16 Bjorksten B, Naaber P, Sepp E, Mikelsaar M. The intestinal microflora in allergic Estonian and Swedish 2-year-old children. *Clin Exp Allergy* 1999;29:342-6.
17 Matricardi PM, Nisini R, Biselli R, D'Amelio R. Evaluation of the overall degree of sensitization to airborne allergens by a single serologic test: implications for epidemiologic studies of allergy. *J Allergy Clin Immunol* 1994;93:68-79.
18 Gruppo di Studio Interdisciplinare Solle Vaccinazioni dell'età Evolutiva. *Le vaccinazioni in Italia*. Milano: Centro Informazione Scientifica Editore, 1996.
19 Dixon WJ, ed. *BMDP statistical software*, vol 2. Berkeley: University of California Press, 1990.
20 Matricardi PM, Franzinelli F, Franco A, Caprio G, Murru F, Cioffi D, et al. Sibship size, birth order, and atopy in 11,371 Italian young men. *J Allergy Clin Immunol* 1998;101:439-44.
21 Beaman MH, McCabe RE, Wong S, Remington JS. Toxoplasma gondii. In: Mandell GL, Bennett JE, Dolin R. *Principles and practice of microbial diseases*, 4th ed. New York: Churchill Livingstone, 1995:2455-75.
22 Goodwin CS, Mendall MM, Northfield TC. Helicobacter pylori infection. *Lancet* 1997;349:265-9.
23 Grubel P, Huang L, Masubuchi N, Stutzenberger F, Cave DR. Detection of Helicobacter pylori DNA in houseflies (Musca domestica) on three continents. *Lancet* 1998;352:788-9.
24 Dore MP, Sepulveda AR, Osato MS, Realdi G, Graham DY. Helicobacter pylori in sheep milk. *Lancet* 1999;354:132.
25 Dominici P, Bellentani S, Di Biase AR, Saccoccio G, Le Rose A, Masutti F, et al. Familial clustering of Helicobacter pylori infection: population based study. *BMJ* 1999;319:537-41.
26 Holt PG, Sly PD, Bjorksten B. Atopic versus infectious diseases in childhood: a question of balance? *Pediatr Allergy Immunol* 1997;8:53-8.
27 Durkin HG, Chice SM, Gaetjens E, Bazin H, Tarcsay L, Dukor P. Origin and fate of IgE-bearing lymphocytes. II. Modulation of IgE isotype expression on Peyer's patch cells by feeding with certain bacteria and bacterial cell wall components or by thymectomy. *J Immunol* 1989;143:1777-83.
28 Sudo N, Sawamura S, Tanaka K, Kubo C, Koga Y. The requirement of intestinal bacterial flora for the development of an IgE production system fully susceptible to oral tolerance induction. *J Immunol* 1997;159:1739-45.
29 Reis e Sousa C, Hieny S, Scharton-Kersten T, Jankovic D, Charest H, Germain RN, et al. In vivo microbial stimulation induces rapid CD40 ligand-independent production of interleukin-12 by dendritic cells and their redistribution to T cell areas. *J Exp Med* 1997;186:1819-29.
30 Alm J, Swartz J, Lilja G, Scheynius A, Pershagen G. Atopy in children of families living with an anthroposophic lifestyle. *Lancet* 1999;353:1485-8.
31 Rook GAW, Stanford JL. Give us this day our daily germs. *Immunol Today* 1998;19:113-6.
32 Hesselmar B, Aberg N, Aberg B, Eriksson B, Bjorksten B. Does early exposure to cat or dog protect against later allergy development? *Clin Exp Allergy* 1999;29:611-7.
33 Braun-Fahrlander C, Gassner M, Grize L, Neu U, Sennhauser FH, Varonier HS, et al. Prevalence of hay fever and allergic sensitization in farmer's children and their peers living in the same rural community. SCARPOL team. *Clin Exp Allergy* 1999;29:28-34.
34 Jarvis D, Chinn S, Luczynska C, Burney P. The association of family size with atopy and atopic disease. *Clin Exp Allergy* 1997;27:240-5.
35 Burney P, Malmberg J, Chinn S, Jarvis D, Luczynska C, Lai E. The distribution of total and specific serum IgE in the European community respiratory health survey. *J Allergy Clin Immunol* 1997:99:314-22.
36 Yemaneberhan H, Bekele Z, Venn A, Lewis S, Parry E, Britton J. Prevalence of wheeze and asthma and relation to atopy in urban and rural Ethiopia. *Lancet* 1997;350:85-90.

(Accepted 1 December 1999)

Reanalysis of epidemiological evidence on lung cancer and passive smoking

J B Copas, J Q Shi

Abstract

Objective To assess the epidemiological evidence for an increase in the risk of lung cancer resulting from exposure to environmental tobacco smoke.
Design Reanalysis of 37 published epidemiological studies previously included in a meta-analysis allowing for the possibility of publication bias.
Main outcome measure Relative risk of lung cancer among female lifelong non-smokers, according to whether her partner was a current smoker or a lifelong non-smoker.
Results If it is assumed that all studies that have ever been carried out are included, or that those selected for review are truly representative of all such studies, then the estimated excess risk of lung cancer is 24%, as previously reported (95% confidence interval 13% to 36%, $P < 0.001$). However, a significant correlation between study outcome and study size suggests the presence of publication bias. Adjustment for such bias implies that the risk has been overestimated. For example, if only 60% of studies have been included, the estimate of excess risk falls from 24% to 15%.
Conclusion A modest degree of publication bias leads to a substantial reduction in the relative risk and to a weaker level of significance, suggesting that the published estimate of the increased risk of lung cancer associated with environmental tobacco smoke needs to be interpreted with caution.

Introduction

Exposure to environmental tobacco smoke (passive smoking) is widely accepted to increase the risk of lung cancer, but different epidemiological studies have produced varying estimates of the size of the relative risk. Hackshaw et al reviewed the results of 37 such studies that estimated the relative risk of lung cancer among female lifelong non-smokers, comparing those whose spouses (or partners) were current smokers with those whose spouses had never smoked.[1] Of the 37 studies, 31 reported an increase in risk, and the increase was significant in seven studies. The remaining six studies reported negative results, but none of these was significant. Pooling these results using a method which allows for statistical heterogeneity between studies, Hackshaw et al concluded that there is an overall excess risk of 24% (95% confidence interval 13% to 36%).[1] This is strong epidemiological evidence for an association between lung cancer and passive smoking ($P < 0.001$).

The approach used by Hackshaw et al does not allow for the possibility of publication bias—that is, the possibility that published studies, particularly smaller ones, will be biased in favour of more positive results. We reanalysed the results and looked for evidence of publication bias.

Methods and results

The figure shows the relative risks from the 37 epidemiological studies analysed by Hackshaw et al[1] plotted against a measure of the uncertainty in that relative risk. This uncertainty (s) decreases as the size of the study increases so that large studies are on the left of the plot and small studies on the right. The plot shows a trend for smaller studies to give more positive results than the larger studies (correlation = 0.35, $P < 0.05$, or $P = 0.012$ by Egger's test[2]). This graph is

Department of Statistics, University of Warwick, Coventry CV4 7AL
J B Copas
professor
J Q Shi
research fellow
Correspondence to:
J B Copas
jbc@stats.warwick.ac.uk

BMJ 2000;320:417–8

Estimated relative risk and number of unpublished smaller and larger studies for various values of publication probability

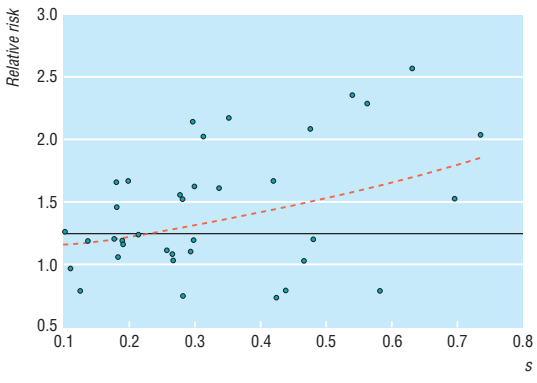
Publication probability	Relative risk (95% CI)	P value	No of unpublished studies*	
			Small	Large
0.6	1.11 (0.97 to 1.27)	0.110	36	24
0.7	1.13 (1.00 to 1.27)	0.052	23	15
0.8	1.15 (1.03 to 1.28)	0.014	14	9
0.9	1.18 (1.07 to 1.31)	0.002	7	4
1.0	1.24 (1.13 to 1.36)	<0.001	0	0

*Smaller studies s>0.4; larger studies s≤0.4.

similar to the funnel plot used in the meta-analysis of clinical trials, when a trend such as this is interpreted as a sign of publication bias.[3] This bias arises when a study is more likely to be written up and submitted to a journal and more likely to be accepted for publication if it reports positive results than if its results are inconclusive or negative. Since it is reasonable to assume that publication is more likely for larger (small s) than smaller (large s) studies, the problem of publication bias will be most evident among the smaller studies, as suggested by the figure. By "publication" we mean the whole process of selecting a study for review.

We reanalysed the results of the 37 epidemiological studies to allow for the trend evident in the figure. Our method describes the apparent relation between relative risk and study size by a curve. This gives a good fit to the observed points. The basic idea of the method is that there is no real relation between study outcome and study size, the relation that we observe is simply an artefact of the process of selecting these studies.

Our method has been published,[4] and further details are available from us on request. The estimated average relative risk depends on a statistical parameter that can be interpreted as the probability that a paper with a certain value of s is published (publication probability. If the publication probability is 1, all papers are published and so there is no possibility of publication bias; the relative risk is then estimated as 1.24 (24% risk excess), agreeing as expected with Hackshaw et al's result.[1] But smaller values of publication probability give smaller estimates of relative risk. We do not know how many unpublished studies have been carried out. Therefore there is no way of estimating the publication probability from any data: all we know is that there is a significant correlation in the funnel plot, so that some degree of publication bias is needed to explain this trend.



Plot of relative risk of lung cancer versus s (standard deviation of log relative risk). Overall weighted average of relative risk is shown as dotted line and fitted value (publication probability=0.8) as dashed line

Key messages

- A systematic review of epidemiological studies on passive smoking estimated the increased risk of lung cancer as 24%
- There is clear evidence of publication bias in these studies
- Reanalysis of the data allowing for the possibility of publication bias substantially lowers the estimate of relative risk

The table gives the estimated relative risk for values of publication probability between 0.6 and 1, together with 95% confidence intervals and P values. The P value is less than 5% only when the publication probability is more than about 0.7. The indirect estimate of 19% excess risk derived from studies on biochemical markers (table 5 of Hackshaw et al's paper[1]) agrees with the epidemiological analysis when the publication probability is about 0.9.

For any given value of publication probability it is possible to estimate the number of studies which have been undertaken but not published. This is shown in the final two columns of the table. If the publication probability is 0.8 then there are a total of 23 unpublished studies so that the 37 selected ones represent a sample of 37/60 = 62% of all such studies that have been undertaken. If this is the case, then the excess risk is likely to be closer to 15% than 24%. The dashed line in the figure shows the fit from our statistical model when the publication probability is 0.8; this curve fits the available evidence well.

Conclusions

Although the trend in the figure seems clear, Bero et al suggest that the number of unpublished studies is unlikely to be large,[5] and so the problem of publication bias may be less severe here than in systematic reviews of other aspects of medicine. However, the possibility of publication bias cannot be ruled out altogether, and at least some publication bias is needed to explain the trend we found. Our results show that the publication probability does not have to fall much below 1.0 before there is quite a substantial reduction in the estimated risk.

Contributors: The reanalysis was done jointly by JBC and JQS. JBS was responsible for the presentation of the paper and is the guarantor.

Funding: Economic and Social Research Council.

Competing interest: None declared.

1 Hackshaw AK, Law MR, Wald NJ. The accumulated evidence on lung cancer and environmental tobacco smoke. *BMJ* 1997;315:980-8.
2 Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple graphical test. *BMJ* 1997;315:629-34.
3 Egger M, Smith GD. Misleading meta-analysis. *BMJ* 1995;310:752-4.
4 Copas JB. What works; selectivity models and meta analysis. *J R Stat Soc Am* 1999;162:95-109.
5 Bero LA, Glantz SA, Rennie D. Publication bias and public health policy on environmental tobacco smoke. *JAMA* 1994;272:133-6.

(Accepted 1 November 1999)