ELSEVIER

Social Science & Medicine 60 (2005) 2833–2843

SOCIAL SCIENCE & MEDICINE

www.elsevier.com/locate/socscimed

Is the urban excess in lung cancer in Scotland explained by patterns of smoking?

Jamie Pearce[a,*], Paul Boyle[b]

[a]*Department of Geography, University of Canterbury, Private Bag 4800, Christchurch 8020, New Zealand*
[b]*School of Geography & Geosciences, University of St Andrews, St Andrews. KY16 9AL, United Kingdom*

Abstract

Numerous studies have shown that lung cancer rates are higher in urban than rural areas, controlling for differences in age and sex profiles. One explanation is that smoking rates are higher in urban areas, although it is not clear whether the variations in smoking behaviour fully account for the observed urban/rural gradient in lung cancer incidence. Indeed, some studies have demonstrated an excess of cases in urban areas, even controlling for smoking behaviour. However, previous studies have been hampered by the lack of small-area smoking estimates which are required if urban/rural variations are to be examined reliably. This paper considers whether there is an urban excess in lung cancer incidence in Scotland, a country with particularly high rates of the disease, for the period 1988–1991. First, we examine whether an urban excess exists in Scotland using Poisson probabilities and a cluster detection technique. Second, regression analysis was then used to test whether any urban excess in lung cancer incidence remained once smoking behaviour was controlled for, using smoking estimates calculated for small areas throughout Scotland. The results demonstrate that the rates of lung cancer were higher in urban areas and that all the significant clusters of cases of lung cancer were located in the large urban centres of Scotland. Smoking behaviour did account for much of this urban excess in lung cancer, although it did not explain the entire effect. These results suggest that there are urban effects that influence the incidence of lung cancer that are not explained entirely by smoking behaviour. Possible explanations include the variations in exposure to air pollution, occupational differences and the legacy of selective migration between urban and rural areas.

Keywords: Lung cancer; Smoking behaviour; Urban excess; Poisson regression; Cluster analysis; Scotland, UK

Introduction

It was estimated that in the year 2000 there were 1.25 million new cases of lung cancer (WHO, 2002), and globally the rate is rising by 0.5% each year (Haugen, 2000). This makes lung cancer one of the most common causes of what has been termed 'avoidable deaths' (Payne, 2001). Distinct geographical differences in lung cancer rates exist, most noticeably between developed and developing countries, but also between quite similar countries (WHO, 2002) that appear to be at different stages in the smoking epidemic (Cavelaars et al., 2000). Whilst rates of lung cancer continue to increase in the developing world, the rates in developed countries have reduced in recent years, mainly in response to changes in smoking behaviour. However, the rate of reduction in

---

*Corresponding author. Tel.: +64 3 364 2897x7943; fax: +64 3 364 2907.
E-mail address: jamie.pearce@canterbury.ac.nz (J. Pearce).

0277-9536/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.socscimed.2004.11.014

lung cancer incidence has not been uniform as there are important localised geographical differences, with markedly different rates between regions in some countries.

Of particular note is the common finding that rates are higher in the most urban areas (e.g. Kadafar, Freedman, Goodall, & Tukey, 1996). This may be because there is an urban bias in certain behaviours or exposure to other causal factors. One suggestion is that smoking is more common in urban than rural areas (Twigg, Moon, & Jones, 2000), although it is not clear whether smoking behaviour fully accounts for the urban/rural variations in lung cancer incidence. Haynes (1988) found an urban excess in lung cancer rates during the period 1980–1983 in districts of England and Wales that was independent of smoking behaviour. It was suggested that this may reflect a combination of several effects including differences in diet, exposure to air pollution and passive smoking and selective migration from urban to rural areas. Similarly, Neuberger, Lynch, Kross, Field, and Woolson (1994) demonstrated that lung cancer incidence was higher in urban counties in Iowa after controlling for smoking. On the other hand, some studies suggest that the urban/rural gradient disappears when behavioural differences between urban and rural inhabitants, especially smoking behaviour, are controlled for (Doll, 1991; Verheij, 1996). However, these papers were geographically crude, using relatively large geographical zones rather than accurate and spatially disaggregated estimates of smoking behaviour.

This paper contributes to this debate by examining whether there is an urban excess in lung cancer in small areas across Scotland, a country with particularly high rates of the disease, and considers whether any significant urban/rural differences in lung cancer incidence remain once smoking prevalence has been controlled for. This is the first study to examine lung cancer incidence in small areas across the country juxtaposed with geographically specific estimates of smoking behaviour.

The rates of lung cancer have been consistently higher in Scotland compared with the rest of the UK. This is demonstrated in Table 1 which shows the lung cancer incidence rates for the UK, England, Wales, Northern Ireland and Scotland in 1996 (WHO, 2002). The table provides the crude incidence rate for males and females (number of cases per 100,000 people) and the age-standardised rate (ASR) using the 'world' population as the standard. Thus, lung cancer incidence is consistently higher in Scotland than in any other part of the UK. Fig. 1 compares the ASRs between 1968 and 1998 for men and women in Scotland and the whole of the UK, and shows that the rates have been consistently higher in Scotland throughout the period. Note, however, that while the rates for males are higher than for females they are converging as the rate for men has declined and the rate for women has increased over time. Indeed, the lung cancer rates for Scottish men and particularly Scottish women are currently among the highest in the world (WHO, 2002), making an examination of the incidence of lung cancer in Scotland especially pertinent.

Table 1
Rates of lung cancer in UK countries in 1996 amongst males and females (WHO, 2002)

	Males		Females	
	Crude rate*	ASR (W) **	Crude rate*	ASR (W) **
Scotland	98.40	60.09	63.54	30.86
England and Wales	77.38	44.22	41.71	19.60
Northern Ireland	65.05	46.00	33.42	19.22
UK	78.86	45.62	43.40	20.61

*Rate per 100,000.

**Age standardised rate (ASR) using 'world' standard population.

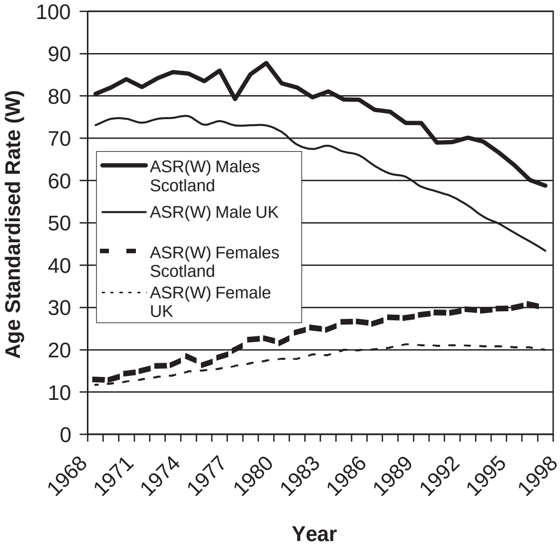

Fig. 1. Age Standardised Rate for trachea, bronchus and lung cancer incidence in the UK per 100,000 person-years at risk ('world' standard population), 1968–1997 (WHO, 2002).

Smoking has long been established as a cause of lung cancer (Royal College of Physicians, 1962) and it is estimated to be responsible for 85% of lung cancer deaths (Schottenfeld, 1996). It is widely thought that falling rates of lung cancer in developing countries over recent years is largely a reflection of falling smoking rates (Cavelaars et al., 2000). Fig. 2 shows the prevalence of smoking in Scotland and for Great Britain as a whole

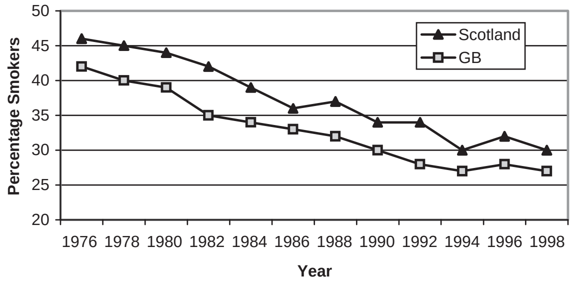

Fig. 2. Rates of smoking in Scotland and the UK as a whole for males and females, 1984 to 1998 (Office for National Statistics, 2000).

Table 2
Percentage of smokers in local authority groupings in Scotland in 2001 (Dudleston et al., 2002)

Local authority grouping	Percentage smokers
Edinburgh	27
Glasgow	36
Fife	28
North Lanarkshire	37
South Lanarkshire	26
Highlands and Islands	26
Grampian	24
Tayside	30
Central	30
Dunbartonshire	25
Renfrewshire and Inverclyde	26
Ayrshire	23
Lothians	31
Southern Scotland	25
All	28

for males and females during the period 1984–1998. Although the prevalence of smoking has fallen, it has remained consistently higher in Scotland than in the rest of the UK, for both males and females (Office for National Statistics, 2000). Smoking behaviour has also been shown to vary between different parts of Scotland. This variation is demonstrated in Table 2 which shows that the percentage of smokers varies between different local authority groupings in the country, as the average proportion of smokers in Scotland is 28% but this varies from 23% (Ayrshire) to 37% (North Lanarkshire) (Dudleston, Hope, Littlewood, Martin, & Ormston, 2002).

In this study, we therefore consider urban/rural variations in lung cancer in Scotland. First, Poisson probabilities are calculated for small geographical areas taking into account the age and sex profiles of the areas. This allows us to examine whether lung cancer is more common in urban than rural areas. Second, a method of cluster detection is used to examine whether there are groups of areas with significantly high and low rates of lung cancer in close proximity, again to see whether such clusters are more common in urban than rural areas. Finally, we use a statistical modelling approach to test whether lung cancer is higher in urban areas, controlling for age, sex and smoking behaviour, using estimates produced by Pearce, Boyle, and Flowerdew (2003). All three analyses were carried out at the output area level (the smallest unit of dissemination of 1991 Scottish census data). This is the first time that lung cancer has been examined for such small areas, enabling localised variations in the disease, and its causes, to be examined.

Data and methods

Lung cancer data in Scotland

An individual level data set on lung cancer incidence was provided by the Information and Statistics Division (ISD) of NHS Scotland for the period 1988–1991. The records included information on age at diagnosis, sex and the census output area (OA) in which each person lived. The date of diagnosis was defined as the date of the first consultation or admission to hospital for the cancer (ISD, 2002). There were 18,632 lung cancer patients in this period, of which 12,073 cases were male and 6559 female. The incidence of lung cancer was low below the age of 51 and over the age of 90 and most common between the ages of 61 and 80, for both males and females (Table 3).

Although it was not possible to gain access to individual postcodes for each patient, due to confidentiality restrictions, the inclusion of the OA identifier allowed census information to be associated with each record. Output areas are the smallest unit of dissemination in the Scottish census (approximately 50 households) and they nest within the larger pseudo-postcode sectors PPSs (approximately 2000 households) (Dale, 1993). At the time of diagnosis, the 18,632 lung cancer patients lived in 13,434 different OAs out of the total of the 38,254 OAs that encompassed Scotland in the 1991 census. The highest number of cases in any OA was 12 and more than half of the OAs (24,820) had no cases during the 3 year period.

Analyses of the geographical pattern of lung cancer in Scotland

Two methods were used to examine the spatial distribution of lung cancer in Scotland: Poisson probabilities and cluster detection. Poisson probability analysis provides an appropriate method for identifying areas of higher or lower than expected incidence of a particular phenomenon when the data have a Poisson

Table 3
Count of cases of lung cancer in age–sex groups, 1988–1991

Age	Male	Female	Total
30 or under	10	7	17
31–40	59	52	111
41–50	423	276	699
51–60	1771	968	2739
61–70	4502	2374	6876
71–80	3955	2070	6025
81–90	1303	743	2046
91–101	50	69	119
All Ages	12073	6559	18632

distribution. It is appropriate to undertake this analysis at the OA level because this may reveal important localised differences and the methodology is capable of addressing the issue of small counts. The probability of an event (X) occurring can be calculated as

$$P(X) = \frac{e^{-u}u^{x}}{X!},$$

where e is a constant (2.781) and u is the mean number of events in the entire population. This provides the Poisson probability for a given value. Low Poisson probabilities indicate a larger observed than expected value. For example, a Poisson probability of 0.05 indicates that there is a significant difference at the 95% level.

Usually the global mean of the number of cases is used to represent u. In the context of this work it was possible to improve upon this and calculate an expected number of lung cancer cases based upon the age structure of the population. The expected number of cases E, in OA i is calculated by

$$Ei = \frac{\Sigma Oi.Pi}{\Sigma Pi},$$

where P is the age- and sex-specific population and O is the observed number of lung cancer cases in each age- and sex-group. This calculation can then be used instead of the standard mean to calculate the Poisson probability.

The Poisson probabilities allow for comparisons to be made between areas but do not consider whether the rates are clustered in nearby areas. Clustering is useful to consider in this context as it may indicate whether broad urban areas have higher incidences of lung cancer than elsewhere. The spatial scan statistic was used to test for the presence of clusters of cases of lung cancer in Scotland and it has been applied to a number of different diseases including around a solid waste incinerator (Viel, Arveux, Baverel, & Cahn, 2000), childhood leukaemia (Hjalmars, Kulldorf, Gustafsson, & Nagarwalla, 1996), breast cancer (Kulldorff, Feuer, Miller, & Freedman, 1997) and amyotrophic lateral sclerosis (Sabel et al., 2003). The spatial scan statistic is designed to evaluate reported spatial or space-time clusters and to see if they are statistically significant; to test whether they are randomly distributed over space and/or time; and to detect areas that have significantly high or low rates (Kulldorff, Rand, Gherman, Williams, & DeFrancesco, 1998). The spatial scan statistic imposes a circular window on the mapped data and lets the centre of the circle to move over the study region so that at each position the window includes different sets of neighbouring administrative units. The window is flexible both in location and size and is able to resolve the issue of multiple hypothesis testing because the software allows the circle size to vary continuously. Instead of testing the significance of each circle separately, a single value for the test of the null hypothesis is provided. If the null hypothesis is rejected then the location of the cluster(s) that caused the rejection is specified. Circles are created around each centroid that are big enough to include a user-defined percentage of the total population (e.g. if 10% were chosen in this Scottish example then this would equate to approximately 500,000 people). The circle is maximised until it reaches this threshold. The software then randomly assigns cases to the coordinates and the analysis is run a further 9999 times (including the 'real' run provides a total of 10,000 runs). The results from each are then compared to see if the real run has been exceeded by any of the 9999 centroids. If the real case is within the 5% highest then it is significant at the 0.05 level.

The spatial scan statistic was used here to examine whether any clusters of lung cancer incidence existed, once the age and population structure of each area had been controlled for. In this case, the software assumes the number of cases in each area to be Poisson distributed. If no covariates are included then the number of cases within each area can be expected to be proportional to its population size. However, in this instance categorical covariates were included that controlled for both the population size of the areas and the age distribution of the population. Therefore an adjusted expected number of cases was used.

The groups analysed were males and females aged 16–24, 25–34, 35–44, 45–54, 55–64, 65–74 and 75 and over. The denominator population count data for each age-sex group was taken from the 1991 census. Controlling for age and sex avoids a scenario where the clusters that are detected are merely clusters of older people, who are more likely to suffer from lung cancer. One problem with this approach is that the lung cancer data set covers the period 1988–1991 and so the overlap with the population data was not exact. It was therefore possible that if the patient was diagnosed prior to the 1991 census then they could have moved into a different age group by the date of enumeration. Alternatively, they could have died by 1991 and therefore would not have been included in the census. As the precise date of diagnosis was not known for confidentiality reasons, it was not possible to calculate the exact age of the lung cancer patient in 1991.

Population-weighted centroids for the OAs were used as the geographical reference for the disease as the precise address of the lung cancer patients was not known. The 'special OAs' (e.g. military bases) that did not have a coordinate were excluded from this analysis. There is no consensus about which percentage of the total population should be included in the window (Kulldorff et al., 1998) but because of the computational problem of dealing with such a large set of areas a 15% window was used.

The methods described above are useful for identifying whether there are geographical differences in the incidence of lung cancer and, particularly, whether there is an urban excess in the distribution. A key question, and the focus of this analysis, was whether any urban excess in lung cancer incidence that was found could then be accounted for by local differences in smoking behaviour.

Smoking behaviour in scotland

Unfortunately, information on smoking behaviour is not routinely available for small areas in Scotland. Smoking behaviour in Scotland was therefore estimated for 12 age-sex groups for each OA using information on smoking from the Scottish Household Survey. Data on 13,784 individuals were extracted and the variables included the person's age, sex, whether they smoked or not, and an OA identifier. These individuals were distributed across 7,127 Scottish OAs. These data were combined with a set of OA and PPS variables from the 1991 census which were used to model the probability of smoking using a multi-level logit model. Once a suitable model had been derived, the parameters from this model were then used to calculate smoking probabilities for OAs across the whole of Scotland. A full explanation is provided in Pearce et al. (2003).

Modelling lung cancer incidence in Scotland

Poisson regression was used to model lung cancer incidence (1988–1991) in Scotland at the OA level. We aimed to test whether there was an urban/rural bias and, if so, whether it could be explained by the age and sex structure of the population and/or variations in smoking behaviour. For each OA the population count and the count of cases of lung cancer were calculated for 12 age–sex groups. Therefore, each OA was represented 12 times in the data set, once for each age–sex group. The log of the age- and sex-specific populations from the 1991 census was included as an offset (Knudsen, 1992). The age and sex groups were chosen to coincide with the age- and sex-specific smoking estimates (males and females aged 16–24, 25–34, 35–44, 45–54, 55–64 and 65 and over). Two categorical variables were included to define which age and sex group each record represented. In addition, the population density was calculated for OAs to provide an urban/rural proxy (Martin, Diamond, Brigham, Roderick, & Barnett, 2000). The smoking behaviour estimates were age- and sex-specific and when the smoking probability was missing, the OA was excluded from the analysis. Smoking probabilities were absent when it was not possible to calculate estimates due to missing census data in a small number of areas where the population was unusually small. A total of 37,907 of the 38,254 OAs in Scotland were included in the analysis and because each OA was represented 12 times there were 454,884 records in the entire data set. The smoking and population density variables were tested individually, both in their raw and logged form, where applicable, and the more significant of the two was used in subsequent models. The goodness of fit of each model was tested by comparing the reduction in the deviance with the critical chi-square value (Lovett and Flowerdew, 1989). Finally, the two-way interactions between all of the significant variables in each model were tested.

Results

Geographical variations in lung cancer in scotland

Although it is clear that Scotland has a high incidence of lung cancer compared with the rest of the UK and also compared with many other countries, the rates are not uniform across the country. Table 4 provides the indirectly standardised incidence ratio (SIR) of lung cancer cases for males and females in urban/rural quintiles, using the population density of each PPS as an urban/rural proxy. The SIR is the ratio between the observed number of cases and the expected number of cases of lung cancer using age-specific rates for the

Table 4
Mean lung cancer SIRs in five urban/rural categories by sex

Urban–Rural quintiles	Males	Females
Most rural	63.78	61.22
	73.93	79.43
Intermediate	89.83	85.73
	101.57	99.09
Most urban	124.48	124.50

Fig. 3. Local government regions and major urban centres in Scotland.

whole of Scotland (Moon et al., 2000). This is a useful summary measure for comparing disease rates between different population groups because it adjusts the data so that the confounding effect of age is removed. The SIR increased from the most rural to most urban category for males and females, which shows that the rates of lung cancer were higher in the urban areas compared with the rural ones, controlling for the age structure. However, this simple analysis masks local level geographical variations which may exist in the pattern of lung cancer incidence.

A map of the local government regions and the major urban centres across Scotland is provided for reference (Fig. 3). Poisson probabilities were calculated for the smaller OAs and Fig. 4 displays those OAs with a significantly higher incidence of lung cancer than expected. The high values were primarily in the major cities of Edinburgh, Glasgow, Dundee and Aberdeen. In addition, though, there was a small number of significantly high values in the Borders around Dumfries, south of Edinburgh and south-west of Glasgow and a few in the north and west of Scotland. Isolated

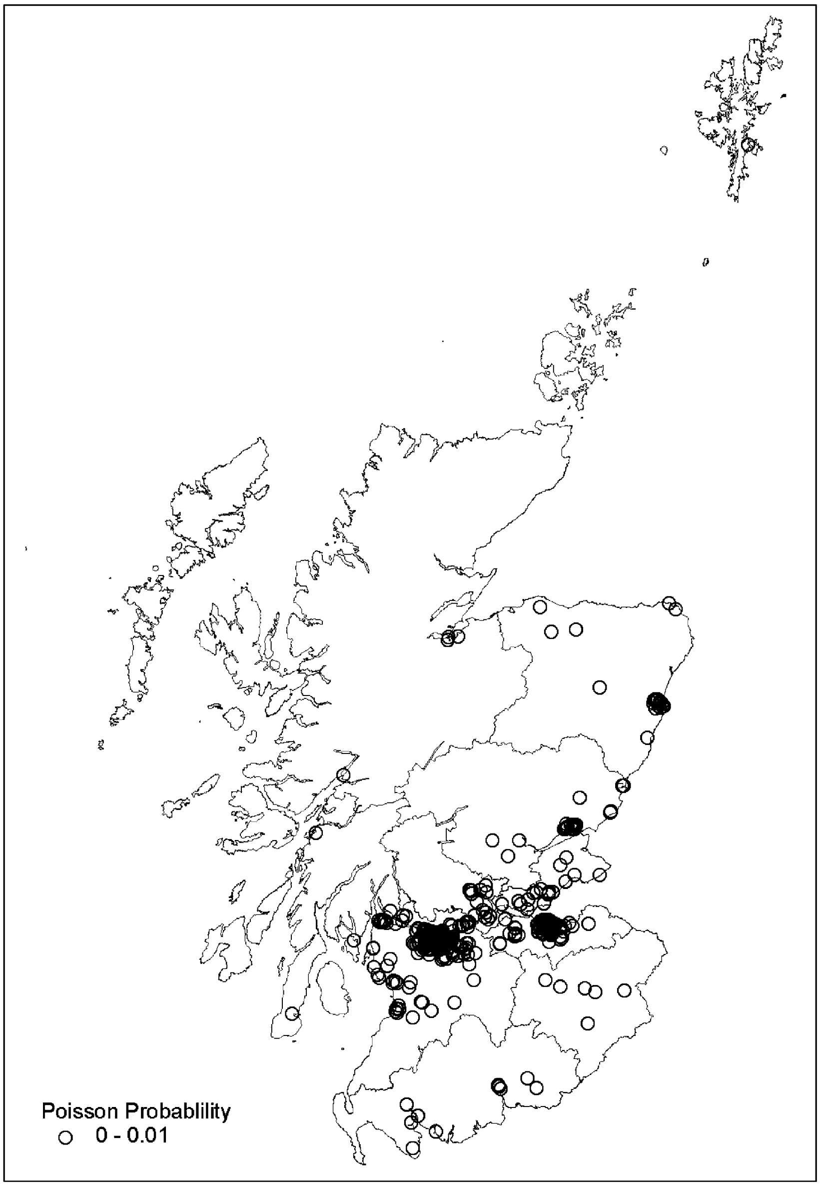

Fig. 4. Most significant Poisson probabilities in Scotland.

significant OAs were also apparent near Oban and Glen Coe as well as the north coast of Aberdeenshire. On the whole, though, rates were considerably higher in the urban areas.

The above results suggest that the incidence of lung cancer is highest in the largest Scottish cities. To test whether these apparent clusters were significant, a cluster detection analysis which also controls for the age/sex distribution was conducted and the results are shown in Fig. 5. Four significant clusters were identified in Aberdeen, Dundee, Edinburgh and Glasgow.

It is clear from these analyses that the incidence of lung cancer is indeed higher in the more urban centres of Scotland. The question that remains is whether this urban/rural gradient can be explained by geographical variations in smoking behaviour.

Geographical variations in smoking behaviour

There are important differences in smoking behaviour between the urban and rural areas of Scotland. This is demonstrated in Table 5, which uses the population

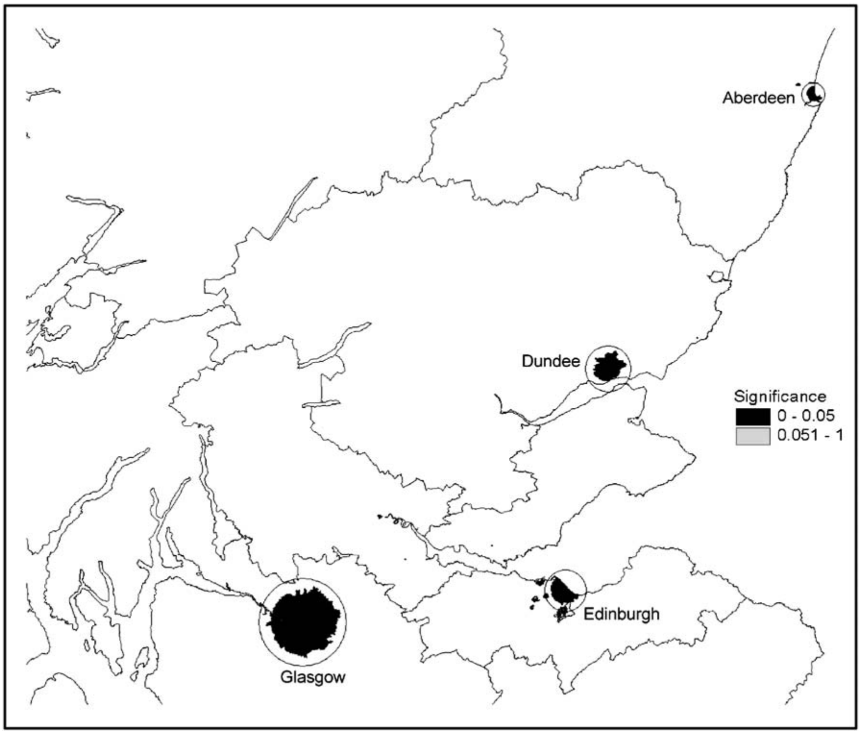

Fig. 5. Clusters of lung cancer incidence in output areas, significant at the 0.05 level, in Scotland.

Table 5
Percentage of smokers and lung cancer SIRs for males and females

Urban–rural quintiles	Percentage male smokers	Percentage female smokers	Male SIR	Female SIR
Most rural	23.39	24.44	70.19	73.69
	26.72	26.39	73.55	82.94
Intermediate	28.80	32.58	90.84	85.40
	33.58	31.97	105.47	102.00
Most urban	34.30	36.13	124.53	129.54

Table 6
Relationship between population-weighted smoking probability and lung cancer SIRs

Weighted smoking probability	SIR (per 100,000)
0.0000–0.2037	59.02
0.2038–0.2653	75.11
0.2654–0.3523	98.18
0.3524–0.4529	119.28
0.4530–1.0000	156.45

density of OAs to classify the areas where data were collected for the SHS in Scotland (7127 of the 38,254 OAs) into five urban/rural categories and shows the percentage of smokers and the SIR for lung cancer for males and females in each one. Among these selected OAs, the prevalence of smokers for males and females are lowest in the most rural areas (23.39 and 24.44) and highest in the most urban areas (34.30 and 36.13) and this coincides with progressively higher SIR values for both males and females. This would seem to suggest that the variations in lung cancer across Scotland may be largely attributed to differences in smoking behaviour.

Employing the estimates of smoking behaviour that were calculated for all OAs using the technique discussed above (see Pearce et al. 2003), all 38,254 OAs were divided into quintiles based on population-weighted average smoking proportions and the lung cancer SIRs are provided for each (Table 6). The rate of lung cancer increased consistently with the smoking probabilities. In the lowest smoking band where the probabilities of smoking were 0.2037 or less, the SIR was 59.02 whereas in the highest band where the smoking probabilities were 0.4530 or more, the SIR was 156.45. The SIR was below 100 (the expected number of cases were greater than the observed number of cases) in the lowest three smoking categories but the SIR was over 100 in the highest two categories (the observed number of cases were higher than the expected number).

Modelling lung cancer, smoking and population density

The results above demonstrate that lung cancer is more common in urban areas. They also show that lung cancer is positively associated with estimated smoking prevalence. Here we examine whether the apparent urban/rural variations in lung cancer remain when smoking is controlled for using Poisson regression.

The results are shown in models 1–4 (Table 7). Model 1 is simply the null model which includes a constant and no other explanatory variables. The deviance from subsequent models can be compared with this to test for goodness of fit. Model 2 shows that the population density variable was significantly and quite strongly associated with lung cancer incidence in Scotland (*t*-value = 18.1). Population density remained significant once age and sex were controlled for (model 3) and the deviance fell by 25.5% from 113,522 for the null model to 84,189. Note, also, that the population density variable remained highly significant even controlling for age and sex; indeed, the *t*-value rose slightly to 21.8.

However, the addition of the smoking variable to this model also reduced the deviance significantly by 1507 and the variable was positive and highly significant (model 4) with a *t*-value of 38.2. The population density variable remained significant, but only just with a *t*-value of 3.9. The substantial reduction in the significance of the population density variable demonstrates that the majority of the urban effect can be accounted for by urban/rural variations in smoking behaviour and this analysis is the first to use small-area smoking estimates and lung cancer incidence data to test this. Even so, the population density variable does remain marginally significant, suggesting that one or more urban-related variables are missing from the model.

Discussion

The Poisson probability calculations for OAs suggested that there is an urban excess in lung cancer incidence. This is supported by the results of the cluster detection analysis, which demonstrated that clusters of cases of lung cancer were located in the large urban areas. Thus, regardless of the approach taken, and despite their different limitations, lung cancer incidence appears to be higher in the urban centres of Scotland.

The univariate regression results demonstrated that the population density variable was significant in explaining lung cancer incidence in Scotland. The variable remained highly significant once age and sex were controlled for. However, the urban/rural differences in lung cancer incidence are likely to be due to one or more of the risk factors associated with the disease and here we examined the most important of these, smoking. The descriptive results in Table 6 demonstrated that the SIRs for lung cancer were higher in areas where the probability of smoking was higher. And, controlling for age and sex, the Poisson regression analysis demonstrated that the smoking probability variable was significant in explaining lung cancer incidence in Scotland. The final regression results demonstrated that the majority of the apparent urban effect could be explained by smoking, although the population density variable remained marginally significant. The results from this small-area analysis of lung cancer incidence therefore provide limited support for an urban excess in lung cancer incidence in Scotland that is independent of smoking behaviour. Previous studies, such as Haynes (1988), were not conducted at this scale of analysis and it is possible that the strong urban effect that he identified was caused by the lack of geographical specificity in both the lung cancer incidence and smoking data.

Table 7

Results from Poisson regression analysis with the age and sex-specific count of lung cancer as the dependent variable.

	Model 1			Model 2			Model 3			Model 4		
Scaled deviance	113,522			113,160			84,189			82,682		
Degrees of freedom	454,883			454,882			454,876			454,875		
	Estimate	Std. error	*t*-value	Estimate	Std. error	*t*-value	Estimate	Std. error	*t*-value	Estimate	Std. error	*t*-value
Intercept	−5.442	0.007364	−739.00	−6.065	0.0357	−169.89	−12.35	0.3473	−35.56	−10.85	0.3485	−31.13
Log pop. density				0.076	0.0042	18.10	0.096	0.0044	21.82	0.018	0.0046	3.91
Age (25–34)							1.589	0.3911	4.06	1.522	0.3905	3.90
Age (35–44)							3.858	0.3511	10.99	3.846	0.3504	10.98
Age (45–54)							5.609	0.3465	16.19	5.569	0.3459	16.10
Age (55–64)							6.977	0.3456	20.19	6.946	0.345	20.13
Age (65+)							7.328	0.3454	21.22	7.725	0.345	22.39
Sex (female)							−0.480	0.0155	−30.97	−0.401	0.0156	−25.71
Log smoking										0.822	0.0215	38.23

Of course, similar to many studies of lung cancer, this research also has limitations. First, there is likely to be a long latency period between carcinogenic exposure and diagnosis with lung cancer (Williams, 1992), and the use of 1991 smoking data may be inappropriate if there have been dramatic variations in smoking patterns during the latency period. However, spatially disaggregated smoking data were not collected in Scotland prior to the early 1990s and therefore it was not possible to calculate geographically detailed historical estimates of smoking behaviour. Furthermore, even if historical estimates of smoking behaviour were available it is not clear whether they would be more sensible than contemporary estimates given the likely levels of migration during the latency period. If a lung cancer patient has migrated during the latency period then historical estimates of smoking behaviour are likely to be less appropriate than the contemporary estimates used in this study (assuming that people who currently live in places with high levels of smoking are also likely to have smoked in the past). Second, although the small-area smoking estimates used here are the best that are available, the models of smoking behaviour did not explain all the variance (Pearce et al., 2003) and hence it is also possible that the estimates do not fully control for the variation in smoking behaviour across Scotland.

These results indicate that there is a small urban excess in lung cancer, having controlled for smoking, and a variety of reasons may explain this. As suggested above, one explanation may be that smoking was inadequately controlled for in this model. An alternative explanation is that there are other features of urban areas that influence the incidence of lung cancer, particularly air pollution (Cohen, 2000). In addition, a number of other factors have previously been identified that may, in combination, contribute to an urban effect (Haynes, 1988). These include passive smoking, particular types of employment that are more common in urban areas, geographically biased mortality statistics and selective migration. Passive smoking may have a greater influence in urban areas than in rural areas because urban areas contain proportionally more smokers than rural populations (Matsukura et al., 1984). Furthermore, people are more likely to be exposed to passive smoking in crowded, poorly ventilated public places in cities than they are in rural areas. There are also possible industrial explanations for the lung cancer urban effect. A number of hazardous industries have been recognised as potential risk factors, particularly those industries that use materials such as asbestos, arsenic, chromium and chloromethyl and these tend to be more concentrated in urban areas (Cohen, 2000).

The third suggestion is that there may be a geographical bias in mortality statistics, which has meant that cancers are more likely to be diagnosed and reported in urban areas than in rural areas (Bentham & Haynes, 1985). However, there is not any direct evidence of widespread under-diagnosis of lung cancer in rural Britain, and Scotland is recognised as having some of the most accurate health data in the world (ISD, 2002).

Finally, any urban effect may be created partially by selective migration (Haynes, 1988). Selective migration has been highlighted as one reason that may help to explain health differences between neighbourhoods, including between different areas on the urban/rural gradient (Verheij, van de Mheen, de Bakker, Groenewegen, & Mackenbach, 1998; Brimblecombe, Dorling, & Shaw, 1999; Boyle, Norman, & Rees, 2002). For example, Brimblecombe et al., (1999) used the 1991 British Household Panel Survey to investigate the extent to which selective migration influences current geographical variations in mortality. They found that local level selective migration since birth significantly altered the geographical pattern of mortality in Britain. Prior to the period of the lung cancer dataset (1988–1991), many rural communities grew at the expense of the urban populations (Champion, 1989), and studies have shown that adults who migrate relatively long distances tend to be healthier than the population that remains (Fox & Goldblatt, 1982). This trend would lead to an apparent improvement in health in rural areas but a worsening of health in urban areas. The effect may also be compounded by the net out-migration from rural areas of the elderly and chronically sick to be closer to relatives or to health and service facilities (Bentham, 1988). Therefore, the effect of selective migration could obscure the geographical distribution of many diseases, especially those with a long latency period such as lung cancer (Haynes, 1988). However, it is difficult to explicitly examine the effects of selective migration between urban and rural areas because the lung cancer data did not include a migration history and other available data on migrants do not tend to be available for OAs, nor do they identify individuals with lung cancer. Even so, it should be stressed that our analysis indicates that the majority of the urban excess in lung cancer incidence appears to be explained by geographical variations in smoking behaviour and that policies to reduce geographical inequalities in lung cancer would be advised to focus on differences in smoking.

References

Bentham, G. (1988). Migration and morbidity: implications for geographical studies of disease. *Social Science and Medicine*, *26*, 49–54.

Bentham, G., & Haynes, R. (1985). Health, personal mobility and the use of health services in rural Norfolk. *Journal of Rural Studies*, *1*, 231–239.

Boyle, P., Norman, P., & Rees, P. (2002). Does migration exaggerate the relationship between deprivation and limiting long-term illness? A Scottish analysis. *Social Science and Medicine*, *55*, 21–31.

Brimblecombe, N., Dorling, D., & Shaw, M. (1999). Mortality and migration in Britain, first results from the British Household Panel Survey. *Social Science and Medicine*, *49*, 981–988.

Cavelaars, A. E. J. M., Kunst, A. E., Geurts, J. J. M., Crialesi, R., Grötvedt, L., Helmert, U., Lahelma, E., Lundberg, O., Matheson, J., Mielck, A., Rasmussen, N. K. R., Regidor, E., do Rosário-Giraldes, M., Spuhler, T. H., & Mackenbach, J. P. (2000). Educational differences in smoking: international comparison. *British Medical Journal*, *320*, 1102–1107.

Champion, A. G. (1989). United Kingdom, population deconcentration as a cyclic phenomenon. In A. G. Champion (Ed.), *Counterurbanization, the changing pace and nature of population deconcentration* (pp. 83–102). London: Arnold.

Cohen, A. J. (2000). Outdoor air pollution and lung cancer. *Environmental Health Perspectives*, *108*, 743–750.

Dale, A. (1993). The content of the 1991 census: change and continuity. In A. Dale, & C. Marsh (Eds.), *The 1991 Census Users Guide* (pp. 16–51). London: HMSO.

Doll, R. (1991). Urban and rural factors in the aetiology of cancer. *International Journal of Cancer*, *47*, 803–810.

Dudleston, A., Hope, S., Littlewood, A., Martin, C. and Ormston, R. (2002). Scotland's people. Results from the 2002 Scottish household survey (vol. 5). Annual report. Scottish Executive National Statistics Publication: Edinburgh.

Fox, A. J., & Goldblatt, P. O. (1982). *Longitudinal study: socio-demographic mortality differentials. OPCS Series LS No. 1*. London: HMSO.

Haugen, A. (2000). Etiology of Lung Cancer. In H. H. Hansen (Ed.), *Textbook of Lung Cancer* (pp. 1–12). Copenhagen: Martin Dunitz.

Haynes, R. (1988). The urban distribution of lung cancer mortality in England and Wales 1980–1983. *Urban Studies*, *25*, 497–506.

Hjalmars, U., Kulldorf, M., Gustafsson, G., & Nagarwalla, N. (1996). Childhood leukemia in Sweden: using GIS and a spatial scan statistic for cluster detection. *Statistics in Medicine*, *15*, 707–715.

ISD (Information and Statistics Division, NHS Scotland) (2002). ISD Online www.show.scot.nhs.uk/isd/ Site accessed April 2002

Kadafar, K., Freedman, L. S., Goodall, C. R., & Tukey, J. W. (1996). Urbanicity-related trends in lung cancer mortality in US counties: white females and white males, 1970–1987. *International Journal of Epidemiology*, *25*, 918–932.

Knudsen, D. C. (1992). Generalizing Poisson regression: Including a priori information using the method of offsets. *Professional Geographer*, *44*, 202–208.

Kulldorff, M., Feuer, E. J., Miller, B. A., & Freedman, L. S. (1997). Breast cancer in the northeastern United States: a geographical analysis. *American Journal of Epidemiology*, *146*, 161–170.

Kulldorff, M., Rand, K., Gherman, G., Williams, G., & DeFrancesco, D. (1998). *SatScan v2.1: software for the spatial and space-time scan statistics*. Bethesda, MD: National Cancer Institute.

Lovett, A., & Flowerdew, R. (1989). Analysis of count data using Poisson regression. *Professional Geographer*, *41*, 190–198.

Martin, D., Diamond, I., Brigham, P., Roderick, P., & Barnett, S. (2000). The (mis)representation of rural deprivation. *Environment and Planning A*, *32*, 735–751.

Matsukura, S., Taminato, T., Kitano, N., Seino, Y., Hamada, H., Uchihashi, M., Nakajima, H., & Hirata, Y. (1984). Effects of environmental tobacco smoke on urinary cotinine excretion in nonsmokers. Evidence for passive smoking. *The New England Journal of Medicine*, *311*, 828–832.

Moon, G., Gould, M., et al. (2000). *Epidemiology—an introduction*. Buckingham: Open University Press.

Neuberger, J. S., Lynch, C. F., Kross, B. C., Field, R. W., & Woolson, R. F. (1994). Residential radon exposure and lung cancer—evidence of an urban factor in Iowa. *Health Physics*, *66*, 263–269.

Office for National Statistics. (2000). *Living in Britain: results from the 1998 general household Survey*. London: The Stationery Office.

Payne, S. (2001). 'Smoke like a man, die like a man'?: a review of the relationship between gender, sex and lung cancer. *Social Science and Medicine*, *53*, 1067–1080.

Pearce, J., Boyle, P., & Flowerdew, R. (2003). Predicting smoking behaviour in census output areas across Scotland. *Health and Place*, *9*, 139–149.

Royal College of Physicians. (1962). *Smoking and Health*. London: Pitman Medical Publishing.

Sabel, C. E., Boyle, P. J., Loytonen, M., Gatrell, A. C., Jokelainen, M., Flowerdew, R., & Maasilta, P. (2003). Spatial clustering of amyotrophic lateral sclerosis in Finland at place of birth and place of death. *American Journal of Epidemiology*, *157*, 898–905.

Schottenfeld, D. (1996). Epidemiology of lung cancer. In J. B. Mitchell, D. H. Johnson, & A. T. Turrisi (Eds.), *Lung Cancer: Principles and Practice* (pp. 305–321). Philadelphia: Lippincott-Raven.

Twigg, L., Moon, G., & Jones, K. (2000). Predicting small-area health related behaviour: a comparison of smoking and drinking indicators. *Social Science and Medicine*, *50*, 1109–1120.

Verheij, R. A. (1996). Explaining urban–rural variations in health: a review of interactions between individuals and environment. *Social Science and Medicine*, *42*, 923–935.

Verheij, R. A., van de Mheen, H. D., de Bakker, D. H., Groenewegen, P. P., & Mackenbach, J. P. (1998). Urban–-rural variations in health in The Netherlands: does selective migration play a part? *Journal of Epidemiology and Community Health*, *52*, 487–493.

Viel, J. F., Arveux, P., Baverel, J., & Cahn, J. Y. (2000). Soft-tissue sarcoma and non-hodgkin's lymphoma clusters around a municipal solid waste incinerator with high dioxin emission levels. *American Journal of Epidemiology*, *152*, 13–19.

WHO (World Health Organization) (2002). www.who.org Site Accessed February 2002.

Williams, C. (1992). *Lung Cancer: The Facts*. Oxford: Oxford University Press.