Socioeconomic position and mortality risk of smoking: evidence from the English Longitudinal Study of Ageing (ELSA)

Dan Lewer, Martin McKee, Antonio Gasparrini, Aaron Reeves, Cesar de Oliveira

Background: It is not clear whether the harm associated with smoking differs by socioeconomic status. This study tests the hypothesis that smoking confers a greater mortality risk for individuals in low socioeconomic groups, using a cohort of 18,479 adults drawn from the English Longitudinal Study of Ageing. Methods: Additive hazards models were used to estimate the absolute smoking-related risk of death due to lung cancer or Chronic Obstructive Pulmonary Disease (COPD). Smoking was measured using a continuous index that incorporated the duration of smoking, intensity of smoking and the time since cessation. Attributable death rates were reported for different levels of education, occupational class, income and wealth. Results: Smoking was associated with higher absolute mortality risk in lower socioeconomic groups for all four socioeconomic indicators. For example, smoking 20 cigarettes per day for 40 years was associated with 898 (95% CI 738, 1058) deaths due to lung cancer or COPD per 100,000 person-years among participants in the bottom income tertile, compared to 327 (95% CI 209, 445) among participants in the top tertile. Conclusions: Smoking is associated with greater absolute mortality risk for individuals in lower socioeconomic groups. This suggests greater public health benefits of smoking prevention or cessation in these groups.

Introduction

Smoking is more common in low socioeconomic groups in almost all high-income countries. In the UK in 2013, for example, 14% of adults in managerial and professional occupations smoked, compared to 29% in routine and manual occupations. Smoking-related diseases such as lung cancer are also more common in low socioeconomic groups. Many studies show that differences in smoking behaviour only partially explain inequalities in smoking-related diseases such as lung cancer. It has been suggested that the harms of smoking differ by socioeconomic status, but this interaction has received little attention. In a literature search we found 13 studies that tested the interaction between socioeconomic status and smoking (details of the search are included in Supplementary Material A), but these produced conflicting results. Four found a statistically significant interaction, with low socioeconomic status associated with increased risks of smoking. For example, a cross-sectional study in Canada found that the difference in health outcomes between smokers and non-smokers was greatest for low income participants. Eight found no evidence of an interaction. For example, a cohort study in Belgium found that the all-cause mortality rate ratios comparing smokers to non-smokers did not differ by level of education. Finally, one cross-sectional study observed that the difference in health between smokers and non-smokers was greater for those in high status occupations than among manual workers. The author suggested that the effects of multiple risk factors in low socioeconomic groups do not always accumulate, limiting the risks specifically attributable to smoking. This is now commonly known as the ‘Blaxter hypothesis’.

These studies compare the relative risk of smoking across socioeconomic groups, which may hide differences in absolute risk because the baseline risk of smoking-related diseases is higher in low socioeconomic groups. The literature has three further limitations. First, studies commonly use simple measures of smoking exposure (such as classifying subjects as current, ex- or never-smokers), overlooking socioeconomic differences in total smoking history. For example, smokers in low socioeconomic groups start younger and...
smoke more cigarettes per day. Second, most studies use general health outcomes such as all-cause mortality, meaning that the interaction between smoking and socioeconomic status may be confounded by other risk factors for general poor health. Third, several of the studies use cross-sectional designs, and the results may be biased by selection if some non-smokers quit due to worsening health.

This study uses data from the English Longitudinal Study of Ageing to test the hypothesis that smoking confers greater absolute risk of mortality in low socioeconomic groups. This may help explain the steep social gradients in smoking-related diseases. The study also aims to use a more complete measure of smoking exposure and use a smoking-specific outcome.

**Methods**

**Data source**

The data source was the English Longitudinal Study of Ageing (ELSA), which includes participants in the Health Survey for England’s 1998, 1999 and 2001 rounds born before 1953. The sampling and research methodologies are described elsewhere. All independent variables were taken from the Health Survey for England data, except for wealth, which was taken from the ELSA follow-up survey conducted in 2002. The interview date was used as the study baseline. Linked mortality data including the month, year and underlying cause of death were provided to ELSA by the UK’s Office of National Statistics, with complete follow-up to April 2013.

**Study variables**

The dependent variable was survival in months from the date of interview. If the underlying cause of death was lung cancer or COPD, the cause was considered ‘smoking related’. These diseases have a high degree of specificity, allowing us to examine specific smoking-related risk, although it is recognised that smoking also contributes to many other conditions.

Participants’ smoking history was summarised using the comprehensive smoking index (CSI). This measure was chosen because it includes the duration of smoking, intensity of smoking (cigarettes per day) and time since cessation, which are important determinants of smoking-related risk and may vary by socioeconomic status. It also reflects the non-linear relationship between these variables and smoking-related risk. The formula is:

\[
CSI = \left(1 - 0.5^{\frac{\text{years since quit}}{5}}\right) \times \left(0.5^{\frac{\text{cigarettes currently smoked per day}}{5}}\right) \times \ln(\text{intensity} + 1)
\]

A ‘half-life’ of 25 years was determined by testing a range of values and maximising model fit. The recommended formula also included a ‘lag term’, as harm does not reduce immediately after cessation. However, this did not improve model fit and was excluded. The range of CSI was 0 (for never-smokers) to 3.7 (for an individual who reported smoking 80 per day for 66 years). CSI was included as a continuous measure in survival models. For reporting baseline characteristics, it was aggregated into 10-year bands.

**Missing data**

The ELSA sample included 18,651 adults. Fifty-six cases had missing data for their smoking status (current, ex- or never smoker), cigarettes currently smoked per day, passive smoking status or education level achieved. One case appeared to have an incorrect date of death. Smoking exposure could not be calculated for 56 ex-smokers who reported less than one year of smoking and 59 individuals who described themselves as current smokers yet reported smoking zero cigarettes per day. These cases were excluded, leaving 18,479 (99% of the original sample).

Income data were not available for 3,346 (18%) and occupation data were not available for 435 (2%). The age of smoking initiation was not available for 1,171 (including 14% of ex-smokers and <1% of current smokers). The number of cigarettes previously smoked per day was not available for 1,157 ex-smokers (14%). These cases were retained. Multiple imputed complete datasets (m = 20) were generated from age, sex, CSI and all socioeconomic variables, using the Amelia II package. For additive hazards models, the results were combined using Rubin’s Rule. Descriptive analyses used ‘missing’ categories rather than multiple imputation.

Wealth data were collected for a sub-sample of 10,922 participants who responded to a follow-up ELSA survey in 2002. These data were analysed using the interview date in 2002 as the baseline.

**Statistical analysis**

Baseline characteristics of study participants were reported by age, sex, all socioeconomic indicators, current smoking status, CSI and passive smoking status. Mortality rates were directly age-adjusted using ten-year age bands and the study population as the reference. Population attributable fractions were calculated for the proportion of smoking-related deaths that could be attributed to smoking, to check how closely the outcome event was related to smoking.

Additive hazards models were used to estimate the number of additional deaths per 100,000 person-years associated with the independent variables. This approach allowed us to compare the absolute risks of smoking and the numbers of deaths that could be avoided by preventing smoking in different socioeconomic groups.

Three models were fitted for each socioeconomic indicator. The dependent variable was the time in months until death due to lung cancer or COPD, with cases censored if they died due to other causes. The first model estimated the additional deaths associated with the socioeconomic indicator, adjusting for age, sex and passive smoking to allow reasonable comparisons between socioeconomic groups. The second model was also adjusted for smoking (using CSI). The third model included an interaction term between CSI and socioeconomic status and reported the effect of CSI for each socioeconomic level. P-values were reported from tests of the difference between the socioeconomic-specific effects of CSI. Models were not adjusted for environmental factors, health...
behaviours, comorbidities and other factors associated with socioeconomic status that may cause a differential impact, as these variables are plausibly on the causal pathway between socioeconomic status and the interaction effect between smoking and mortality. The analyses were repeated using mortality due to causes other than lung cancer or COPD and all-cause mortality as the events of interest. Kolmogorov–Smirnov and Cramer–von Mises tests were used to evaluate the assumption of time-invariance.26 There was evidence that the effect of age varied over follow-up. Age was therefore included as a time-varying regressor. The results report socioeconomic-specific effects of CSI from the third model, with results from all models reported in Supplementary Material B. As a sensitivity analysis, we re-ran the additive hazards models excluding all participants with CSI greater than 2.5 (equivalent to 40 cigarettes per day for 40 years), as smoking indices may perform poorly at extreme values.

As a second sensitivity analysis, we fitted Cox’s proportional hazards models to assess whether multiplicative and additive models produce different results. The results of the Cox models are shown in Supplementary Material C.

Results

The analysis included 18 479 adults and 223 641 years of follow-up (mean = 12.1 years). Table 1 shows the baseline characteristics of the sample and the age-adjusted mortality rates. Older people, men, those of lower socioeconomic status and smokers had higher rates of our measure of smoking-related mortality (i.e. death due to lung cancer or COPD), consistent with other research.

All analyses and production of graphics were carried out using R version 3.3.1.
mortality. For example, the rate of smoking-related mortality in participants with less than £76 000 of wealth was estimated to be 248 (95% CI 165, 330) per 100 000 person-years higher than in participants with more than £202 000 of wealth. Adjusting for smoking only partially explained this gradient for all socioeconomic indicators (results of models without interaction terms are shown in Supplementary Material B). We then included an interaction term between CSI and socioeconomic group to test whether the association between smoking and mortality varied across these groups, finding that the effect of smoking was greater in lower socioeconomic groups (see table 2; figure 1). For example, a one-unit increase in CSI was associated with an increase in smoking-related mortality of 160 (95% CI 102, 218) per 100 000 person-years among participants in the top income tertile, compared to 440 (95% CI 362, 519) among participants in the bottom tertile. To illustrate this, smoking 20 cigarettes per day for 40 years (CSI = 2.0) was associated with 327 (95% CI 209, 445) smoking-related deaths per 100 000 person-years in the top income tertile and 898 (95% CI 738, 1058) in the bottom tertile. While increases in smoking were also associated with increases in the rate of mortality due to other causes, the effects did not differ significantly between socioeconomic groups. In the sensitivity analysis excluding participants with CSI greater than 2.5, the point estimates for the number of deaths per 100 000 person-years attributable to a one-unit increase in CSI reduced by 15% or less. There was no apparent relationship between the change and socioeconomic level, and no changes in the gradients or significance of results.

**Discussion**

This study investigated socioeconomic differences in the relationship between smoking and mortality. The results show that substantially more deaths due to lung cancer or COPD can be attributed to smoking in low socioeconomic groups, even after accounting for the higher rates of smoking in these groups.

Many studies have shown that smoking-related diseases have a steep social gradient that is only partially explained by smoking behaviour, and future research could investigate which pathways lead to greater absolute risk in smokers of low socioeconomic status.

Table 2 Deaths per 100 000 person-years associated with a one-unit increase in the Comprehensive Smoking Index, stratified by socioeconomic status (95% CIs)

<table>
<thead>
<tr>
<th>Cause of death</th>
<th>Lung cancer or COPD</th>
<th>Other</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A-level</td>
<td>152 (89, 215)</td>
<td>206 (54, 359)</td>
<td>358 (192, 525)</td>
</tr>
<tr>
<td>GCSE</td>
<td>258 (172, 344)</td>
<td>361 (189, 534)</td>
<td>619 (426, 812)</td>
</tr>
<tr>
<td>None</td>
<td>383 (322, 445)***</td>
<td>293 (153, 432)</td>
<td>675 (519, 831)***</td>
</tr>
<tr>
<td>Other</td>
<td>357 (220, 494)***</td>
<td>223 (-68, 514)</td>
<td>579 (269, 889)***</td>
</tr>
<tr>
<td>Occupation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-manual</td>
<td>241 (184, 297)</td>
<td>377 (248, 507)</td>
<td>618 (479, 757)</td>
</tr>
<tr>
<td>Skilled manual</td>
<td>366 (295, 436)***</td>
<td>178 (27, 328)</td>
<td>542 (378, 707)***</td>
</tr>
<tr>
<td>Manual</td>
<td>363 (273, 453)***</td>
<td>268 (74, 462)</td>
<td>630 (413, 847)***</td>
</tr>
<tr>
<td>Income</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;£22k</td>
<td>160 (102, 218)</td>
<td>175 (33, 317)</td>
<td>335 (181, 488)</td>
</tr>
<tr>
<td>£12k–£22k</td>
<td>314 (244, 383)***</td>
<td>360 (204, 506)</td>
<td>673 (502, 845)***</td>
</tr>
<tr>
<td>&lt;£12k</td>
<td>440 (362, 519)***</td>
<td>317 (141, 494)</td>
<td>756 (564, 948)***</td>
</tr>
<tr>
<td>Wealth</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;£202k</td>
<td>217 (121, 312)</td>
<td>251 (54, 449)</td>
<td>468 (249, 686)</td>
</tr>
<tr>
<td>£76k–£202k</td>
<td>215 (128, 302)</td>
<td>283 (87, 480)</td>
<td>498 (286, 710)</td>
</tr>
<tr>
<td>&lt;£76k</td>
<td>446 (338, 554)***</td>
<td>264 (25, 503)</td>
<td>709 (445, 973)***</td>
</tr>
</tbody>
</table>

a: ‘A-level’ means A-level or higher (including higher education); ‘GCSE’ means GCSE or equivalent (e.g. O-level).

*: P < 0.05.

**: P < 0.01.

***: P < 0.001.
P-values test the difference between each value and the value for the highest socioeconomic group (e.g. ‘GCSE’ vs. ‘A-level’).
of smoking are therefore consistent with similar relative risks (or even a reversed gradient in relative risks), due to differing baseline mortality rates across socioeconomic groups. A study comparing mortality rates of manual and non-manual workers in Scotland found that the absolute difference between smokers and non-smokers was similar in the two groups. This study looked at all-cause mortality and a wider definition of ‘smoking-related mortality’ than lung cancer and COPD, including cardiovascular disease, respiratory infections and some non-respiratory cancers. These results may therefore be consistent with the lack of difference in attributable deaths found for all-causes and ‘other causes’ of death when comparing occupational classes in the present study. Differences in the social gradients in smoking-related risk for other causes of death may be an area for further research.

By using an outcome that is closely related to smoking (mortality due to lung cancer or COPD) and a smoking exposure variable that captures important elements of risk, this study overcomes some of the issues with existing studies. It is a study of older people, which means that a large proportion of participants’ lifetime smoking is likely to be in the past. The mean duration of smoking was 42 years for current smokers, while mean follow-up was 12.1 years. This limits the risk of bias due to unmeasured smoking during follow-up and smoking cessation due to ill health. The study population also includes the vast majority of smoking-related deaths, with 99% of deaths due to COPD and 98% of deaths due to lung cancer in England and Wales occurring in people aged over 50.

Smoking was based on self-report and a limited number of variables (duration of smoking, intensity of smoking and time since cessation). This simplifies smoking behaviours, which are likely to change over time. Recall may be inaccurate due to the difficulty of remembering the average number of cigarettes smoked or social desirability bias. Studies that compare biomarkers such as cotinine and exhaled carbon monoxide with self-report data have found that self-report data tend to underestimate prevalence of current smoking, but underreporting is not associated with socioeconomic status. This suggests that social gradients in the effects of smoking are unlikely to be strongly affected by recall bias.

The study used cause-specific mortality data, which may be susceptible to misclassification. Death due to lung cancer is likely to be well defined and cancers are often histologically confirmed. Diagnosis of death due to COPD may have less certainty and misclassification may occur. Analyses of cause-specific mortality may also be susceptible to bias from competing risks, for which we have not accounted. Other research suggests that smokers in lower socioeconomic groups are more likely to suffer premature death due to cardiovascular causes and may be less likely to survive until they develop lung cancer or COPD. This scenario would mean that the observed differences in the risks of smoking are understated.

The Health Survey for England (from which the ELSA sample is drawn) may have selection bias due to non-response. The response rate is difficult to estimate because the study population is a subset of Health Survey for England participants. Response rates for the full Health Survey for England sample were 69, 70 and 67% in 1998, 1999 and 2001, respectively. The results for wealth tertiles are likely to change over time. Recall may be inaccurate due to the difficulty of remembering the average number of cigarettes smoked or social desirability bias. Studies that compare biomarkers such as cotinine and exhaled carbon monoxide with self-report data have found that self-report data tend to underestimate prevalence of current smoking, but underreporting is not associated with socioeconomic status.

The study used cause-specific mortality data, which may be susceptible to misclassification. Death due to lung cancer is likely to be well defined and cancers are often histologically confirmed. Diagnosis of death due to COPD may have less certainty and misclassification may occur. Analyses of cause-specific mortality may also be susceptible to bias from competing risks, for which we have not accounted. Other research suggests that smokers in lower socioeconomic groups are more likely to suffer premature death due to cardiovascular causes and may be less likely to survive until they develop lung cancer or COPD. This scenario would mean that the observed differences in the risks of smoking are understated.

The Health Survey for England (from which the ELSA sample is drawn) may have selection bias due to non-response. The response rate is difficult to estimate because the study population is a subset of Health Survey for England participants. Response rates for the full Health Survey for England sample were 69, 70 and 67% in 1998, 1999 and 2001, respectively. The results for wealth tertiles are likely to change over time. Recall may be inaccurate due to the difficulty of remembering the average number of cigarettes smoked or social desirability bias. Studies that compare biomarkers such as cotinine and exhaled carbon monoxide with self-report data have found that self-report data tend to underestimate prevalence of current smoking, but underreporting is not associated with socioeconomic status.

The study used cause-specific mortality data, which may be susceptible to misclassification. Death due to lung cancer is likely to be well defined and cancers are often histologically confirmed. Diagnosis of death due to COPD may have less certainty and misclassification may occur. Analyses of cause-specific mortality may also be susceptible to bias from competing risks, for which we have not accounted. Other research suggests that smokers in lower socioeconomic groups are more likely to suffer premature death due to cardiovascular causes and may be less likely to survive until they develop lung cancer or COPD. This scenario would mean that the observed differences in the risks of smoking are understated.

The Health Survey for England (from which the ELSA sample is drawn) may have selection bias due to non-response. The response rate is difficult to estimate because the study population is a subset of Health Survey for England participants. Response rates for the full Health Survey for England sample were 69, 70 and 67% in 1998, 1999 and 2001, respectively. The results for wealth tertiles are likely to change over time. Recall may be inaccurate due to the difficulty of remembering the average number of cigarettes smoked or social desirability bias. Studies that compare biomarkers such as cotinine and exhaled carbon monoxide with self-report data have found that self-report data tend to underestimate prevalence of current smoking, but underreporting is not associated with socioeconomic status.

The study used cause-specific mortality data, which may be susceptible to misclassification. Death due to lung cancer is likely to be well defined and cancers are often histologically confirmed. Diagnosis of death due to COPD may have less certainty and misclassification may occur. Analyses of cause-specific mortality may also be susceptible to bias from competing risks, for which we have not accounted. Other research suggests that smokers in lower socioeconomic groups are more likely to suffer premature death due to cardiovascular causes and may be less likely to survive until they develop lung cancer or COPD. This scenario would mean that the observed differences in the risks of smoking are understated.

The Health Survey for England (from which the ELSA sample is drawn) may have selection bias due to non-response. The response rate is difficult to estimate because the study population is a subset of Health Survey for England participants. Response rates for the full Health Survey for England sample were 69, 70 and 67% in 1998, 1999 and 2001, respectively. The results for wealth tertiles are likely to change over time. Recall may be inaccurate due to the difficulty of remembering the average number of cigarettes smoked or social desirability bias. Studies that compare biomarkers such as cotinine and exhaled carbon monoxide with self-report data have found that self-report data tend to underestimate prevalence of current smoking, but underreporting is not associated with socioeconomic status.

The study used cause-specific mortality data, which may be susceptible to misclassification. Death due to lung cancer is likely to be well defined and cancers are often histologically confirmed. Diagnosis of death due to COPD may have less certainty and misclassification may occur. Analyses of cause-specific mortality may also be susceptible to bias from competing risks, for which we have not accounted. Other research suggests that smokers in lower socioeconomic groups are more likely to suffer premature death due to cardiovascular causes and may be less likely to survive until they develop lung cancer or COPD. This scenario would mean that the observed differences in the risks of smoking are understated.

The Health Survey for England (from which the ELSA sample is drawn) may have selection bias due to non-response. The response rate is difficult to estimate because the study population is a subset of Health Survey for England participants. Response rates for the full Health Survey for England sample were 69, 70 and 67% in 1998, 1999 and 2001, respectively. The results for wealth tertiles are likely to change over time. Recall may be inaccurate due to the difficulty of remembering the average number of cigarettes smoked or social desirability bias. Studies that compare biomarkers such as cotinine and exhaled carbon monoxide with self-report data have found that self-report data tend to underestimate prevalence of current smoking, but underreporting is not associated with socioeconomic status.

The study used cause-specific mortality data, which may be susceptible to misclassification. Death due to lung cancer is likely to be well defined and cancers are often histologically confirmed. Diagnosis of death due to COPD may have less certainty and misclassification may occur. Analyses of cause-specific mortality may also be susceptible to bias from competing risks, for which we have not accounted. Other research suggests that smokers in lower socioeconomic groups are more likely to suffer premature death due to cardiovascular causes and may be less likely to survive until they develop lung cancer or COPD. This scenario would mean that the observed differences in the risks of smoking are understated.

The Health Survey for England (from which the ELSA sample is drawn) may have selection bias due to non-response. The response rate is difficult to estimate because the study population is a subset of Health Survey for England participants. Response rates for the full Health Survey for England sample were 69, 70 and 67% in 1998, 1999 and 2001, respectively. The results for wealth tertiles are likely to change over time. Recall may be inaccurate due to the difficulty of remembering the average number of cigarettes smoked or social desirability bias. Studies that compare biomarkers such as cotinine and exhaled carbon monoxide with self-report data have found that self-report data tend to underestimate prevalence of current smoking, but underreporting is not associated with socioeconomic status.

The study used cause-specific mortality data, which may be susceptible to misclassification. Death due to lung cancer is likely to be well defined and cancers are often histologically confirmed. Diagnosis of death due to COPD may have less certainty and misclassification may occur. Analyses of cause-specific mortality may also be susceptible to bias from competing risks, for which we have not accounted. Other research suggests that smokers in lower socioeconomic groups are more likely to suffer premature death due to cardiovascular causes and may be less likely to survive until they develop lung cancer or COPD. This scenario would mean that the observed differences in the risks of smoking are understated.

The Health Survey for England (from which the ELSA sample is drawn) may have selection bias due to non-response. The response rate is difficult to estimate because the study population is a subset of Health Survey for England participants. Response rates for the full Health Survey for England sample were 69, 70 and 67% in 1998, 1999 and 2001, respectively. The results for wealth tertiles are likely to change over time. Recall may be inaccurate due to the difficulty of remembering the average number of cigarettes smoked or social desirability bias. Studies that compare biomarkers such as cotinine and exhaled carbon monoxide with self-report data have found that self-report data tend to underestimate prevalence of current smoking, but underreporting is not associated with socioeconomic status.