the children are exposed to this agent,

(e) the agent is closely related to social class, so that 80% of those at risk from the agent are in social class "low" and only 20% are in social class "high",

(f) the agent is correlated with maternal smoking, so that 60% of those at risk from the agent have mothers who smoke while only 40% of those not at risk from the agent have mothers who smoke.

We can then set up a table of expected results as follows:

		---- Observable ----			
Group Number	Exposure to agent	Social class	Maternal smoking	Number in group	Risk* of group
1	+	Low	Yes	240	5
2	+	Low	No	160	5
3	+	High	Yes	60	5
4	+	High	No	40	5
5	-	Low	Yes	40	1
6	-	Low	No	60	1
7	-	High	Yes	160	1
8	-	High	No	240	1

(* relative to same constant)

If we are epidemiologists given the observable expected data, what do we conclude? First, we start off by looking at the simple

relationship between maternal smoking and risk.

Groups	Maternal smoking	Number in group	Mean Risk	Relative Risk
1+3+5+7	Yes	500	3.40	1.31
2+4+6+8	No	500	2.60	1

We note that there is an association between maternal smoking and risk. Next, we look for potential confounders. We note that social class is related to risk.

Groups	Social class	Number in group	Mean Risk	Relative Risk
1+2+5+6	Low	500	4.20	2.33
3+4+7+8	High	500	1.80	1

The next step, therefore, is to see if standardisation for social class removes the relationship between maternal smoking and risk. We compute two tables as follows:

Groups	Social class	Maternal Smoking	Number in group	Mean Risk	Relative Risk
1+5	Low	Yes	280	4.43	1.13
2+6		No	220	3.91	1
3+7	High	Yes	220	2.09	1.33
4+8		No	280	1.57	1

We note that within each social class group, there still remains a positive association with maternal smoking, which averaged overall (by direct standardisation to the overall social class distribution) gives a relative risk figure of 1.19.

Three conclusions, of general application, can be made from this example:

(a) the observed association for the inaccurately measured true risk factor is less (relative risk 2.33) than the true

association for the actual risk factor (relative risk 5).

(b) the association between the true risk factor and maternal smoking results in an apparent association (relative risk 1.31) between maternal smoking and the condition of interest, when no true association exists.

(c) this apparent association is reduced by standardisation for the inaccurately measured risk factor (to 1.19), but is not eliminated. An apparent association still exists even though maternal smoking has been assumed in our example to have no actual association with risk.

The second situation relates to the studies of Hirayama, Trichopoulos, Garfinkel, etc., in which the smoking habit of the spouse is related to risk of lung cancer in non-smoking women. Here we know that active smoking is strongly related to risk and that smokers tend to marry smokers and non-smokers to marry non-smokers. If we assume some smokers deny their smoking, can this result in an apparent association between spouse's smoking and lung cancer when no true association exists? Formally, then, let us look at an example with the following assumptions:

(a) we study a population of 100,000 married women, in which half the wives smoke,

(b) half their husbands smoke also, but their habits are associated with their wives, so that in women who smoke 60% of husbands smoke, while in women who do not smoke 40% of husbands smoke.

(c) active smoking in women multiplies risk of lung cancer by 20,

(d) passive smoking has no effect on risk,

(e) 5% of men and women who smoke deny smoking on interview.

Again setting up a table of expected results, we have:

Group Number	Actual smoking habits Wife	Husband	Number in group	Self-reported smoking habits Wife	Husband	Number in group	Risk of group
1A	Smoker	Smoker	30,000	Smoker	Smoker	27075	20
1B	"	"		"	Nonsmoker	1425	20
1C	"	"		Nonsmoker	Smoker	1425	20
1D	"	"		"	Nonsmoker	75	20
2A	"	Nonsmoker	20,000	Smoker	Nonsmoker	19000	20
2B	"	"		Nonsmoker	Nonsmoker	1000	20
3A	Nonsmoker	Smoker	20,000	Nonsmoker	Smoker	19000	1
3B	"	"		Nonsmoker	Nonsmoker	1000	1
4	"	Nonsmoker	30,000	Nonsmoker	Nonsmoker	30000	1

Given the self-reported data one can make three observations:

Groups	Wife	Husband	Number in group	Mean Risk	Relative Risk
1A,1B,2A	Smoker	Either	47500	20	10.50
All others	Nonsmoker	Either	52500	1.90	1
1A	Smoker	Smoker	27075	20	1
1B,2A	Smoker	Nonsmoker	20425	20	1
1C,3A	Nonsmoker	Smoker	20425	2.33	1.42
1D,2B,3B,4	Nonsmoker	Nonsmoker	32075	1.64	1

In words, we observe

(a) the observed association between self-reported active smoking and risk of lung cancer (10.5) is less marked than that assumed to actually exist between actual active smoking and risk of lung cancer (20).

(b) In self-reported smokers, an unbiassed estimate of the effect of husband's smoking (1) is obtained.

(c) In self-reported non-smokers, an apparent effect of husband's smoking is seen (relative risk 1.42) when no true effect exists.

It is also of interest to look at more Japanese style assumptions. For the sake of illustration, let us assume

(a) we study a population of 100,000 married women in which 15% of the wives smoke,

(b) 75% of their husbands smoke, but their habits are associated, so that for women who smoke we find that 13/15 (87%) of their husbands smoke whereas for women who do not smoke only 62/85 (73%) of their husbands smoke.

(c) active smoking in women multiplies risk of lung cancer by 20,

(d) passive smoking has no effect on risk,

(e) men do not deny smoking but 20% of women do (smoking being an unacceptable habit for women in Japan).

Here the table is:

Group Number	Actual smoking habits: Wife	Husband	Number in group	Self-reported smoking habits: Wife	Husband	Number in group	Risk of group
1A	Smoker	Smoker	13000	Smoker	Smoker	10400	20
1B	"	"		Nonsmoker	Smoker	2600	20
2A	"	Nonsmoker	2000	Smoker	Nonsmoker	1600	20
2B	"	"		Nonsmoker	Nonsmoker	400	20
3	Nonsmoker	Smoker	62000	Nonsmoker	Smoker	62000	1
4	"	Nonsmoker	23000	Nonsmoker	Nonsmoker	23000	1

Here the "Hirayama" observations, based on the self-reported data, produce the following:

Groups	Wife	Husband	Number in group	Mean Risk	Relative Risk
1B,3	Nonsmoker	Smoker	64600	1.76	1.33
2B,4	Nonsmoker	Nonsmoker	23400	1.35	1

Again one ends up with an apparent relationship with passive smoking when no real relationship actually exists.

Crucial to the extent of the apparent relationship are both the extent of lying about smoking and the degree of correlation between husbands' and wives' smoking habits. The figures given are pure guesstimates, which seemed

intuitively reasonable to me at the time. It is of obvious importance to try to get data here though the difficulty of obtaining it is not to be underestimated - the extent of lying will doubtless depend on the way in which the smoking data are collected. Another important point is that my calculations depend on lying about smoking by the wife being independent of the husband's smoking habit. It is relatively easy to see that if women only lied if their husbands were non-smokers, a bias in the reverse direction would obtain. However, this does not seem all that plausible especially if the data were collected with the husband not present.

Obviously this note leaves many unanswered questions. However, its purpose is merely to indicate that when in fact no true effect of passive smoking actually exists at all, apparently reasonable assumptions about inaccuracy of measurement error can lead to observing artefactual associations between passive smoking and disease of a magnitude not dissimilar to those reported in the literature.