Risk of lung cancer in relation to type of cigarette smoked (filter/plain and tar level)

A summary of the data with meta-analyses and heterogeneity analyses

Authors: P N Lee and Jan Hamling
Date: 21 May 2004

Executive summary

This is an updated version of a report written in 2002.

It provides a set of estimates of the lung cancer risk associated with a) smoking filter rather than plain cigarettes and b) smoking low tar rather than high tar cigarettes. The data set includes estimates unadjusted and adjusted for number of cigarettes smoked per day. Also included are results of various meta-analyses and heterogeneity analyses.

For the filter/plain comparison, overall relative risk estimates are in the range 0.61-0.66 and are highly statistically significant ($p<0.001$), whether based on fixed-effects or random-effects meta-analysis or whether or not allowance is made for the existence of overlapping populations in some of the estimates presented. The advantage to filter cigarettes was clearly and quite similarly evident in both sexes, in studies conducted in different continents and in studies conducted in different time periods. It was evident, regardless of whether adjustments were or were not made for age, cigarettes smoked per day or other factors.

Highly significant ($p<0.001$) heterogeneity between the filter/plain relative risk estimates adjusted for cigarettes smoked per day could be well explained by a multivariate model based on seven study characteristics. The major contributors to the model were the greater advantage to filter apparent in studies in which 100% histological confirmation of diagnosis was required and in studies which adjusted for age. Highly significant ($p<0.001$) heterogeneity was also seen between the filter/plain relative risk estimates unadjusted for amount smoked. Here however this could not be fully explained by the multivariate model used. However, the tendency for the advantage to filter to be greater in studies in which 100% histological confirmation of diagnosis was required was again evident.

For the low tar/high tar comparison, overall relative risk estimates were all statistically significantly below 1.00, but varied more depending on whether fixed-effects or random-effects meta-analysis was used, whether exclusion for overlaps was made and whether data unadjusted or adjusted for amount smoked was considered.

Relative risks adjusted for amount smoked did not show marked heterogeneity and there was little evidence that they varied by study characteristics. Relative risks unadjusted for amount smoked showed significant ($p<0.001$) heterogeneity, but this seemed more attributable to specific studies giving very low or very high estimates. The number of estimates for the low tar/high tar comparison was less than the number for the filter/plain comparison and did not readily allow a detailed investigation of the sources of heterogeneity.

Direct comparison within study of adjusted and unadjusted relative risk estimates did not demonstrate any clear increase in either the filter/plain or the low tar/high tar relative risk following adjustment. Some analyses showed a slight increase, but this was substantially less than the estimated reduction in risk associated with smoking filter or low tar cigarettes.

Although several new studies have been added, these analyses are little changed from those reported in 2002. The results strengthen the conclusions reached in Peter Lee's paper published in Inhalation Toxicology (2001; 13, 951-976).

Index

1. Introduction

In 2001 P.N. Lee published[1] a review of the evidence relating lung cancer to type of cigarette smoked, with major emphasis on the results relating to the smoking of filter or plain cigarettes and to the smoking of high or low tar cigarettes. At about the same time, a review of similar evidence was published in chapter 4 of Smoking and Tobacco Control Monograph 13[2]. The monograph claimed that much, if not all, of the advantage to filter or low tar cigarettes was spurious, caused by adjusting for number of cigarettes smoked per day in analysis, when switching to lower yield cigarettes actually resulted in an increase in amount smoked.

The first version of this document was produced in March 2002. It was not intended as a critique of the contents of Monograph 13[2] but as an update and extension of the tables and analyses presented in the review paper[1]. This is the second version of the report and has been updated to include papers published up to May 2004.

The data presented in this report differ from those given in the review paper[1] for two main reasons:

(i) They include data from more recent publications and also from other publications missed originally and

(ii) They present, if possible, two relative risks rather than one for each study/sex combination with relevant data. One relative risk would be adjusted for as many variables as possible including cigarettes/day (or in some studies pack-years). The other relative risk would be adjusted for as many variables as possible not including cigarettes/day, pack-years or inhalation (which may also be affected by brand switching).

The relative risks shown give the results of the most extreme comparison available. Thus, if risk by tar level in the source paper was given by, say, four levels of tar, relative risks for the lowest versus highest level of tar are shown. Similarly, if risk was categorised in smokers of, say, filter cigarettes only, mixed and plain cigarettes only, the relative risk presented is for the filter only versus plain only comparison. Relative risks are always

presented in the same direction (with low tar or filter as the numerator and high tar or plain as the denominator). Where the source paper presented relative risk estimates adjusted for varying sets of potential confounding variables, this report generally only includes the estimate adjusted for the most potential confounders. In these respects this report follows the practices of the review paper[1] and of the first version of this report.

In some cases the relative risks or confidence intervals presented differ from those given in the review paper. This followed further cross-checking of data against the original source, against data as presented in Monograph 13[2] and against data recorded in our in-house IESLC database. It should be noted that the majority of the estimates given in this report are not presented directly in the source paper, but had to be calculated from data provided. Sometimes this was only using simple reciprocals to convert e.g. plain/filter estimates to filter/plain estimates, but often more complex calculations had to be used, as described in the review paper[1].

The objectives of this document are three-fold:

(i) to present the updated estimates,

(ii) to present a number of meta-analyses, and

(iii) to investigate how risk varies by study attributes in a heterogeneity analysis.

2. The data

Table 1 presents the main data. For each of the 144 relative risk estimates given (numbered 1-144), the table shows:

(i) The name of the study from which it derived;

(ii) Details of the reference group (denominator) and comparison group (numerator). The first set of estimates relate to the filter/plain comparison and the information presented attempts to make it clear not only the filter/plain groups compared (e.g. only filter versus only plain, ever filter versus only plain, mainly filter versus mainly plain) but also the period the brand information relates to (lifetime, defined period, or "current", i.e. brand smoked at latest time point, typically at baseline for prospective studies). The other set of estimates relate to the low tar/high tar comparison and the information presented gives the tar groupings compared;

(iii) Details of which factors the estimate was adjusted for, with separate columns indicating whether the estimate was adjusted for age, cigarettes/day (or pack-years) or for other factors, as well as which these other factors were;

(iv) The gender the estimate relates to;

(v) The number of lung cancer cases considered in the analysis from which the relative risk estimates derived;

(vi) The relative risk estimates themselves with their 95% lower and upper confidence limits. Note that confidence limits were not available for some of the relative risk estimates; and

(vii) Whether the estimates are to be excluded from all meta-analyses (indicated by "yes") or only from some meta-analyses (indicated by "some"). Estimates are excluded from all analyses if they have no confidence interval or if alternative probably more appropriate estimates from the same study are already available in Table 1. Estimates are excluded from some analyses to avoid probable double-counting of cases which appeared in more than one study. Fuller details of these "overlaps" between studies are given in section 4.

(viii) Additional information (under “notes) including whether the relative risks and/or confidence limits were actually given in the paper or had to be calculated based on data provided.

Table 2 gives additional data for the same estimates. It shows the detailed source of the data, the country in which the study was carried out, the year the study started and finished, the study type (prospective or case-control) and the type of controls used (healthy, diseased/dead or both types). It also shows whether proxy respondents were involved in the study, whether the lung cancer cases had or had not all been confirmed histologically and whether the population considered in the analysis was restricted to cigarette only smokers or included smokers of pipes and/or cigars. Note that estimates for women are assumed to be of cigarette only smokers.

3. Meta-analyses

Of the 101 estimates for the filter/plain comparison, 85 are less than 1.00, with 46 statistically significant reductions at $p<0.05$. None of the 16 estimates greater than 1.00 are statistically significant. Of the 13 estimates from the newly added (recently published) studies, only one is greater than 1.00 and 9 are statistically significantly decreased.

Of the 43 estimates for the low tar/high tar comparison, 38 are less than 1.00, 1 is equal to 1.00 and 4 are greater than 1.00. 18 of the decreases and none of the increases are statistically significant at $p<0.05$.

To assess the apparent advantage of filter and low tar more quantitatively, Table 3 shows the results of various meta-analyses.

For each of four classes of relative risk,

(a) filter/plain - unadjusted for cigs/day
(b) filter/plain - adjusted for cigs/day
(c) low tar/high tar - unadjusted for cigs/day
(d) low tar/high tar - adjusted for cigs/day

results are shown

(i) using all available data (i.e. including estimates in Table 1 with exclude = "No" or "Some"),
(ii) excluding overlaps (i.e. only including estimates in Table 1 with exclude = "No"),
(iii) as (i) but only for studies providing both adjusted and unadjusted results, and
(iv) as (ii) but again only for studies providing a corresponding pair of adjusted and unadjusted results.

For the filter/plain comparisons, all the overall relative risk estimates, based on either fixed-effects or random-effects analysis are in the range 0.61-

0.66, and all are highly statistically significant, with the upper 95% confidence limit no greater than 0.70 for the fixed-effects analysis and no greater than 0.76 for the random-effects analyses. All the analyses show highly significant ($p<0.001$) heterogeneity. Exclusions because of some overlap made essentially no difference to the answers. Estimates adjusted for cigs/day, compared with unadjusted estimates, were very similar in the fixed-effects analyses but slightly greater, by about 0.05, in the random-effects analyses. The newly added studies made little difference to the estimates of relative risk but merely reduced the width of the confidence intervals slightly.

For the low tar/high tar comparisons, all the meta-analysis estimates shown were statistically significantly below 1.00, but there was more variation between the estimates than was the case for the filter/plain comparison. For the estimates unadjusted for cigs/day, there was highly significant heterogeneity ($p<0.001$) and the random-effects estimates were substantially lower than the corresponding fixed-effects estimates, by 0.07 to 0.10 depending on the exclusions made. For the estimates adjusted for cigs/day, the heterogeneity was less marked and the random-effects estimates were only slightly lower than the fixed-effects estimates. The effect of excluding the overlaps was generally to increase the relative risk estimates by about 0.05. The effect of restricting attention to estimates where there was a corresponding unadjusted and adjusted estimate was generally to decrease the relative risk estimate by a small amount. Comparing adjusted with corresponding unadjusted estimates, the adjusted estimates tended to be higher, slightly for the fixed-effects estimates and more so for the random-effects estimates. This was mainly due to the data for the Speizer study where the unadjusted relative risk was 0.50 (95% CI 0.36-0.67) and the adjusted relative risk was 1.00 (95% CI 0.71-1.43). This large difference seems implausible, as discussed in the earlier review[1]. The addition of the studies new to this version of the report had the effect of reducing the estimates slightly.

4. Heterogeneity analyses

4.1 Introduction

The intention was to investigate to what extent the significant heterogeneity noted in many of the analyses in Table 3 could be explained by characteristics of the relative risk estimates. As the heterogeneity analyses are quite lengthy, and one wished to look at filter/plain unadjusted, filter/plain adjusted, low tar/high tar unadjusted and low tar/high tar adjusted data separately, attention was limited to only one set of exclusions. It was decided to consider the set excluding overlaps. Exclusions on the basis of whether or not there were paired unadjusted and adjusted estimates available in the study seemed irrelevant to heterogeneity analyses. Leaving in partially overlapping studies would have led to some double counting, which may have affected the heterogeneity analyses.

Before going on to the results of the analyses, it is relevant to note the details of the partial overlaps. The studies excluded, with the study of which they seem to be a subset shown in brackets, are as follows: Hawthorne (Lubin and Tang), Benhamou (Lubin), Vutuc (Lubin), AHF2 (AHF3), Agudo (Simonato), Higenbottam (Tang), Gillis (Lubin) and Kreuzer (Simonato). There remains very slight overlap of the Lubin and Tang estimates but excluding either would lose substantial data.

4.2 Univariate heterogeneity analyses for the filter/plain comparison

Table 4 presents the results of analyses comparing filter/plain relative risks by level of various factors, considered one at a time. The table presents the estimated relative risks (and 95% confidence limits) by level of the 13 different factors considered, as well as the coded significance of each relative risk (p_1) and of the heterogeneity between risks (p_2). The estimates within level of each factor are based on fixed-effects meta-analysis.

For the filter/plain relative risks, estimates for any level of any factor were always below 1.0, and, except for studies with 100 or fewer lung cancers, significant at $p<0.001$. A number of the analyses showed evidence of

significant heterogeneity by level. Consistent patterns for data unadjusted or adjusted for cigs/day were the tendency for risk to:

(i) vary by study size (marginally significant for unadjusted for cigs/day, $p<0.001$ adjusted for cigs/day) with estimates tending to be highest (i.e. a smaller decrease associated with filter use) in the smallest studies, involving 100 cases or fewer;

(ii) be significantly ($p<0.001$) lower in estimates based on full histological confirmation;

(iii) be lower in case-control than prospective studies, more clearly for adjusted data ($p<0.001$) than for unadjusted data ($p<0.05$);

(iv) be lower in studies not restricted to cigarette only smokers ($p<0.01$ and $p<0.05$ for unadjusted and adjusted estimates respectively) and

(v) vary by type of filter/plain comparison ($p<0.01$). Here the estimates were split into three groups, with those based on a filter/plain split at a single time point going into one group, those based on a filter only/plain only split from a lifetime history going into a second group, and others going into a third group. Relative risks tended to be low for the third group and high for the first group, though the pattern was not completely clear.

The analysis of studies using proxies or not showed significant heterogeneity for both unadjusted and adjusted analyses but the direction of effect was different for the unadjusted and adjusted data.

For some other variables (gender, year of start, year of finish, type of controls), heterogeneity significant at least at the 95% level was seen, but not for both unadjusted and adjusted analyses.

The advantage to filter cigarettes was clearly and quite similarly evident in studies of males and of females, in studies conducted in different continents, and in studies conducted in different time periods. It was evident regardless of whether adjustment was carried out or not for age, cigs/day smoked and other factors.

4.3 Further heterogeneity analyses for the filter-plain comparison

Before going into the details of the multivariate analyses it is useful to consider the effect of individual study estimates on the overall heterogeneity of the data.

For the 46 filter/plain relative risk estimates unadjusted for cigs/day (and excluding overlaps), the overall fixed-effects estimate of 0.63 (0.60-0.67) showed highly significant heterogeneity (χ^2 = 164.99 on 45 d.f., p<0.001). Major contributors to the heterogeneity were the following estimates:

Ref	Study	Sex	RR (95% CI)	Q	p
4	Migrant	M	1.01 (0.69-1.48)	5.78	<0.05
46	Sidney	M	1.15 (0.68-1.94)	5.00	<0.05
53	Khuder	M	0.46 (0.36-0.59)	6.38	<0.01
56	Pezzotto	M	0.23 (0.16-0.34)	27.68	<0.001
60	Matos	M	1.49 (0.86-2.57)	9.41	<0.01
62	Jöckel	M	0.31 (0.17-0.56)	5.50	<0.05
65	Dean	M	0.32 (0.19-0.54)	6.54	<0.05
66	Dean	F	0.31 (0.16-0.62)	4.26	<0.05
71	Texas	F	1.34 (0.80-2.23)	8.24	<0.01
79	Wicken	F	3.12 (0.65-15.00)	3.97	<0.05
82	Choi	M	0.06 (0.01-0.30)	7.37	<0.01
86	Simonato	M	0.92 (0.76-1.10)	15.78	<0.001
141	Brooks	M+F	0.34 (0.27-0.42)	30.33	<0.001

Q contribution to chisquared for overall heterogeneity; p values assume this is distributed as chisquared on 1 d.f.

Some studies, particularly Pezzotto, give relative risks quite far from the overall mean. However, the overall heterogeneity is too large to be explained by any single estimate.

For the 33 filter/plain relative risk estimates adjusted for cigs/day (and excluding overlaps), the overall fixed-effects estimate of 0.65 (0.61-0.70) also showed highly significant heterogeneity (χ^2 = 95.98 on 32 d.f., p<0.001). Major contributors to the heterogeneity are shown in the table below. Again no single estimate can explain the overall observed heterogeneity.

Ref	Study	Sex	RR (95% CI)	Q	p
6	Migrant	M	1.17 (0.79-1.72)	8.68	<0.01
12	Tang	M	0.94 (0.75-1.18)	10.01	<0.01
25	Lubin	M	0.48 (0.40-0.56)	12.74	<0.001
40	Alderson	M	1.48 (0.81-2.69)	7.17	<0.01
44	AHF3	M	0.92 (0.65-1.29)	3.88	<0.05
55	Armadans-Gil	M	0.41 (0.30-0.70)	4.61	<0.05
57	Pezzotto	M	0.29 (0.20-0.42)	18.33	<0.001
67	Dean	M	0.35 (0.21-0.59)	5.57	<0.05
68	Dean	F	0.32 (0.16-0.64)	4.05	<0.05

Q contribution to chisquared for overall heterogeneity; p values assume this is distributed as chisquared on 1 d.f.

Note that the great majority of "outlying" estimates for both unadjusted and adjusted data are in males. This reflects partly the greater number of estimates for males, and also the fact that for a given study confidence intervals tend to be wider for females than for males.

To investigate whether the heterogeneity can be explained by a systematic contribution of one or more of the study characteristics rather than by unusual results in individual studies, multivariate analysis was carried out by linear regression of the log of the relative risk weighted on the inverse of the variance of the individual estimate.

Table 5 summarizes the results of attempts to reach a "best" model for data adjusted for cigs/day. The table shows the deviance (approximate chisquared) for various fitted models that include one or more of the 13 factors considered in Table 4. The procedure is as follows:

(i) Start with the null model (common mean) with deviance of 95.98 on 32 d.f.

(ii) Then try in turn all the single-factor models by including one only of the factors considered. This gives the deviances under the column headed "Individual". The corresponding degrees of freedom are 32 minus the degrees of freedom (1 or 2) for the factor included.

(iii) Choose the factor which causes the most significant drop in deviance. Here this is STYP = Study type (case control or prospective).

(iv) Having chosen this "best" one-factor model, choose the factor which now causes the most significant drop in deviance when added to the model - see the columns headed "Top down models". Continue adding factors as long as a drop in deviance significant at least at $p<0.1$ can be achieved, and stop when it cannot. The column headed "Sequence" identifies the order in which the factors were selected for inclusion in the model. The "Deviance" column shows the deviance for the model which includes the corresponding factor and all previously-added factors. The P column shows the significance of the difference made to the models by adding the factor. In all, eight factors were chosen giving a final deviance of 27.40.

(v) The columns under "Exclude a factor" show the effect of excluding any one of the factors in the final "top down" model. Note that, once the other factors are included, Study type no longer makes a significant contribution to the model. We should therefore consider the final model to be the "topdown" model with study type excluded.

(vi) The columns under "Exclude a factor from the model without study type" relate to the top down model modified to exclude study type, i.e. our final model. They show the effect of excluding each of the remaining factors in turn.

The final model, which had a deviance of 28.01 on 22 d.f. (NS), was as follows:

Parameter	Level	Estimate	S.E.	exp (estimate)
Mean		-1.22	0.27	0.295
Full histological conf. (2)	No	+0.33	0.10	1.391
Used proxies (2)	No	+0.47	0.24	1.600
Year started study (2)	1971-1980	+0.01	0.12	1.010
Year started study (3)	1981+	-0.31	0.11	0.733
Year finished study (2)	1981-1990	+0.33	0.12	1.391
Year finished study (3)	1991+	+0.46	0.13	1.584
Standardised for age (2)	No	+0.48	0.13	1.616
No of lung cancers (2)	101-300	-0.20	0.16	0.819
No of lung cancers (3)	301+	+0.07	0.13	1.073
Sex studied (2)	Females or both	-0.18	0.10	0.835

The "mean" corresponds to the estimate for a study with the first level of all the characteristics included in the model. The "exp (estimate)" for a given level of a characteristic is an estimate of the relative risk compared with the first level of that characteristic.

From this analysis, and additional output not shown, one can conclude the following:

(i) the final model fits - the deviance is not significant. No study gives an extreme estimate except, perhaps, the Pezzotto study. [Three estimates gave residuals of over 2 standard errors; refs 57 (Pezzotto, males, -2.3 S.E.s), 2 (Bross, -2.15 S.E.s) and 61 (Matos, 2.05 S.E.s)];

(ii) deleting any of the factors included made the model fit worse, although Study sex and Proxy use were of marginal significance;

(iii) the most significant factors, at $p<0.001$, were Full histological confirmation and Standardised for age. The filter/plain relative risk estimates tended to be lower (i.e. showed more advantage to filter) in studies that were standardised for age and in which 100% histological confirmation of diagnosis was required;

(iv) other less significant factors included in the model were Year started study ($p<0.01$), Year finished study ($p<0.01$) and Number of lung cancers ($p<0.05$). More advantage to filter cigarette use tended to be seen in studies that started later, and, paradoxically, also in studies finishing before 1980. A greater advantage to filter cigarette use was seen in studies with 101-300 lung cancers than in those with either more or fewer. For all of these three factors the "significant" results may not actually imply any meaningful relationship.

Table 6 similarly shows the results of multivariate weighted regression analyses of data unadjusted for cigs/day. Here, the final model fitted, which had a deviance of 109.95 on 40 d.f. ($p<0.001$) was as follows:

Parameter	Level	Estimate	S.E.	exp (estimate)
Mean		-0.79	0.10	0.454
Full histological conf. (2)	No	+0.39	0.06	1.477
Year started study (2)	1971-1980	+0.28	0.09	1.323
Year started study (3)	1981+	+0.05	0.08	1.051
Sex studied (2)	Females or both	-0.24	0.07	0.787
Type of controls used (2)	Included diseased	+0.13	0.07	1.139

From these results we can see that the model does not fit – the deviance is highly significant ($p<0.001$) and a number of study estimates show significant deviations from the value predicted by the model. Two estimates gave residuals over 3 standard errors - refs 56 (Pezzotto, -4.43 S.E.s) and 15 (AHF2, males, 3.01 S.E.s); and six estimates gave residuals of over 2 standard errors - refs 82 (Choi, -2.98 S.E.s), 65 (Dean, males, -2.78 S.E.s), 140 (Rachtan, 2.54 S.E.s), 64 (Doll, -2.23 S.E.s), 60 (Matos, 2.24 S.E.s) and 79 (Wicken, females, +2.06 S.E.s) and a number of other studies gave residuals of almost 2 S.E.s;

4.4 Univariate heterogeneity analyses for the low tar/high tar comparison

Table 7 presents the results of heterogeneity analyses for the low tar/high tar comparison considering one factor at a time. These correspond to Table 4 for the filter/plain analyses.

The data adjusted for cigs/day, which gave an overall fixed effects estimate of 0.76 (0.69-0.83), showed no significant heterogeneity (χ^2 = 25.22 on 18 d.f.) based on all 19 estimates, though study reference 112 showed marginal significance (Sidney, females, RR 1.49 (0.76-2.94), Q 3.85, p<0.05). The data also showed no evidence of heterogeneity by the individual factors.

The data unadjusted for cigs/day, which gave an overall fixed effect estimate of 0.74 (0.66-0.83), did show significant heterogeneity (χ^2 = 54.94 on 13 d.f.) based on all the 14 available estimates. This heterogeneity predominantly arose because of six estimates, listed below:

Ref	Study	Gender	RR (CI)	Q	p
97	Lubin	males	0.98 (0.75-1.28)	4.08	<0.05
107	AHF	males	1.32 (0.89-1.95)	8.20	<0.01
110	Sidney	females	1.49 (0.78-2.87)	4.36	<0.05
115	Kaufman	males	0.11 (0.04-0.30)	13.83	<0.001
116	Kaufman	females	0.17 (0.07-0.43)	10.16	<0.01
119	Speizer	females	0.50 (0.36-0.67)	6.30	<0.05

Q contribution to chisquared for overall heterogeneity; p values assume this is distributed as chisquared on 1 d.f.

Rejecting as many as the three most extreme estimates out of 11 as "outliers" does not seem appropriate. However the existence of these very low and very high estimates does mean that interpretation of the results in Table 7 is not straightforward.

Significant variation by level of various factors – seen most notably for the factors year of start, number of lung cancers and full histological confirmation – depends mainly on the six heterogeneous estimates noted

above. Thus, for example, the highly significant ($p<0.001$) heterogeneity by year of start was because the estimate for 1971-80 included the high Lubin, AHF and Sidney estimates while the estimate for 1981+ included the two low Kaufman estimates and the low Speizer estimate.

There is really too much variation in the low tar/high tar estimates unadjusted for cigs/day and too few estimates available to make any reliable inferences. There is certainly no point in conducting multivariate heterogeneity analysis here (and also for the estimates adjusted for cigs/day where there is no real heterogeneity to explain).

5. Further comparisons of relative risks unadjusted and adjusted for cigs/day

In Table 3 meta-analysis estimates have been calculated separately for relative risk estimates that are unadjusted and adjusted for cigs/day. In order to provide greater comparability, some estimates have been based on data restricted to studies which presented both unadjusted and adjusted estimates.

In principle, if one has pairs of unadjusted and adjusted estimates from a number of studies, a more powerful test of the effect of adjustment can be obtained by combining over studies individual estimates of the effect of adjustment. Table 8 presents results of analyses meta-analysing the ratio of adjusted to unadjusted relative risks. Results are presented:

(a) separately for the filter/plain and the low tar/high tar relative risk,

(b) separately for "all available data" and "excluding overlaps", and

(c) of unweighted analyses and of analyses weighted on the inverse of the variance of the adjusted relative risk.

It can be seen that, for both the filter/plain and the low tar/high tar comparison, adjustment for amount smoked more often increased than decreased the relative risk estimate. However, this tendency was not marked. Using unweighted analysis, the overall estimate of the effect of adjustment was to increase the relative risk by about 10%, while using weighted analysis eliminated this increase. The difference in conclusion between the unweighted and weighted analysis arises because increases following adjustment were predominantly seen in small studies, with some large studies showing decreases.

Whatever the relative merits of unweighted and weighted analysis, it is clear that any effect of adjustment for amount smoked is quite small. It is certainly much less than the magnitude of the advantage seen in Table 3 for filter and lower tar cigarettes. This tends to refute the view of Monograph 13[2] that the apparent advantage of filters and tar reduction is an artefact caused by inappropriate adjustment for amount smoked.

TABLE 1: Relative risks and confidence intervals for filter/plain and low tar/high tar comparison

RR no	Study	Reference group	Comparison group	Adj. for age	Adj. for cigs/day	Adjusted for other factors	Gender	Lung cancer cases	Relative risk	Lower 95% limit	Upper 95% limit	Exclude	Notes
ESTIMATES FOR FILTER/PLAIN COMPARISON													
1	Bross	Current plain	Current filter	no	no	no	male	265	0.59	0.40	0.88	no	r,c
2	Bross	Current plain	Current filter	no	yes	no	male	265	0.57	0.39	0.85	no	r,c
3	Hawthorne	Current plain	Current filter	yes	yes	study	male	88	0.83	0.53	1.31	some	r.c
4	Migrant	Current plain only	Current filter only	no	no	no	male	99	1.01	0.69	1.48	no	r,c
5	Migrant	Current plain only	Current filter only	no	no	no	female	21	0.92	0.39	2.16	no	r,c
6	Migrant	Current plain only	Current filter only	no	yes	no	male	99	1.17	0.79	1.72	no	r,c
7	Migrant	Current plain only	Current filter only	no	yes	no	female	21	1.01	0.43	2.37	no	r,c
8	Engeland	Current plain only	Current filter only	no	no	no	male	45	0.53	0.24	1.18	no	r,c
9	Engeland	Current plain only	Current filter only	no	no	no	female	24	0.58	0.25	1.34	no	r,c
10	Engeland	Current plain only	Current filter only	yes	yes	ages,occ,pipe,urb	male	45	0.67	0.30	1.43	no	r,c
11	Engeland	Current plain only	Current filter only	yes	yes	ages,pipe,urb	female	24	0.91	0.40	2.00	no	r,c
12	Tang	Current plain	Current filter	yes	yes	study	male	366	0.94	0.75	1.18	no	r,c
13	AHF 1	Current plain	Filter for 10+ yrs	no	no	no	male	226	0.52	0.35	0.76	no	r,c
14	AHF 1	Current plain	Filter for 10+ yrs	no	yes	no	male	226	0.51	0.34	0.76	no	r,c
15	AHF 2	Current plain	Filter for 10+ yrs	no	no	no	male	293	0.76	0.59	0.98	some	r,c,k1
16	AHF 2	Current plain	Filter for 10+ yrs	no	no	no	female	63	0.74	0.40	1.40	some	r,c,k1
17	AHF 2	Current plain	Filter for 10+ yrs	no	yes	no	male	293	0.77	0.59	0.99	some	r,c,k1
18	AHF 2	Current plain	Filter for 10+ yrs	no	yes	no	female	63	0.73	0.39	1.39	some	r,c,k1
19	AHF 3	Lifetime only plain	Lifetime only filter	yes	yes	educ,inh	male	2085	0.77	0.46	1.30	yes	r,c
20	AHF 3	Lifetime only plain	Lifetime only filter	yes	yes	educ,inh	female	1012	0.87	0.56	1.33	yes	r,c
21	Rimington	Current plain	Current filter	yes	no	no	male	104	0.65	0.44	0.96	no	r,c
22	Rimington	Current plain	Current filter	yes	yes	no	male	104	0.62	0.42	0.91	no	r,c
23	Lubin	Lifetime only plain	Lifetime only filter	no	no	no	male	6626	0.66	0.57	0.76	no	r,c
24	Lubin	Lifetime only plain	Lifetime only filter	no	no	no	female	551	0.52	0.34	0.82	no	r,c
25	Lubin	Lifetime only plain	Lifetime only filter	yes	yes	yrsq	male	6626	0.48	0.40	0.56	no	r,c
26	Lubin	Lifetime only plain	Lifetime only filter	yes	yes	yrsq	female	551	0.43	0.22	0.85	no	r,c
27	Texas	Plain undefined	Filter undefined	no	no	no	male	475	0.92	N.A.	N.A.	yes	r
28	Texas	Plain undefined	Filter undefined	no	no	no	female	460	1.17	N.A.	N.A.	yes	r
29	Benhamou	Lifetime only plain	Lifetime only filter	yes	no	no	male	1114	0.38	0.24	0.62	some	r,c
30	Benhamou	Lifetime only plain	Lifetime only filter	yes	yes	cur,dur,hos,inh,ivr,tar,ttyp	male	1114	0.63	0.35	1.10	some	r,c
31	Benhamou	Lifetime only plain	Lifetime mainly filter	yes	no	hos,ivr	female	46	0.16	0.04	0.61	some	r,c
32	Benhamou	Lifetime only plain	Lifetime mainly filter	yes	yes	ages,dur,hos,inh,ivr	female	46	0.28	0.05	1.47	some	r,c
33	Berrino	Lifetime only plain	Lifetime only filter	yes	yes	curr,res,ttyp	male	1101	0.91	N.A.	N.A.	yes	r
34	Lange	Current plain	Current filter	yes	no	no	male	90	0.82	0.47	1.44	no	g
35	Lange	Current plain	Current filter	yes	no	no	female	39	0.61	0.21	1.79	no	g
36	Lange	Current plain	Current filter	yes	as pyrs	no	male	90	0.90	0.60	1.40	no	r,c
37	Lange	Current plain	Current filter	yes	as pyrs	no	female	39	0.70	0.40	1.40	no	r,c

TABLE 1: Relative risks and confidence intervals for filter/plain and low tar/high tar comparison (continued)

RR no	Study	Reference group	Comparison group	Adj. for age	Adj. for cigs/day	Adjusted for other factors	Gender	Lung cancer cases	Relative risk	Lower 95% limit	Upper 95% limit	Exclude	Notes
38	Alderson	Lifetime only plain	Lifetime only filter	no	no	no	male	312	1.01	0.56	1.85	no	r,c
39	Alderson	Lifetime only plain	Lifetime only filter	no	no	no	female	410	0.80	0.49	1.30	no	r,c
40	Alderson	Lifetime only plain	Lifetime only filter	yes	yes	no	male	312	1.48	0.81	2.69	no	r,c
41	Alderson	Lifetime only plain	Lifetime only filter	yes	yes	no	female	410	0.85	0.52	1.38	no	r,c
42	AHF 3	Lifetime only plain	Lifetime only filter	no	no	no	male	1442	0.69	0.51	0.94	no	r,c
43	AHF 3	Lifetime only plain	Lifetime only filter	no	no	no	female	850	0.63	0.39	1.01	no	r,c
44	AHF 3	Lifetime only plain	Lifetime only filter	yes	yes	educ	male	1442	0.92	0.65	1.29	no	r,c
45	AHF 3	Lifetime only plain	Lifetime only filter	yes	yes	educ	female	850	0.68	0.39	1.19	no	r,c
46	Sidney	Usual brand plain	Usual brand filter	yes	no	no	male	98	1.15	0.68	1.94	no	r,c
47	Sidney	Usual brand plain	Usual brand filter	yes	no	no	female	83	0.44	0.22	0.88	no	r,c
48	Sidney	Usual brand plain	Usual brand filter	yes	yes	dur,educ,race	male	98	1.03	0.61	1.75	no	g
49	Sidney	Usual brand plain	Usual brand filter	yes	yes	dur,educ,race	female	83	0.65	0.32	1.31	no	g
50	Sidney	Lifetime only plain	Filter for 20+ yrs	yes	yes	dur,educ,race	male	93	1.04	0.58	1.87	yes	g
51	Sidney	Lifetime only plain	Filter for 20+ yrs	yes	yes	dur,educ,race	female	73	0.36	0.18	0.75	yes	g
52	Pathak	Lifetime only plain	Lifetime only filter	yes	yes	dur,race,sex	both	320	0.80	N.A.	N.A.	yes	g
53	Khuder	Lifetime only plain	Lifetime ever filter	no	no	no	male	457	0.46	0.36	0.59	no	r,c
54	Armadans-Gil	Ever plain last 20 yrs	Only filter for 20+ yrs	no	no	no	male	267	0.43	0.27	0.67	no	r,c
55	Armadans-Gil	Ever plain last 20 yrs	Only filter for 20+ yrs	yes	as pyrs	dur,ttyp,ses	male	267	0.41	0.30	0.70	no	g
56	Pezzotto	Lifetime plain or both	Lifetime ever filter	yes	no	hos	male	211	0.23	0.16	0.34	no	r,c
57	Pezzotto	Lifetime plain or both	Lifetime ever filter	yes	yes	dur,hos	male	211	0.29	0.20	0.42	no	r,c
58	De Stefani 1	Lifetime ever plain	Lifetime only filter	yes	no	educ,res,urb	male	470	0.72	0.54	0.96	no	r,c
59	Agudo	Lifetime ever plain	Lifetime only filter	yes	no	hosp,res	female	22	0.22	0.04	1.27	some	r,c
60	Matos	Lifetime mainly plain	Lifetime mainly filter	no	no	no	male	185	1.49	0.86	2.57	no	r,c
61	Matos	Lifetime mainly plain	Lifetime mainly filter	yes	yes	hosp,yrsq	male	185	1.25	0.67	2.50	no	r,c
62	Jockel	Ever plain last 20 yrs	Only filter for 20+ yrs	no	no	no	male	137	0.31	0.17	0.56	no	r,c
63	Jockel	Ever plain last 20 yrs	Only filter for 20+ yrs	yes	as pyrs	yrsq	male	137	0.41	0.21	0.81	no	r,c
64	Doll	Lifetime only plain	Lifetime ever filter	no	no	no	male	504	0.18	0.05	0.63	no	r,c
65	Dean	Plain 1954 and 1969	Filter 1954 and 1969	yes	no	no	male	262	0.32	0.19	0.54	no	c
66	Dean	Plain 1954 and 1969	Filter 1954 and 1969	yes	no	no	female	81	0.31	0.16	0.62	no	c
67	Dean	Plain 1954 and 1969	Filter 1954 and 1969	yes	yes	no	male	262	0.35	0.21	0.59	no	c
68	Dean	Plain 1954 and 1969	Filter 1954 and 1969	yes	yes	no	female	81	0.32	0.16	0.64	no	c
69	Hirayama	Current plain	Current filter	yes **	yes **	no	N.K.	N.K.	0.51	N.A.	N.A.	yes	g
70	Ockene	Current plain	Current filter	yes	yes	ages,alc,bp,chol,nic,tar,thio	male	106	0.53	0.24	1.17	no	r,c
71	Texas	Lifetime only plain	Lifetime only filter	no	no	no	female	176	1.34	0.80	2.23	no	r,c
72	Correa	Current plain	Current filter	yes	no	sex	both	1338	0.55	0.35	0.85	no	r,c
73	CPS II	Lifetime 60+% plain	Lifetime only filter	yes	yes	inh	female	570	0.66	0.57	0.78	yes	r,c
74	CPS II	Lifetime mainly plain	Lifetime mainly filter	yes	no	no	male	N.K.	0.47	N.A.	N.A.	yes	r
75	CPS II	Lifetime mainly plain	Lifetime mainly filter	yes	no	no	female	N.K.	0.51	N.A.	N.A.	yes	r

TABLE 1: Relative risks and confidence intervals for filter/plain and low tar/high tar comparison
(continued/2)

RR no	Study	Reference group	Comparison group	Adj. for age	Adj. for cigs/day	Adjusted for other factors	Gender	Lung cancer cases	Relative risk	Lower 95% limit	Upper 95% limit	Exclude	Notes
76	De Stefani 2	Current plain	Current filter	yes	no	bmi,edu,fhis,res,sex,urb	both	300	0.73	0.51	1.05	no	r,c
77	Zemla	Current plain	Current filter	no	no	no	male	210	0.97	N.A.	N.A.	yes	r
78	Wicken	Current plain	Current filter	no	no	no	male	678	0.97	0.50	1.86	no	r,c
79	Wicken	Current plain	Current filter	no	no	no	female	62	3.12	0.65	15.00	no	r,c
80	Wakai	Current plain	Current filter	no	no	no	male	179	1.16	0.38	3.51	no	r,c
81	Wakai	Current plain	Current filter	yes	yes	ages,frac,inh	male	179	1.02	0.31	3.33	no	r,c
82	Choi	Current plain only	Current filter only	no	no	no	male	267	0.06	0.01	0.30	no	r,c
83	Segi	Current plain	Current filter	yes	no	no	male	240	0.62	0.45	0.85	no	r,c
125	Segi	Current plain	Current filter	yes	yes	ages	male	240	0.62	0.45	0.85	no	r,c
84	Sobue	Current plain	Current filter	no	no	no	male	609	0.55	0.34	0.90	no	r,c
85	Sobue	Current plain	Current filter	yes	yes	dur,frac,inh	male	609	0.67	0.38	1.11	no	r,c
86	Simonato	Lifetime only plain	Lifetime only filter	yes	no	educ,study	male	3562	0.92	0.76	1.10	no	r,c
87	Simonato	Lifetime only plain	Lifetime only filter	yes	no	educ,study	female	860	0.49	0.32	0.75	no	r,c
126	Blizzard	Lifetime plain/mixed	Lifetime only filter	no	no	no	male	92	0.56	0.32	0.98	no	r,c
127	Blizzard	Lifetime plain/mixed	Lifetime only filter	no	no	no	female	56	1.05	0.46	2.43	no	r,c
132	Marugama	Lifetime plain/mixed	Lifetime only filter	yes	no	reg	male	356	0.70	0.40	1.15	no	g
133	Marugama	Lifetime plain/mixed	Lifetime only filter	yes	yes	reg	male	356	0.70	0.41	1.21	no	g
136	CPS II	Current plain	Current filter	yes	no	ages,educ,race,mar,occ,nut,asb	male	1194	0.71	0.60	0.84	no	r,c
137	CPS II	Current plain	Current filter	yes	yes	ages,educ,race,mar,occ,nut,asb	male	1194	0.71	0.59	0.84	no	r,c
138	CPS II	Current plain	Current filter	yes	no	ages,educ,race,mar,occ,nut,asb	female	849	0.59	0.46	0.77	no	r,c
139	CPS II	Current plain	Current filter	yes	yes	ages,educ,race,mar,occ,nut,asb	female	849	0.60	0.46	0.77	no	r,c
140	Rachtan	Current plain	Current filter	yes	no	no	female	188	0.95	0.46	1.93	no	r,c
141	Brooks	Current plain	Current filter	no	no	no	both	643	0.34	0.27	0.42	no	r,c
142	Kreuzer	Lifetime plain/mixed	Lifetime only filter	no	no	no	male	3642	0.74	0.65	0.83	some	r,c
143	Kreuzer	Lifetime plain/mixed	Lifetime only filter	no	no	no	female	614	0.61	0.46	0.81	some	r,c
144	Agudo	Lifetime plain/mixed	Lifetime only filter	yes	no	reg	female	1301	0.46	0.36	0.58	some	r,c
ESTIMATES FOR TAR COMPARISON													
88	CPS I 1960-66	2.0+ mg nic 25.8+ mg tar	<1.2mg nic, <17.6 mg tar	yes	yes	ages.educ,hchd,hlc,occ,race,urb	male	341	0.83	0.64	1.08	no	r,c
89	CPS I 1960-66	2.0+ mg nic 25.8+ mg tar	<1.2mg nic, <17.6 mg tar	yes	yes	ages.educ,hchd,hlc,race,urb	female	117	0.57	0.36	0.91	no	r,c
90	CPS I 1966-72	High tar/nicotine *	Low tar/nicotine *	yes	yes	ages.educ,hchd,hlc,occ,race,urb	male	245	0.79	0.58	1.08	no	r,c
91	CPS I 1966-72	High tar/nicotine *	Low tar/nicotine *	yes	yes	ages.educ,hchd,hlc,race,urb	female	137	0.62	0.41	0.94	no	r,c
92	Higenbottam	>32 mg tar	<24 mg tar	yes	yes	inh,occ	male	143	0.56	0.36	0.86	some	r,c
93	MRFIT	20+ mg tar	<15 mg tar	no	no	no	male	95	0.73	0.38	1.41	no	r,c
94	MRFIT	20+ mg tar	<15 mg tar	yes	yes	bp,chol	male	95	0.88	0.52	1.49	no	r,c
95	Lubin	High tar (mean 29.8 mg)	Low tar (mean 15.6 mg)	no	no	curr,dur	male	2650	0.98	0.75	1.28	no	r,c
96	Lubin	High tar (mean 25.2 mg)	Low tar (mean 15.6 mg)	no	no	curr,dur	female	313	0.66	0.37	1.16	no	r,c

TABLE 1: Relative risks and confidence intervals for filter/plain and low tar/high tar comparison
(continued/3)

RR no	Study	Reference group	Comparison group	Adj. for age	Adj. for cigs/day	Adjusted for other factors	Gender	Lung cancer cases	Relative risk	Lower 95% limit	Upper 95% limit	Exclude	Notes
97	Lubin	High tar (mean 29.8 mg)	Low tar (mean 15.6 mg)	no	yes	curr,dur	male	2650	0.71	0.55	0.93	no	r,c
98	Lubin	High tar (mean 25.2 mg)	Low tar (mean 15.6 mg)	no	yes	curr,dur	female	313	0.67	0.38	1.18	no	r,c
99	Benhamou	>75% 30+ mg tar	Light imported (tar NK)	no	no	no	male	1101	0.26	0.14	0.48	some	r,c
100	Benhamou	>75% 30+ mg tar	Light imported (tar NK)	yes	yes	curr,dur,fp,inh,ttyp	male	1101	0.30	0.10	0.91	some	r,c
101	Vutuc	>24 mg tar	<15 mg tar	no	no	no	female	188	0.32	0.09	1.17	some	r,c
102	Vutuc	>24 mg tar	<15 mg tar	yes	yes	dur	female	188	0.29	0.09	0.95	some	r,c
103	Vutuc	>24 mg tar	<15 mg tar	no	no	no	male	248	0.32	0.09	1.08	some	r,c
104	Vutuc	>24 mg tar	<15 mg tar	yes	yes	dur	male	248	0.30	0.11	0.81	some	r,c
105	Gillis	>22 mg tar	22- mg tar	no	no	no	male	490	0.73	0.52	1.01	some	r,c
106	Gillis	>22 mg tar	22- mg tar	no	yes	no	male	490	0.74	0.53	1.03	some	r,c
107	AHF	15+ mg tar	<10 mg tar	no	no	no	male	682	1.32	0.89	1.95	no	r,c
108	AHF	15+ mg tar	<10 mg tar	no	no	no	female	492	0.93	0.61	1.42	no	r,c
109	Sidney	>18 mg tar	<11 mg tar	yes	no	no	male	82	0.92	0.48	1.73	no	r,c
110	Sidney	>18 mg tar	<11 mg tar	yes	no	no	female	76	1.49	0.78	2.87	no	r,c
111	Sidney	>18 mg tar	<11 mg tar	yes	yes	dur,educ,race	male	82	0.79	0.41	0.50	no	r,c
112	Sidney	>18 mg tar	<11 mg tar	yes	yes	dur,educ,race	female	76	1.49	0.76	2.94	no	r,c
113	Wilcox	21.1-28.0 mg tar	14- mg tar	no	no	no	male	373	0.53	0.29	0.97	no	g
114	Wilcox	21.1-28.0 mg tar	14- mg tar	no	yes	no	male	373	0.61	0.32	1.13	no	g
115	Kaufman	29+ mg tar	<22 mg tar	no	no	no	male	119	0.11	0.04	0.30	no	r,c
116	Kaufman	29+ mg tar	<22 mg tar	no	no	no	female	51	0.17	0.07	0.43	no	r,c
117	Kaufman	29+ mg tar	<22 mg tar	yes	yes	ages,educ,race,reg,rel,yrin	male	119	0.25	0.08	0.82	no	r,c
118	Kaufman	29+ mg tar	<22 mg tar	yes	yes	ages,educ,race,reg,rel,yrin	female	51	0.21	0.05	0.93	no	r,c
119	Speizer	Top quartile	Bottom quartile	yes	no	ages	female	593	0.50	0.36	0.67	no	r,c
120	Speizer	Top quartile	Bottom quartile	yes	yes	ages	female	593	1.00	0.71	1.43	no	r,c
121	Alderson	>28 mg tar	<23 mg tar	yes	yes	no	male	299	0.83	0.55	1.24	no	r,c
122	Alderson	>28 mg tar	<23 mg tar	yes	yes	no	female	386	1.12	0.74	1.70	no	r,c
123	Tang	RR for 15mg tar decrease	Not applicable	yes	yes	study	male	366	0.75	0.52	1.09	no	g
124	Tang	RR for 15mg tar decrease	Not applicable	yes	yes	inh	female	1006	0.63	N.A.	N.A.	yes	r,c
128	CPS II	≥22 mg tar	0-7 mg tar	yes	no	ages,educ,race,mar,occ,nut,asb	male	1194	0.81	0.63	1.05	no	r,c
129	CPS II	≥22 mg tar	0-7 mg tar	yes	yes	ages,educ,race,mar,occ,nut,asb	male	1194	0.78	0.60	1.00	no	r,c
130	CPS II	≥22 mg tar	0-7 mg tar	yes	no	ages,educ,race,mar,occ,nut,asb	female	849	0.60	0.44	0.81	no	r,c
131	CPS II	≥22 mg tar	0-7 mg tar	yes	yes	ages,educ,race,mar,occ,nut,asb	female	849	0.56	0.41	0.76	no	r,c
134	SHHS	≥15 mg tar	<10 mg tar	yes	no	sex	both	79	0.31	0.11	0.86	no	r,c
135	SHHS	≥15 mg tar	<10 mg tar	yes	yes	sex,ses,dur,vit	both	79	0.43	0.15	1.23	no	r,c

Footnotes

RR no	Relative risk number
Study	See table 2 for reference and fuller details of study
Reference group	Denominator of relative risk
Comparison group	Numerator of relative risk
Adjusted for	Shows whether the relative risk estimate was adjusted for age, cigarettes/day (or pack-years = pyrs) or other variables with abbreviations used shown below
Lung cancer cases	(Usually) The number considered in the analysis which may have included smoking groups other than those shown
Exclude	Yes = not included in any meta-analyses, because no confidence interval available, alternative estimate from same study available or result implausible Some = excluded from some meta-analyses due to "overlap" - see text
Notes	c = confidence limits estimated from data presented g = data as given in the source shown in Table 2 k1 = data for Kreyberg 1 lung cancer only r = relative risk estimated from data presented
Other abbreviations	N.A. = not available, N.K. = not known
*	For 1960-66 subjects were divided into 3 groups based on the 1959-60 questionnaire (Q) with the high and low definition being as shown in the table with medium intermediate For 1966-72 "high T/N" included subjects in the high group from the 1959-60 Q and in the high or medium groups on the 1965-66 Q,. with "low" including those in the low group from the 1959-60 Q and in either the high or medium group on the 1965-66 Q or in the low group on both the 1961-62 and the 1965-66 Qs.
**	It has been assumed the estimates were adjusted for age and cigs/day but details were not actually given

Abbreviations for adjustment factors

ages	age of starting to smoke
alc	alcohol
asb	asbestos exposure
bmi	body mass index
bp	blood pressure
chol	cholesterol
curr	current or ex smoking
dur	duration of smoking
educ	education
fhis	family history of lung cancer
frac	fraction of cigarette smoked
fp	filter/plain
hchd	history of CHD
hlc	history of lung cancer
hosp	hospital
inh	inhalation
ivr	interviewer
mar	marital status
nic	nicotine
nut	nutrition
occ	occupation
pipe	pipe smoking
race	race
reg	region
rel	religion
res	residence
ses	socioeconomic status
sex	sex (male/female)
tar	tar
thio	thiocyanate
ttyp	tobacco type (light/dark)
urb	urban/rural
vit	vitamins in diet
yrin	year of interview
yrsq	years quit

TABLE 2: Characteristics of studies

RR no	Study	Reference number	Detailed source	Country	Period start	Finish	Study type	Control type	Proxy use	Full hist. conf.	Cigarette only smokers
1	Bross	[3]	Table 1	USA	1960	1966	case-control	diseased/dead	no	no	no
2	Bross	[3]	Table 1	USA	1960	1966	case-control	diseased/dead	no	no	no
3	Hawthorne	[4]	Table 10	UK	1965	1977	prospective	healthy	no	no	no
4	Migrant	[5]	Table 18	UK	1974	1977	prospective	healthy	no	no	yes
5	Migrant	[5]	Table 18	UK	1974	1977	prospective	healthy	no	no	yes
6	Migrant	[5]	Table 18	UK	1974	1977	prospective	healthy	no	no	yes
7	Migrant	[5]	Table 18	UK	1974	1977	prospective	healthy	no	no	yes
8	Engeland	[6]	Table 3	Norway	1964	1993	prospective	healthy	no	no	no
9	Engeland	[6]	Table 4	Norway	1964	1993	prospective	healthy	no	no	yes
10	Engeland	[6]	Table 3	Norway	1964	1993	prospective	healthy	no	no	no
11	Engeland	[6]	Table 4	Norway	1964	1993	prospective	healthy	no	no	yes
12	Tang	[7]	Table 3	UK	1967	1982	prospective	healthy	no	no	yes
13	AHF 1	[8]	Figure 9	USA	1966	1969	case-control	diseased/dead	no	yes	no
14	AHF 1	[8]	Figure 9	USA	1966	1969	case-control	diseased/dead	no	yes	no
15	AHF 2	[9]	Figure 1	USA	1969	1976	case-control	diseased/dead	no	yes	no
16	AHF 2	[9]	Figure 1	USA	1969	1976	case-control	diseased/dead	no	yes	yes
17	AHF 2	[9]	Figure 1	USA	1969	1976	case-control	diseased/dead	no	yes	no
18	AHF 2	[9]	Figure 1	USA	1969	1976	case-control	diseased/dead	no	yes	yes
19	AHF 3	[10]	Table 2	USA	1969	1991	case-control	diseased/dead	no	yes	no
20	AHF 3	[10]	Table 2	USA	1969	1991	case-control	diseased/dead	no	yes	yes
21	Rimington	[11]	Table 1	UK	1970	1976	prospective	healthy	no	no	yes
22	Rimington	[11]	Table 1	UK	1970	1976	prospective	healthy	no	no	yes
23	Lubin	[12]	Table V	Various *	1976	1980	case-control	diseased/dead	no	yes	no
24	Lubin	[12]	Table V	Various *	1976	1980	case-control	diseased/dead	no	yes	yes
25	Lubin	[12]	Table V	Various *	1976	1980	case-control	diseased/dead	no	yes	no
26	Lubin	[12]	Table V	Various *	1976	1980	case-control	diseased/dead	no	yes	yes
27	Texas	[13]	Table 6	USA	1976	1980	case-control	both	yes	no	no
28	Texas	[13]	Table 6	USA	1976	1980	case-control	both	yes	no	yes

TABLE 2: Characteristics of studies (continued)

RR no	Study	Reference number	Detailed source	Country	Period start	Finish	Study type	Control type	Proxy use	Full hist. conf.	Cigarette only smokers
29	Benhamou	[14]	Table 2	France	1976	1980	case-control	diseased/dead	no	yes	no
30	Benhamou	[14]	Table 3	France	1976	1980	case-control	diseased/dead	no	yes	no
31	Benhamou	[15]	Table III	France	1976	1980	case-control	diseased/dead	no	yes	yes
32	Benhamou	[15]	Table IV	France	1976	1980	case-control	diseased/dead	no	yes	yes
33	Berrino	[16]	Table 5	Italy	1977	1980	case-control	diseased/dead	no	yes	no
34	Lange	[17]	Table 3	Denmark	1976	1989	prospective	healthy	no	no	yes
35	Lange	[17]	Table 3	Denmark	1976	1989	prospective	healthy	no	no	yes
36	Lange	[17]	Table 4	Denmark	1976	1989	prospective	healthy	no	no	yes
37	Lange	[17]	Table 4	Denmark	1976	1989	prospective	healthy	no	no	yes
38	Alderson	[18]	Table 5	UK	1977	1982	case-control	diseased/dead	no	no	no
39	Alderson	[18]	Table 5	UK	1977	1982	case-control	diseased/dead	no	no	yes
40	Alderson	[18]	Table 5	UK	1977	1982	case-control	diseased/dead	no	no	no
41	Alderson	[18]	Table 5	UK	1977	1982	case-control	diseased/dead	no	no	yes
42	AHF 3	[19]	Table 3	USA	1977	1995	case-control	diseased/dead	no	yes	no
43	AHF 3	[19]	Table 3	USA	1977	1995	case-control	diseased/dead	no	yes	yes
44	AHF 3	[19]	Table 3	USA	1977	1995	case-control	diseased/dead	no	yes	no
45	AHF 3	[19]	Table 3	USA	1977	1995	case-control	diseased/dead	no	yes	yes
46	Sidney	[20]	Table 3	USA	1979	1987	prospective	healthy	no	no	yes
47	Sidney	[20]	Table 3	USA	1979	1987	prospective	healthy	no	no	yes
48	Sidney	[20]	Table 3	USA	1979	1987	prospective	healthy	no	no	yes
49	Sidney	[20]	Table 3	USA	1979	1987	prospective	healthy	no	no	yes
50	Sidney	[20]	Table 4	USA	1979	1987	prospective	healthy	no	no	yes
51	Sidney	[20]	Table 4	USA	1979	1987	prospective	healthy	no	no	yes
52	Pathak	[21]	Table 7	USA	1980	1982	case-control	healthy	yes	no	no
53	Khuder	[22]	Table 2	USA	1985	1987	case-control	healthy	no	yes	no
54	Armadans-Gil	[23]	Table 4	Spain	1986	1990	case-control	diseased/dead	no	yes	no
55	Armadans-Gil	[23]	Table 4	Spain	1986	1990	case-control	diseased/dead	no	yes	no

TABLE 2: Characteristics of studies
(continued/2)

RR no	Study	Reference number	Detailed source	Country	Period start	Finish	Study type	Control type	Proxy use	Full hist. conf.	Cigarette only smokers
56	Pezzotto	[24]	Table 7	Argentina	1987	1991	case-control	diseased/dead	no	yes	yes
57	Pezzotto	[24]	Table 7	Argentina	1987	1991	case-control	diseased/dead	no	yes	yes
58	DeStefani 1	[25]	Table 2	Uruguay	1988	1994	case-control	diseased/dead	no	yes	no
59	Agudo	[26]	Table III	Spain	1989	1992	case-control	diseased/dead	no	no	yes
60	Matos	[27]	Table 3	Argentina	1994	1996	case-control	diseased/dead	no	no	no
61	Matos	[27]	Table 3	Argentina	1994	1996	case-control	diseased/dead	no	no	no
62	Jockel	[28]	Table 3	Germany	1985	1986	case-control	both	no	yes	no
63	Jockel	[28]	Table 3	Germany	1985	1986	case-control	both	no	yes	no
64	Doll	[29]	Table VIII	UK	1948	1952	case-control	diseased/dead	no	no	no
65	Dean	[30]	Table10S	UK	1963	1972	case-control	healthy	yes	no	yes
66	Dean	[30]	Table10S	UK	1963	1972	case-control	healthy	yes	no	yes
67	Dean	[30]	Table10S	UK	1963	1972	case-control	healthy	yes	no	yes
68	Dean	[30]	Table10S	UK	1963	1972	case-control	healthy	yes	no	yes
69	Hirayama	[31]	Textp178	Japan	1965	1981	prospective	healthy	no	no	no
70	Ockene	[32]	Table 2	USA	1973	1985	prospective	healthy	no	no	no
71	Texas	[33]	AppCT29	USA	1976	1980	case-control	both	yes	no	yes
72	Correa	[34]	Table 4	USA	1979	1981	case-control	diseased/dead	yes	no	no
73	CPS II	[35]	Textp6954	USA	1982	1988	prospective	healthy	no	no	yes
74	CPS II	[36]	Figure 2	USA	1982	1988	prospective	healthy	no	no	yes
75	CPS II	[36]	Figure 3	USA	1982	1988	prospective	healthy	no	no	yes
76	De Stefani 2	[37]	Table 4	Uruguay	1993	1996	case-control	diseased/dead	no	no	no
77	Zemla	[38]	Table 1	Poland	N.K.	N.K.	case-control	diseased/dead	no	no	no
78	Wicken	[39]	Table 8	UK	1960	1962	case-control	diseased/dead	yes	no	no
79	Wicken	[39]	Table 8	UK	1960	1962	case-control	diseased/dead	yes	no	yes
80	Wakai	[40]	Table 4	Japan	1988	1991	case-control	healthy	no	yes	no
81	Wakai	[40]	Table 4	Japan	1988	1991	case-control	healthy	no	yes	no
82	Choi	[41]	Table 3	Korea	1985	1988	case-control	diseased/dead	no	no	no

TABLE 2: Characteristics of studies
(continued/3)

RR no	Study	Reference number	Detailed source	Country	Period start	Finish	Study type	Control type	Proxy use	Full hist. conf.	Cigarette only smokers
83	Segi	[42]	Table 7	Japan	1962	1970	case-control	diseased/dead	no	no	no
84	Sobue	[43]	Table III	Japan	1986	1988	case-control	diseased/dead	no	yes	no
85	Sobue	[43]	Table III	Japan	1986	1988	case-control	diseased/dead	no	yes	no
86	Simonato	[44]	Table X	Various *	1988	1994	case-control	both	yes	no	yes
87	Simonato	[44]	Table X	Various *	1988	1994	case-control	both	yes	no	yes
88	CPS I	[45]	Slide B7	USA	1960	1966	prospective	healthy	no	no	yes
89	CPS I	[45]	Slide B7	USA	1960	1966	prospective	healthy	no	no	yes
90	CPS I	[45]	Slide B7	USA	1966	1972	prospective	healthy	no	no	yes
91	CPS I	[45]	Slide B7	USA	1966	1972	prospective	healthy	no	no	yes
92	Higenbottam	[46]	Table 4	UK	1967	1969	prospective	healthy	no	no	no
93	MRFIT	[47]	Table 8	USA	1973	1976	prospective	healthy	no	no	no
94	MRFIT	[47]	Table 8	USA	1973	1976	prospective	healthy	no	no	no
95	Lubin	[12]	Table X	Various *	1976	1980	case-control	diseased/dead	no	yes	no
96	Lubin	[12]	Table X	Various *	1976	1980	case-control	diseased/dead	no	yes	yes
97	Lubin	[12]	Table X	Various *	1976	1980	case-control	diseased/dead	no	yes	no
98	Lubin	[12]	Table X	Various *	1976	1980	case-control	diseased/dead	no	yes	yes
99	Benhamou	[14]	Table 2	France	1976	1980	case-control	diseased/dead	no	yes	no
100	Benhamou	[14]	Table 3	France	1976	1980	case-control	diseased/dead	no	yes	no
101	Vutuc	[48]	Table 1	Austria	1976	1980	case-control	diseased/dead	no	no	yes
102	Vutuc	[48]	Table 1	Austria	1976	1980	case-control	diseased/dead	no	no	yes
103	Vutuc	[49]	Table 1	Austria	1976	1980	case-control	diseased/dead	no	no	yes
104	Vutuc	[49]	Table 1	Austria	1976	1980	case-control	diseased/dead	no	no	yes
105	Gillis	[50]	Table 5	UK	1976	1981	case-control	diseased/dead	no	no	yes
106	Gillis	[50]	Table 5	UK	1976	1981	case-control	diseased/dead	no	no	yes
107	AHF	[51]	Figure 1	USA	1977	1984	case-control	diseased/dead	no	yes	yes
108	AHF	[51]	Figure 2	USA	1977	1984	case-control	diseased/dead	no	yes	yes

TABLE 2: Characteristics of studies
(continued/4)

RR no	Study	Reference number	Detailed source	Country	Period start	Finish	Study type	Control type	Proxy use	Full hist. conf.	Cigarette only smokers
109	Sidney	[20]	Table 3	USA	1979	1987	prospective	healthy	no	no	yes
110	Sidney	[20]	Table 3	USA	1979	1987	prospective	healthy	no	no	yes
111	Sidney	[20]	Table 3	USA	1979	1987	prospective	healthy	no	no	yes
112	Sidney	[20]	Table 3	USA	1979	1987	prospective	healthy	no	no	yes
113	Wilcox	[52]	Table 2	USA	1980	1981	case-control	both	yes	yes	no
114	Wilcox	[52]	Table 2	USA	1980	1981	case-control	both	yes	yes	no
115	Kaufman	[53]	Table 5	USA	1981	1986	case-control	diseased/dead	no	no	no
116	Kaufman	[53]	Table 5	USA	1981	1986	case-control	diseased/dead	no	no	yes
117	Kaufman	[53]	Table 5	USA	1981	1986	case-control	diseased/dead	no	no	no
118	Kaufman	[53]	Table 5	USA	1981	1986	case-control	diseased/dead	no	no	yes
119	Speizer	[54]	Text p477	USA	1982	1986	prospective	healthy	no	no	yes
120	Speizer	[54]	Text p477	USA	1982	1986	prospective	healthy	no	no	yes
121	Alderson	[55]	Table 13F	UK	1977	1982	case-control	healthy	no	no	no
122	Alderson	[55]	Table 13F	UK	1977	1982	case-control	diseased/dead	no	no	yes
123	Tang	[7]	Table 5	UK	1967	1982	prospective	healthy	no	no	yes
124	CPS II	[35]	Text p6954	USA	1982	1986	prospective	healthy	no	no	yes
125	Segi	[42]	Table 7	Japan	1962	1970	case-control	diseased/dead	no	no	no
126	Blizza2003A	[56]	Tables 4,5	Australia	1994	1997	case-control	healthy	yes	no	yes
127	Blizza2003A	[56]	Tables 4,5	Australia	1994	1997	case-control	healthy	yes	no	yes
128	CPS II	[57]	Table 1	USA	1982	1988	prospective	healthy	no	no	yes
129	CPS II	[57]	Table 1	USA	1982	1988	prospective	healthy	no	no	yes
130	CPS II	[57]	Table 2	USA	1982	1988	prospective	healthy	no	no	yes
131	CPS II	[57]	Table 2	USA	1982	1988	prospective	healthy	no	no	yes
132	Maruga2004	[58]	Table 4	Japan	1996	1998	case-control	diseased/dead	no	yes	yes
133	Maruga2004	[58]	Table 4	Japan	1996	1998	case-control	diseased/dead	no	yes	yes
134	SHHS	[59]	Table 3	Scotland	1984	1999	prospective	healthy	no	no	no
135	SHHS	[59]	Table 3	Scotland	1984	1999	prospective	healthy	no	no	no

TABLE 2: Characteristics of studies
(continued/5)

RR no	Study	Reference number	Detailed source	Country	Period start	Finish	Study type	Control type	Proxy use	Full hist. conf.	Cigarette only smokers
136	CPS II	[57]	Table 1	USA	1982	1988	prospective	healthy	no	no	yes
137	CPS II	[57]	Table 1	USA	1982	1988	prospective	healthy	no	no	yes
138	CPS II	[57]	Table 2	USA	1982	1988	prospective	healthy	no	no	yes
139	CPS II	[57]	Table 2	USA	1982	1988	prospective	healthy	no	no	yes
140	Rachtan	[60]	Table 2	Poland	1991	1997	case-control	healthy	no	yes	yes
141	Brooks	[61]	Table 1	USA	1981	2000	case-control	diseased/dead	no	yes	no
142	Kreuzer	[62]	Table 2	Various *	1988	1994	case-control	both	no	no	yes
143	Kreuzer	[62]	Table 2	Various *	1988	1994	case-control	both	no	no	yes
144	Agudo	[63]	Table II	Various *	1988	1994	case-control	both	no	yes	yes

RR no — Relative risk number (see Table 1 for the relative risks)
Control type — Case-control studies are classified as having used healthy (population) controls, diseases or dead controls or both types; Prospective studies are classified here as healthy
Proxy use — Yes = some questions answered by proxy respondent
Full hist. conf. — Yes = all lung cancer cases in study confirmed by histology
Cigarette only smokers — Yes = smokers of pipes/cigars excluded (or study of women)
N.K. — Not known
* — The Lubin study was conducted in Austria, France, Germany, Italy and Scotland
The Simonato study was conducted in France, Germany, Italy and Spain (and in other centres with no relevant data)
The Kreuzer study was conducted in Germany and Italy
The Agudo study was conducted in France, Germany, Italy and Spain

TABLE 3: Meta-analyses

Analysis/data set	N	Fixed-effects meta-analysis RR (95% CI)	Heterogeneity χ^2	Heterogeneity p	Random-effects meta-analysis RR (95% CI)
Filter/plain unadjusted for cigs/day					
All available data	52	0.63 (0.60-0.67)	188.00	<0.001	0.61 (0.55-0.68)
Exclude overlaps	46	0.63 (0.60-0.67)	164.99	<0.001	0.63 (0.56-0.71)
All available data with corresponding adjusted data	33	0.63 (0.59-0.67)	89.85	<0.001	0.61 (0.53-0.69)
Both exclusions	31	0.64 (0.60-0.68)	81.42	<0.001	0.62 (0.55-0.70)
Filter/plain adjusted for cigs/day					
All available data	36	0.65 (0.61-0.70)	98.04	<0.001	0.66 (0.59-0.75)
Exclude overlaps	33	0.65 (0.61-0.70)	95.98	<0.001	0.66 (0.58-0.76)
All available data with corresponding adjusted data	33	0.63 (0.59-0.68)	85.73	<0.001	0.65 (0.57-0.74)
Both exclusions	31	0.63 (0.59-0.68)	84.85	<0.001	0.65 (0.57-0.75)
Low tar/high tar unadjusted for cigs/day					
All available data	18	0.71 (0.64-0.79)	68.91	<0.001	0.61 (0.49-0.78)
Exclude overlaps	14	0.74 (0.66-0.83)	54.94	<0.001	0.67 (0.52-0.86)
All available data with corresponding adjusted data	16	0.66 (0.59-0.74)	56.41	<0.001	0.56 (0.44-0.72)
Both exclusions	12	0.69 (0.61-0.78)	44.23	<0.001	0.60 (0.46-0.80)
Low tar/high tar adjusted for cigs/day					
All available data	24	0.73 (0.67-0.80)	35.13	<0.1	0.72 (0.64-0.81)
Exclude overlaps	19	0.76 (0.69-0.83)	25.22	NS	0.75 (0.67-0.85)
All available data with corresponding adjusted data	16	0.71 (0.63-0.80)	26.20	<0.05	0.68 (0.57-0.81)
Both exclusions	12	0.73 (0.65-0.83)	18.47	<0.1	0.73 (0.60-0.87)

N number of estimates meta-analysis based on
RR relative risk
CI confidence interval

TABLE 4: Filter/plain relative risk by level of some study characteristics using data excluding overlaps

	Unadjusted for cigs smoked				Adjusted for cigs smoked			
Characteristic/level	N	RR (95%) CI	p_1	p_2	N	RR (95%) CI	p_1	p_2
Gender								
Males	28	0.66(0.62-0.70)	---		23	0.65(0.61-0.70)	---	
Females	15				10			
Combined	3				0			
Females + combined	18	0.55(0.50-0.62)	---	**	10	0.64(0.54-0.75)	---	NS
Continent								
North America	14	0.60(0.55-0.65)	---		11	0.70(0.63-0.77)	---	
Europe	21	0.69(0.63-0.75)	---		16	0.64(0.58-0.71)	---	
South America	4				2			
Asia	5				4			
Australasia	2				0			
S America, Asia or Australasia	11	0.61(0.53-0.70)	---	(*)	6	0.56(0.46-0.68)	---	NS
Year of start								
1948-70	13	0.61(0.54-0.69)	---		11	0.68(0.61-0.77)	---	
1971-80	14	0.71(0.65-0.79)	---		13	0.67(0.60-0.75)	---	
1981+	19	0.60(0.55-0.64)	---	*	9	0.60(0.53-0.68)	---	NS
Year of finish								
1952-80	16	0.66(0.61-0.73)	---		12	0.58(0.52-0.64)	---	
1981-90	14	0.61(0.55-0.67)	---		13	0.73(0.66-0.81)	---	
1991+	16	0.62(0.56-0.68)	---	NS	8	0.65(0.54-0.78)	---	**
Number of lung cancers								
1-100	13	0.75(0.63-0.90)	--		10	0.84(0.70-1.01)	(-)	
101-300	15	0.59(0.53-0.66)	---		12	0.55(0.49-0.63)	---	
301+	18	0.63(0.59-0.68)	---	(*)	11	0.67(0.61-0.73)	---	***
Full histological confirmation								
Yes	17	0.55(0.51-0.60)	---		13	0.55(0.50-0.61)	---	
No	29	0.72(0.67-0.78)	---	***	20	0.74(0.68-0.81)	---	***
Study type								
Prospective	11	0.71(0.63-0.79)	---		13	0.77(0.70-0.85)	---	
Case-control	35	0.61(0.57-0.65)	---	*	20	0.57(0.52-0.63)	---	***
Type of controls used								
Healthy/Prospective	18	0.64(0.58-0.70)	---		16	0.74(0.67-0.81)	---	
Diseased	24				16			
Both	4				1			
Diseased or both	28	0.63(0.59-0.67)	---	NS	17	0.59(0.53-0.64)	---	***
Used proxies								
Yes	10	0.75(0.66-0.86)	---		2	0.34(0.22-0.51)	---	
No	36	0.61(0.58-0.65)	---	**	31	0.66(0.62-0.71)	---	**
Standardized for age								
Yes	18	0.66(0.61-0.72)	---		27	0.64(0.59-0.68)	---	
No	28	0.61(0.57-0.66)	---	NS	6	0.74(0.63-0.87)	---	(*)
Standardized for other factors								
Yes	9	0.67(0.62-0.74)	---		18	0.68(0.62-0.75)	---	
No	37	0.61(0.57-0.65)	---	(*)	15	0.62(0.56-0.68)	---	NS

TABLE 4: (continued) Filter/plain relative risk by level of some study characteristics using data excluding overlaps

	Unadjusted for cigs smoked				Adjusted for cigs smoked			
Characteristic/level	N	RR (95%) CI	p_1	p_2	N	RR (95%) CI	p_1	p_2
Cigarettes only								
Yes	24	0.69(0.64-0.75)	---		19	0.70(0.64-0.76)	---	
No	22	0.59(0.55-0.63)	---	**	14	0.60(0.55-0.67)	---	*
Filter/plain comparison								
F/P one timepoint	19	0.68(0.62-0.75)	---		15	0.74(0.67-0.81)	---	
F Only/P Only lifetime	10	0.65(0.60-0.71)	---		6	0.59(0.52-0.68)	---	
Other	17	0.55(0.50-0.61)	---	**	12	0.56(0.49-0.64)	---	**
All estimates	46	0.63(0.60-0.67)	---	***	33	0.65(0.61-0.70)	---	***

Footnotes:

N	number of estimates used in meta-analysis
RR (95% CI)	relative risk (95% confidence interval) from fixed-effects meta-analysis
p_1	coded p value for significance of relative risk --- $p<0.001$, -- $p<0.01$, - $p<0.05$, (-) $p<0.1$, NS $p\geq 0.1$
p_2	coded p value for significance of heterogeneity of relative risk over levels *** $p<0.001$, ** $p<0.01$, * $p<0.05$, (*) $p<0.1$, NS $p\geq 0.1$

TABLE 5: Multivariate weighted linear regression analysis for log of filter/plain relative risks adjusted for cigs/day using data excluding overlaps

Factor*	D.F.	Models including factors**				Adjustments to the final top down model***			
		Individual	Top down models			Exclude a factor		Exclude a factor from the model without study type	
			Sequence	Deviance	P	Deviance	P	Deviance	P
None†		95.98 (32)		95.98 (32)		27.40 (21)		28.01 (22)	
GEN2	1	95.90	8	27.40	(*)	31.17	(*)	31.38	(*)
CON3	2	91.76	-	-	-	-	-	-	-
YRSG	2	93.27	7	31.17	**	35.88	*	37.93	**
YRFG	2	85.43	6	41.12	(*)	40.92	**	41.41	**
NLCG	2	81.39	5	47.03	**	34.32	*	34.37	*
FHIS	1	77.61	4	57.75	**	37.08	**	39.00	***
STYP	1	77.11	1	77.11	***	28.01	NS	-	-
HCO2	1	84.36	-	-	-	-	-	-	-
PROX	1	86.15	3	64.97	*	31.76	*	31.76	(*)
SAGE	1	93.17	2	70.12	**	40.34	***	41.06	***
SOTH	1	93.76	-	-	-	-	-	-	-
CONL	1	91.59	-	-	-	-	-	-	-
FPCG	2	82.38	-	-	-	-	-	-	-

* Factors are GEN2 = Gender, CON3 = Continent, YRSG = Year of study start, grouped, YRFG = Year of study finish, grouped, NLCG = Number of lung cancers group, FHIS = full histological confirmation, STYP = Study type, HCO2 = healthy controls, PROX = proxy use, SAGE = Standardized for age, SOTH = Standardized for other factors, CONL = Cigarette only smoker, FPCG = Filter/plain comparison group; see Table 4 for fuller details. Number of groups for each factor is one more than the degrees of freedom (D.F.) indicated.

** These columns show the deviances for successive models. The "Individual" values relate to models each of which includes a single explanatory variable. The columns for the "Top down models" show the results of successively adding an extra explanatory variable, the modelling software choosing the next most significant variable each time. The "Sequence" column indicates the order in which the variables were chosen and the "Deviance" column gives the deviance for the model that includes that variable and all previously chosen variables. Therefore the final top down model included 8 variables and had a deviance of 27.40. The top down process stopped at this point because the addition of any of the remaining variables did not make a significant difference to the model.

*** These columns show the effect of adjusting the final top down model (model sequence number 8) by removing variables. The last two columns show the effect on the model of each variable once Study type was excluded – note that removing Study type made no significant difference to the Top down model.

† The row "none" shows the deviance (and degrees of freedom in brackets) for the models.

TABLE 6: Multivariate weighted linear regression analysis for log of filter/plain relative risks unadjusted for cigs/day using data excluding overlaps

Factor*	D.F.	Models including factors**				Adjustments to the final top down model***	
		Individual	Top down models			Exclude a factor	
			Sequence	Deviance	P	Deviance	P
None†		165.00 (45)		165.00 (45)		109.95 (40)	
GEN2	1	157.26	3	113.40	***	123.20	***
CON3	2	159.28	-	-	-	-	-
YRSG	2	156.70	2	125.57	***	126.05	***
YRFG	2	163.22	-	-	-	-	-
NLCG	2	160.15	-	-	-	-	-
FHIS	1	142.16	1	142.16	***	148.14	***
STYP	1	160.09	-	-	-	-	-
HCO2	1	164.93	4	109.95	(*)	113.40	(*)
PROX	1	157.46	-	-	-	-	-
SAGE	1	162.93	-	-	-	-	-
SOTH	1	161.99	-	-	-	-	-
CONL	1	157.03	-	-	-	-	-
FPCG	2	155.40	-	-	-	-	-

* Factors are GEN2 = Gender, CON3 = Continent, YRSG = Year of study start, grouped, YRFG = Year of study finish, grouped, NLCG = Number of lung cancers group, FHIS = full histological confirmation, STYP = Study type, HCO2 = healthy controls, PROX = proxy use, SAGE = Standardized for age, SOTH = Standardized for other factors, CONL = Cigarette only smoker, FPCG = Filter/plain comparison group; see Table 4 for fuller details. Number of groups for each factor is one more than the degrees of freedom (D.F.) indicated.

** These columns show the deviances for successive models. The "Individual" values relate to models each of which includes a single explanatory variable. The columns for the "Top down models" show the results of successively adding an extra explanatory variable, the modelling software choosing the next most significant variable each time. The "Sequence" column indicates the order in which the variables were chosen and the "Deviance" column gives the deviance for the model that includes that variable and all previously chosen variables. Therefore the final top down model included 4 variables and had a deviance of 109.95. The top down process stopped at this point because the addition of any of the remaining variables did not make a significant difference to the model.

*** These columns show the effect of adjusting the final top down model (model sequence number 4) by removing individual variables.

† The row "none" shows the deviance (and degrees of freedom in brackets) for the models.

TABLE 7: Low tar/high tar relative risk by level of some study characteristics using data excluding overlaps

	Unadjusted for cigs smoked				Adjusted for cigs smoked			
Characteristic/level	N	RR (95%) CI	p_1	p_2	N	RR (95%) CI	p_1	p_2
Gender								
Males	7	0.86(0.74-1.00)	(-)		10	0.77(0.69-0.86)	---	
Females	6				8			
Combined	1				1			
Females + Combined	7	0.61(0.52-0.73)	---	**	9	0.74(0.63-0.86)	---	NS
Continent								
North America	11	0.71(0.63-0.81)	---		13	0.75(0.67-0.84)	---	
Europe	3	0.86(0.68-1.09)	NS	NS	6	0.78(0.66-0.91)	--	NS
Year of start								
1948-70	0				5	0.74(0.64-0.87)	---	
1971-80	8	0.95(0.81-1.12)	NS		8	0.82(0.70-0.96)	-	
1981+	6	0.58(0.50-0.68)	---	***	6	0.71(0.60-0.84)	---	NS
Year of finish								
1952-80	3	0.89(0.71-1.12)	NS		7	0.74(0.65-0.84)	---	
1981-90	10	0.71(0.62-0.81)	---		11	0.78(0.69-0.89)	---	
1991+	1	0.31(0.11-0.87)	-	(*)	1	0.43(0.15-1.23)	NS	NS
Number of lung cancers								
1-100	5	0.70(0.51-0.98)	-		5	0.84(0.61-1.16)	NS	
101-300	1	0.11(0.04-0.30)	---		5	0.70(0.58-0.84)	---	
301+	8	0.77(0.68-0.87)	---	***	9	0.77(0.69-0.86)	---	NS
Full histological confirmation								
Yes	5	0.94(0.79-1.12)	NS		3	0.69(0.55-0.86)	--	
No	9	0.63(0.55-0.73)	---	***	16	0.77(0.70-0.85)	---	NS
Study type								
Prospective	7	0.68(0.59-0.79)	---		12	0.76(0.68-0.85)	---	
Case-control	7	0.83(0.70-0.99)	-	(*)	7	0.75(0.63-0.89)	---	NS
Type of controls used								
Healthy/Prospective	7	0.68(0.59-0.79)	---		13	0.77(0.69-0.85)	---	
Diseased/Dead or both	7	0.83(0.70-0.99)	-	(*)	6	0.73(0.60-0.88)	--	NS
Used proxies								
Yes	1	0.53(0.29-0.97)	-		1	0.61(0.32-1.15)	NS	
No	13	0.75(0.67-0.84)	---	NS	18	0.76(0.69-0.83)	---	NS
Standardised for age								
Yes	6	0.68(0.58-0.79)	---		16	0.77(0.70-0.85)	---	
No	8	0.83(0.70-0.98)	-	(*)	3	0.69(0.55-0.86)	--	NS
Standardized for other factors								
Yes	4	0.63(0.54-0.75)	---		16	0.74(0.67-0.82)	---	
No	10	0.86(0.74-1.00)	(-)	**	3	0.89(0.68-1.16)	NS	NS

TABLE 7: (continued) Low tar/high tar relative risk by level of some study characteristics using data excluding overlaps

	Unadjusted for cigs smoked				Adjusted for cigs smoked			
Characteristic/level	N	RR (95%) CI	p_1	p_2	N	RR (95%) CI	p_1	p_2
Cigarettes only								
Yes	9	0.74(0.65-0.85)	--		13	0.77(0.69-0.86)	---	
No	5	0.75(0.60-0.93)	---	NS	6	0.71(0.59-0.86)	---	NS
All estimates	14	0.74(0.66-0.83)	---	**	19	0.76(0.69-0.83)	---	NS

Footnotes:

N	number of estimates used in meta-analysis
RR (95% CI)	relative risk (95% confidence interval) from fixed-effects meta-analysis
p_1	coded p value for significance of relative risk --- $p<0.001$, -- $p<0.01$, - $p<0.05$, (-) $p<0.1$, NS $p\geq 0.1$
p_2	coded p value for significance of heterogeneity of relative risk excess levels *** $p<0.001$, ** $p<0.01$, * $p<0.05$, (*) $p<0.1$, NS $p\geq 0.1$

TABLE 8: Linear regression of the ratio of the adjusted to the unadjusted relative risk

Comparison	Data*	Weighted**	Ratio (95% CI)	Estimates <1	=1	>1	Total
Filter/plain	All available	No	1.10 (1.03-1.18)	10	3	20	33
		Yes	1.00 (0.93-1.07)				
	Excluding overlaps	No	1.07 (1.01-1.14)	10	3	18	31
		Yes	0.99 (0.92-1.06)				
Low tar/high tar	All available	No	1.12 (0.97-1.30)	6	1	9	16
		Yes	1.02 (0.91-1.14)				
	Excluding overlaps	No	1.16 (0.96-1.40)	4	1	7	12
		Yes	1.02 (0.90-1.16)				

* "All available" includes estimates with exclude code (see Table 1) of "none" or "some", while "exclude overlaps" includes only estimates with exclude code "none".

** Weighting is on the inverse of the variance of the adjusted relative risk.

References

1. Lee PN. Lung cancer and type of cigarette smoked. *Inhal Toxicol* 2001;**13**:951-76.

2. National Cancer Institute. Shopland DR, Burns DM, Benowitz NL, Amacher RH, editors. *Risks associated with smoking cigarettes with low machine-measured yields of tar and nicotine*. Bethesda, MD: US Department of Health and Human Services, National Institutes of Health, National Cancer Institute; 2001. (Smoking and Tobacco Control Monograph No. 13.)

3. Bross IDJ, Gibson R. Risks of lung cancer in smokers who switch to filter cigarettes. *Am J Public Health* 1968;**58**:1396-403.

4. Hawthorne VM, Fry JS. Smoking and health: the association between smoking behaviour, total mortality, and cardiorespiratory disease in west central Scotland. *J Epidemiol Community Health* 1978;**32**:260-6.

5. Lee PN. *Mortality from smoking-associated diseases in Great Britain. A statistical analysis of British data from the U.S.A.-U.K.-Norway migrant study*. Sutton, Surrey: P N Lee Statistics and Computing Ltd; 1979. www.pnlee.co.uk

6. Engeland A, Haldorsen T, Andersen A, Tretli S. The impact of smoking habits on lung cancer risk: 28 years' observation of 26,000 Norwegian men and women. *Cancer Causes Control* 1996;**7**:366-76.

7. Tang J-L, Morris JK, Wald NJ, Hole D, Shipley M, Tunstall-Pedoe H. Mortality in relation to tar yield of cigarettes: a prospective study of four cohorts. *BMJ (Clinical Research ed)* 1995;**311**:1530-3.

8. Wynder EL. Etiology of lung cancer. Reflections on two decades of research. *Cancer* 1972;**30**:1332-9.

9. Wynder EL, Stellman SD. Impact of long-term filter cigarette usage on lung and larynx cancer risk. A case-control study. *J Natl Cancer Inst* 1979;**62**:471-7.

10. Kabat GC. Aspects of the epidemiology of lung cancer in smokers and nonsmokers in the United States. *Lung Cancer* 1996;**15**:1-20.

11. Rimington J. The effect of filters on the incidence of lung cancer in cigarette smokers. *Environ Res* 1981;**24**:162-6.

12. Lubin JH, Blot WJ, Berrino F, Flamant R, Gillis CR, Kunze M, *et al*. Patterns of lung cancer risk according to type of cigarette smoked. *Int J Cancer* 1984;**33**:569-76.

13. Buffler PA, Pickle LW, Mason TJ, Contant C. The causes of lung cancer in Texas. In: Mizell M, Correa P, editors. *Lung cancer: causes and prevention, Proceedings of the International Lung Cancer Update Conference, New Orleans, Louisiana, March 3-5, 1983*. Deerfield Beach, Florida: Verlag Chemie International, Inc, 1984;83-99.

14. Benhamou S, Benhamou E, Auquier A, Flamant R. Differential effects of tar content, type of tobacco and use of a filter on lung cancer risk in male cigarette smokers. *Int J Epidemiol* 1994;**23**:437-43.

15. Benhamou E, Benhamou S, Flamant R. Lung cancer and women: results of a French case-control study. *Br J Cancer* 1987;**55**:91-5.

16. Benhamou E, Benhamou S. Black (air-cured) and blond (flue-cured) tobacco and cancer risk VI: lung cancer. *Eur J Cancer* 1993;**29A**:1778-80.

17. Lange P, Nyboe J, Appleyard M, Jensen G, Schnohr P. Relationship of the type of tobacco and inhalation pattern to pulmonary and total mortality. *Eur Respir J* 1992;**5**:1111-7.

18. Alderson MR, Lee PN, Wang R. Risks of lung cancer, chronic bronchitis, ischaemic heart disease, and stroke in relation to type of cigarette smoked. *J Epidemiol Community Health* 1985;**39**:286-93.

19. Stellman SD, Muscat JE, Thompson S, Hoffmann D, Wynder EL. Risk of squamous cell carcinoma and adenocarcinoma of the lung in relation to lifetime filter cigarette smoking. *Cancer* 1997;**80**:382-8.

20. Sidney S, Tekawa IS, Friedman GD. A prospective study of cigarette tar yield and lung cancer. *Cancer Causes Control* 1993;**4**:3-10.

21. Pathak DR, Samet JM, Humble CG, Skipper BJ. Determinants of lung cancer risk in cigarette smokers in New Mexico. *J Natl Cancer Inst* 1986;**76**:597-604.

22. Khuder SA, Dayal HH, Mutgi AB, Willey JC, Dayal G. Effect of cigarette smoking on major histological types of lung cancer in men. *Lung Cancer* 1998;**22**:15-21.

23. Armadans-Gil L, Vaqué-Rafart J, Rosselló J, Olona M, Alsedà M. Cigarette smoking and male lung cancer risk with special regard to type of tobacco. *Int J Epidemiol* 1999;**28**:614-9.

24. Pezzotto SM, Mahuad R, Bay ML, Morini JC, Poletto L. Variation in smoking-related lung cancer risk factors by cell type among men in Argentina: a case-control study. *Cancer Causes Control* 1993;**4**:231-7.

25. de Stefani E, Fierro L, Correa P, Fontham E, Ronco A, Larrinaga M, *et al*. *Mate* drinking and risk of lung cancer in males: a case-control study from Uruguay. *Cancer Epidemiol Biomarkers Prev* 1996;**5**:515-9.

26. Agudo A, Barnadas A, Pallares C, Martinez I, Fabregat X, Rossello J, *et al*. Lung cancer and cigarette smoking in women. A case-control study in Barcelona (Spain). *Int J Cancer* 1994;**59**:165-9.

27. Matos E, Vilensky M, Boffetta P, Kogevinas M. Lung cancer and smoking: a case-control study in Buenos Aires, Argentina. *Lung Cancer* 1998;**21**:155-63.

28. Jöckel K-H, Ahrens W, Wichmann H-E, Becher H, Bolm-Audorff U, Jahn I, *et al*. Occupational and environmental hazards associated with lung cancer. *Int J Epidemiol* 1992;**21**:202-13.

29. Doll R, Hill AB. A study of the aetiology of carcinoma of the lung. *Br Med J* 1952;**2**:1271-86.

30. Dean G, Lee PN, Todd GF, Wicken AJ. *Report on a second retrospective mortality study in North-East England - Part I. Factors related to mortality from lung cancer, bronchitis, heart disease and stroke in Cleveland County, with particular emphasis on the relative risks associated with smoking filter and plain cigarettes*. London: Tobacco Research Council; 1977. Research Paper 14.

31. Hirayama T. Lung cancer in Japan: effects of nutrition and passive smoking. In: Mizell M, Correa P, editors. *Lung cancer: causes and prevention, Proceedings of the International Lung Cancer Update Conference, New Orleans, Louisiana, March 3-5, 1983*. Deerfield Beach, Florida: Verlag Chemie International, Inc, 1984;175-95.

32. Ockene JK, Kuller LH, Svendsen KH, Meilahn E. The relationship of smoking cessation to coronary heart disease and lung cancer in the Multiple Risk Factor Intervention Trial (MRFIT). *Am J Public Health* 1990;**80**:954-8.

33. Ives JC. *Environmental exposures and lung cancer risk among women in Harris County, Texas, 1977-1980* [Thesis]. Houston, Texas: University of Texas, Health Science Centre; 1984.

34. Correa P, Pickle LW, Fontham E, Dalager N, Lin Y, Haenszel W, *et al*. The causes of lung cancer in Louisiana. In: Mizell M, Correa P, editors. *Lung Cancer Causes and Prevention*. Verlag Chemie International Inc., 1984;73-82.

35. Garfinkel L, Stellman SD. Smoking and lung cancer in women: findings in a prospective study. *Cancer Res* 1988;**48**:6951-5.

36. Thun MJ, Heath CW, Jr. Changes in mortality from smoking in two American Cancer Society prospective studies since 1959. *Prev Med* 1997;**26**:422-6.

37. de Stefani E, Deneo-Pellegrini H, Carzoglio JC, Ronco A, Mendilaharsu M. Dietary nitrosodimethylamine and the risk of lung cancer: a case-control study from Uruguay. *Cancer Epidemiol Biomarkers Prev* 1996;**5**:679-82.

38. Zemła B, Zielonka I, Kołosza Z. Tobacco smoking and exposure to dust and gas pollution in the place of work and lung cancer risk. *Neoplasma* 1988;**35**:135-43.

39. Wicken AJ. *Environmental and personal factors in lung cancer and bronchitis mortality in Northern Ireland, 1960-1962*. London: Tobacco Research Council; 1966. Research Paper 9.

40. Wakai K, Ohno Y, Genka K, Ohmine K, Kawamura T, Tamakoshi A, *et al*. Smoking habits, local brand cigarettes and lung cancer risk in Okinawa, Japan. *J Epidemiol* 1997;**7**:99-105.

41. Choi S-Y, Lee K-H, Lee T-O. A case-control study on risk factors in lung cancer. *Korean J Epidemiol* 1989;**11**:66-80.

42. Segi M, Kurihara M, Ishikawa S, Haenszel W. Epidemiological survey on lung cancer and smoking. *Lung Cancer* 1979;**19**:157-65.

43. Sobue T, Suzuki T, Fujimoto I, Matsuda M, Doi O, Mori T, *et al*. Case-control study for lung cancer and cigarette smoking in Osaka, Japan: comparison with the results from western Europe. *Jpn J Cancer Res* 1994;**85**:464-73.

44. Simonato L, Agudo A, Ahrens W, Benhamou E, Benhamou S, Boffetta P, *et al*. Lung cancer and cigarette smoking in Europe: an update of risk estimates and an assessment of inter-country heterogeneity. *Int J Cancer* 2001;**91**:876-87.

45. Hammond EC, Garfinkel L, Seidman H, Lew EA. Some recent findings concerning cigarette smoking. In: *Presented at a meeting on "The origins of human cancer" at Cold Springs Harbor Laboratory on September 14, 1976*. 1976;-24.

46. Higenbottam T, Shipley MJ, Rose G. Cigarettes, lung cancer, and coronary heart disease: the effects of inhalation and tar yield. *J Epidemiol Community Health* 1982;**36**:113-7.

47. Kuller LH, Ockene JK, Meilahn E, Wentworth DN, Svendsen KH, Neaton JD. Cigarette smoking and mortality. *Prev Med* 1991;**20**:638-54.

48. Vutuc C, Kunze M. Lung cancer risk in women in relation to tar yields of cigarettes. *Prev Med* 1982;**11**:713-6.

49. Vutuc C, Kunze M. Tar yields of cigarettes and male lung cancer risk. *J Natl Cancer Inst* 1983;**71**:435-7.

50. Gillis CR, Hole DJ, Boyle P. Cigarette smoking and male lung cancer in an area of very high incidence. I. Report of a case-control study in the West of Scotland. *J Epidemiol Community Health* 1988;**42**:38-43.

51. Wynder EL, Kabat GC. The effect of low-yield cigarette smoking on lung cancer risk. *Cancer* 1988;**62**:1223-30.

52. Wilcox HB, Schoenberg JB, Mason TJ, Bill JS, Stemhagen A. Smoking and lung cancer: risk as a function of cigarette tar content. *Prev Med* 1988;**17**:263-72.

53. Kaufman DW, Palmer JR, Rosenberg L, Stolley P, Warshauer E, Shapiro S. Tar content of cigarettes in relation to lung cancer. *Am J Epidemiol* 1989;**129**:703-11.

54. Speizer FE, Colditz GA, Hunter DJ, Rosner B, Hennekens C. Prospective study of smoking, antioxidant intake, and lung cancer in middle-aged women (USA). *Cancer Causes Control* 1999;**10**:475-82.

55. Alderson MR, Lee PN, Wang R. *Risk of lung cancer, chronic bronchitis, ischaemic heart disease, and stroke in relation to type of cigarette smoked, passive smoking and other factors.* Internal report.

56. Blizzard L, Dwyer T. Case-control study of lung cancer during 1994-1997 in the birth cohort in Tasmania, Australia, with an excess of female cases during 1983-1992. *Cancer Causes Control* 2003;**14**:123-9.

57. Harris JE, Thun MJ, Mondul AM, Calle EE. Cigarette tar yields in relation to mortality from lung cancer in the cancer prevention study II prospective cohort, 1982-8 [Abridged version]. *BMJ (Clinical Research ed)* 2004;**328**:72-6. Full version available at http://bmj.bmjjournals.com/cgi/reprint/328/7431/72

58. Marugame T, Sobue T, Nakayama T, Suzuki T, Kuniyoshi H, Sunagawa K, *et al*. Filter cigarette smoking and lung cancer risk; a hospital-based case-control study in Japan. *Br J Cancer* 2004;**90**:646-51.

59. Woodward M. Is compulsory restriction of tar yield of cigarettes a worthwhile public health policy? *Am J Prev Med* 2001;**21**:284-90.

60. Rachtan J. Smoking, passive smoking and lung cancer cell types among women in Poland. *Lung Cancer* 2002;**35**:129-36.

61. Brooks DR, Palmer JR, Strom BL, Rosenberg L. Menthol cigarettes and risk of lung cancer. *Am J Epidemiol* 2003;**158**:609-16.

62. Kreuzer M, Boffetta P, Whitley E, Ahrens W, Gaborieau V, Heinrich J, *et al*. Gender differences in lung cancer risk by smoking: a multicentre case-control study in Germany and Italy. *Br J Cancer* 2000;**82**:227-33.

63. Agudo A, Ahrens W, Benhamou E, Benhamou S, Boffetta P, Darby SC, *et al*. Lung cancer and cigarette smoking in women: a multicenter case-control study in Europe. *Int J Cancer* 2000;**88**:820-7.