

## DETERMINANTS OF REGIONAL DIFFERENCES IN LUNG CANCER MORTALITY IN THE NETHERLANDS

A. E. KUNST, C. W. N. LOOMAN and J. P. MACKENBACH

Department of Public Health, Faculty of Medicine and Health Sciences, Erasmus University Rotterdam,  
P.O. Box 1738, 3000 DR Rotterdam, The Netherlands

**Abstract**—Although regional differences in lung cancer mortality are likely to be attributable to regional differences in tobacco smoking, studies in various countries found only weak relationships. This paper aimed at explaining regional differences in lung cancer mortality in the Netherlands.

In a first step, clues for the role of smoking were obtained from a detailed description of regional mortality differences. These differences were found to be strongly determined by cohort effects: they vary between birth cohorts, and have been stable for over 30 years. Regional mortality differences reflect a diffusion of the lung cancer epidemic from high-income regions to low-income regions. These findings are suggestive of a relationship with regional differences in trends in cigarette smoking.

In a second step, by means of multiple regression analysis, mortality differences in 1980-84 were related to available data on cigarette smoking and two other possible risk factors: work in transport and manufacturing industry, and air contamination. The independent variables referred to the 1970s. Positive associations with various smoking measures were found for women, but for men the associations were weak or non-existent. Mortality differences among men 45-64 years were associated with work in transport and manufacturing industry. Strong associations with air contamination were found for men and women 65 years and older.

Additional analysis showed that regional differences in lung cancer among old men were strongly associated with smoking in 1930, i.e. half a century before. Changes in the regional pattern of tobacco consumption between 1930 and 1970 explain why smoking in 1970 is not associated with mortality differences in the 1980s. This shows that the explanation of regional lung cancer death rates sometimes has to go far back in time.

The strong association with air contamination does not constitute firm evidence of an effect, but should be considered as a stimulus to carry out further analysis on the excess lung cancer mortality in polluted regions.

*Key words*—mortality, lung cancer, regional differences, determinants

### INTRODUCTION

In most countries, lung cancer death rates vary markedly between regions [1-11]. Large health or mortality differentials within national populations often are of concern to health policy makers. With regard to potential preventive policies, such differentials raise questions on the causes of excess mortality or morbidity in one population group over another.

The answer may seem to be easy when lung cancer is concerned. As the lung cancer epidemic is predominantly tobacco-related [12-14], regional differences in lung cancer mortality are likely to be attributable to differences in tobacco consumption. Regional data on current smoking are indeed associated with lung cancer mortality rates in various countries [6, 15]. However, often a large part of the variation is left unexplained [1, 2, 4, 8, 16, 17].

This unexplained variation invites a further search for the determinants of the regional differences in lung cancer mortality. The role of smoking should be explored further. An obvious disadvantage of the previously applied smoking measures is that they did not allow for a time-lag between exposure to tobacco smoke and death from lung cancer. Applying

measures on the current level of smoking would be justified if regional differences in tobacco consumption had not changed considerably over time and, as a consequence, current smoking differences reflect past differences in smoking. Since this may not be the case, it is preferable to use data on past smoking directly.

Although it is likely that most of the variation in the occurrence of lung cancer mortality can be explained by regional differences in tobacco smoking, other factors may play a role too. Even though relative risks and population attributable risks of factors like air contamination [18, 19] and work in specific occupations [20] are relatively small, these factors would contribute substantially to variation in death rates between regions if their prevalence rates vary strongly among regions.

The analysis reported here aims at finding clues to the explanation of regional differences in lung cancer mortality in the Netherlands in 1980-1984. Regional mortality differences will be related to available data on past smoking and two possible risk factors for lung cancer: outdoor air contamination and occupational exposure. Accurate and detailed data on past smoking and on air pollution were available for

Table 1. Age- and sex-specific death rates for lung cancer, 1980–84

Age group	Females		Males	
	Number of deaths	Per 100,000 population	Number of deaths	Per 100,000 population
45–54	567	15	2603	68
55–64	897	25	8224	257
65–74	1037	37	14190	653
75–84	959	56	9382	927
85+	345	78	1562	708
all 45+	3805	31	35961	345

1970s, which implies assuming a time-lag of about 10 years.

Occupational exposure had to be measured by a proxy variable on work in manufacturing industry and transport.

The available data have, as is usual in regional analyses, a number of limitations. For example, although this study is able to use data on past instead of present smoking, regional differences in smoking more than 10 years back in time cannot be measured directly. In the view of these limited data, an attempt is made to obtain additional clues for the role of smoking vs other determinants by means of a detailed description of mortality differences. This description will focus on two aspects.

First, by comparing birth cohorts and, within birth cohorts, periods. National mortality trends are strongly determined by cohort effects [21], which are principally related to cohort differences in tobacco consumption history [13]. It would be suggestive of the role of smoking if cohort effects were also found in regional mortality patterns. Therefore, it will be investigated whether regional mortality patterns vary between birth cohorts and whether, within birth cohorts, regional mortality differences have been constant over time.

Second, by relating regional differences in lung cancer mortality to socio-demographic variables. Surveys in the Netherlands in the 1960s and 1970s showed higher smoking prevalences in specific socio-cultural and socio-economic groups like Roman Catholic men and women, lower-class men, and higher-class women [22]. If regional differences in lung cancer mortality in 1980–84 are strongly related to these same socio-cultural and socio-economic variables, this would again be suggestive of the role of smoking.

#### MATERIALS AND METHODS

Numbers of deaths from lung cancer, by age, sex, year of death, and region of residence, were extracted from a large data file supplied by the Dutch Central Bureau of Statistics (CBS). All deaths from cancer of

the trachea, bronchus and lung were selected. The corresponding ICD codes were for ICD-6/7: 162, 163; for ICD-8: 162, 163.0, 163.9; and for ICD-9: 162, 163, 165. Four 5-year periods (1950–54, 1960–64, 1970–74, and 1980–84) were distinguished. The data refer to 39 so-called 'COROP' regions, which are statistical, nodal regions with a median population size of about 300 thousand inhabitants in 1980–1984. Corresponding numbers of population-at-risk were calculated from population data that are based on the national continuous population register.

Five age groups were distinguished in the basic data. National death rates for these age groups are presented in Table 1. Age groups younger than 45 years were not included because of very large fluctuations in regional numbers of deaths. In some analyses, the age groups 45–54 and 85+ years were added to, respectively, the age groups 55–64 and 75–84 years, in order to reduce random fluctuations in regional numbers of death.

The selection of socio-demographic variables was guided by a previous analysis of a wide range of demographic, socio-cultural and socio-economic indicators with a potential relevance for the explanation of mortality patterns. Factors relating to climate, health care, and genetic predisposition were considered to be of minor relevance [23]. Principal component analysis showed that the selected factors can be reduced to three dimensions: a socio-economical dimension (measured here by the average income per capita), an urbanisation dimension (measured by the percentage of inhabitants living in large cities), and a socio-cultural dimension (measured by the percentage of the population belonging to the Roman Catholic denomination). A variable on religious denomination was used because the major socio-cultural division within the Netherlands is between the Protestant north and the Roman Catholic south [24, 25]. We have previously shown that this socio-cultural contrast separates two distinct mortality regimes [23]. The socio-demographic data used in this analysis refer to the beginning of the 1970s.

Accurate smoking data were obtained for as early a period as possible. One of the first large Dutch surveys with questions on smoking was the National Press Survey of 1972. The total number of respondents (30746) was sufficiently large for regional disaggregation. Regional samples of respondents were representative of the total population according to age, sex, household size and size of municipality. For each age- and sex-group, the following measures could be calculated:

—The percentage of respondents currently smoking cigarettes. Essentially the same results were found

Fig. 1. Geographic pattern of lung cancer mortality in 1980–84, by sex and age-group. Legend: SMR □ 0.00–0.75; ▨ 0.76–0.90; ▩ 0.91–1.00; ■ 1.01–1.15; ■ 1.16+. The class division is centred around 0.955, which is the (unweighted) average SMR. This average is below the 1.00 because high SMRs are concentrated to a few regions with large population sizes.

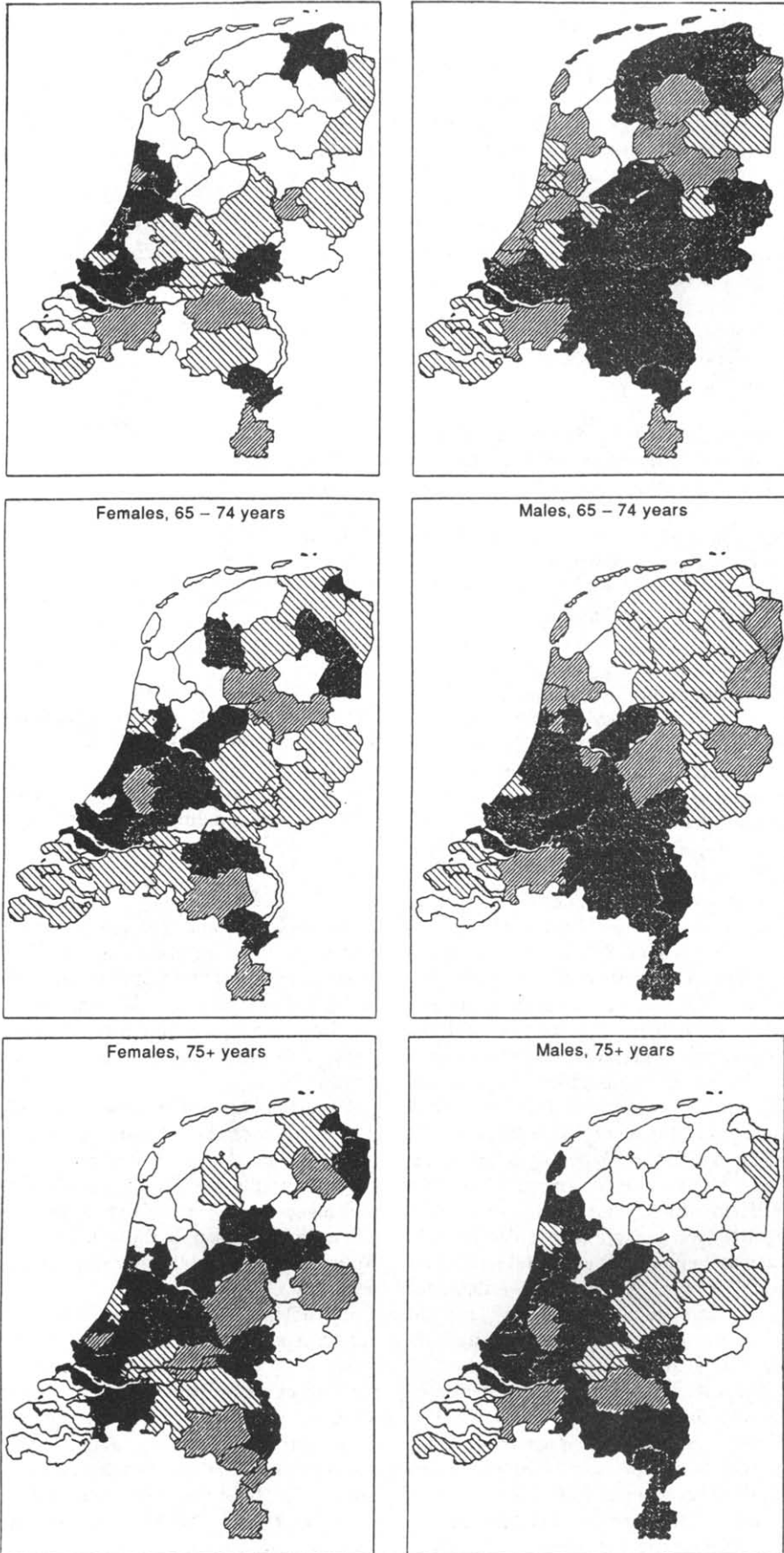


Fig. 1. See caption opposite.

in analyses with measures including pipe and cigar smokers, or excluding filter cigarette smokers.

—Mean number of cigarettes smoked per respondent. The number of cigarettes was shown to be underreported by about 23% in the whole country [26].

—The percentage of respondents that has ever smoked cigarettes (i.e. current and ex-smokers). This variable was considered as an approximate measure of life-time exposure. From a comparison with a national survey in 1958 [26], we found that the number of ex-smokers was underreported by men 55 years and older.

Younger age-groups (35–54, 55–64, 65+) were distinguished than with the mortality data, because the survey data refer to a time approx. 10 years before 1980–84.

Data on work in work in manufacturing industry and transport by age, sex, and region, were obtained from the population census of 1971. Data on 1971 were used because it is likely that there is a time-lag of about 10 years or more between occupational exposure and death from lung cancer. The 1971 census-data concerned the age groups 35–54 and 55–64 years (for men), and 35–64 years (for women). Regional differences in past exposure of persons that were 65 years or more in 1971 were measured by data for younger age groups.

Regional levels of air contamination were measured by the 50-percentile density of SO<sub>2</sub> in April 1979–April 1980. The data were generated by a monitoring system with 206 monitoring stations in 1979–1980, i.e. on the average about 5 stations per region [27]. For each region, SO<sub>2</sub> density was calculated as the weighted average of density values for individual stations. Stations were weighted by the size of the population that they represent, so that the weighted index reflects more accurately the average exposure of inhabitants to outdoor air contamination. The measure employed here correlated strongly with regional differences in the 95-percentile density of SO<sub>2</sub>, and with the SO<sub>2</sub> density in the winter season of 1979–80. Because SO<sub>2</sub> density was also strongly correlated with regional variation in NO<sub>2</sub>, particulate matter, and benzo(a)pyrene density, it is considered here as an index of the general level of air contamination. No extensive data were available for earlier years, but there are no reasons to believe that the regional pattern of air contamination has radically changed during the 1970s.

The mortality level in a region was measured by means of Standardised Mortality Ratios (SMRs), with national death rates by 10-year age group as the standard. The significance of overall mortality differences was established by means of Gail's test [28]. SMRs were related to regional characteristics by means of product-moment correlations and multiple regression analysis. SMRs were transformed to their natural logarithms so as to normalise the distribution

Table 2(A). Correlation between age-specific lung cancer mortality rates in 1950–54 and 1980–84, by sex and age-group

Age group	Product-moment correlation <sup>a</sup>	
	Females	Males
45–54	0.30	0.02
55–64	0.06	–0.06
65–74	0.10	0.45**
75–84	–0.15	0.34*
85+	0.33	0.06

<sup>a</sup>Product-moment correlation between log-transformed SMRs, weighted by population size, *N* = 39 regions.

\**P* < 0.05; \*\**P* < 0.01.

Table 2(B). Correlation between age-specific lung cancer mortality rates in 1950–54 and 1980–84, by birth cohort

Birth cohort	Product-moment correlation <sup>a</sup>	
	Females	Males
ca 1890 (55–64 year in 1950–54, 85–94 year in 1980–84)	–0.17	0.74**
ca 1900 (45–54 year in 1950–54, 75–84 year in 1980–84)	0.27	0.79**

<sup>a</sup>Product-moment correlation between log-transformed SMRs, weighted by population size, *N* = 39 regions.

\**P* < 0.05; \*\**P* < 0.01.

of residuals. Regions were weighted to their population size in order to reduce effects of random fluctuations in numbers of (age specific) deaths and in smoking measures. As independent variables were not strongly correlated (the highest correlation was 0.55), no strong multicollinearity problem existed.

## RESULTS

For each age and sex group, overall mortality differences were significant at the 5% level. The geographic patterns of mortality variation are shown in Fig. 1. Large mortality differences exist for women in all age groups and for men 75 years and older, while differences among younger men are much smaller.

All female age groups have a similar geographic pattern of mortality differences, with the highest death rates in the central-western part of the Netherlands. Mortality rates in the central-western part are also higher among men 75 years and older, but not among younger men. In the south-eastern part of the Netherlands, men in all age groups experience higher death rates.

In Table 2(A) the pattern of regional mortality variation in 1980–84 is compared to its pattern in 1950–54, by age and sex group. The low correlations for females are probably due to large random fluctuations in the small numbers of deaths in 1950–54. Low correlations are also found when men aged 45–64 years in 1980–84 are compared to the same age group in 1950–54 (i.e. birth cohorts that were 75–94 years in 1980–84). Correlations for men in the age group 65–84 years are moderate.

In Table 2(B) periods are compared by birth cohort instead of age group. Birth cohorts that were younger

Table 3. Relationship between lung cancer mortality in 1980-84 and socio-demographic variables around 1970

Sex Variable <sup>b</sup>	Regression coefficient <sup>a</sup>			
	45-64	65-74	75+	all 45+
<b>Females</b>				
Average income per capita (in Dfl. 1000)	0.15	0.17*	0.19*	0.16**
Proportion of population in large cities	0.66**	0.20	0.28	0.39**
Proportion of population Roman-catholic	0.23	0.11	0.29*	0.20**
% variance explained	60**	40**	46**	72**
<b>Males</b>				
Average income per capita (in Dfl. 1000)	-0.12**	0.01	0.17**	0.02
Proportion of population in large cities	0.24**	0.14	0.21	0.19*
Proportion of population Roman-catholic	-0.01	0.24**	0.48**	0.23**
% variance explained	38**	34**	64**	50**

<sup>a</sup>From multiple regression analysis of ln (SMR) on three socio-demographic variables, weighted by population size,  $N = 39$  regions.

<sup>b</sup>Data are neither sex- nor age-specific but apply to the total regional population.

than 75 year in 1980-84 are excluded because of large fluctuations in regional numbers of deaths in 1950-54. Correlations are again low for females, but for males correlations are strong, which indicates stable mortality patterns within male birth cohorts. Mortality patterns for 1960-64 and 1970-74 were also found to be strongly correlated with those of 1980-84 (results not shown here).

The geographic mortality patterns in 1980-84 are reflected in the relationships with socio-demographic variables (Table 3). Among most birth cohorts, lung cancer mortality is positively related to high income and high degree of urbanisation (characteristic of central-western regions), and to percentage Roman Catholic (characteristic of south-eastern regions). For men 45-64 years, lung cancer is inversely instead of positively related to income.

Univariate correlations with smoking measures are given in Table 4. Lung cancer mortality among females has a moderately strong association with the

percentage of cigarette smokers in 1972. Associations are stronger with the percentage of ever-smokers of cigarettes, and with the average number of cigarettes smoked. Unexpectedly low correlations are however found for men, in particular for men younger than 75 years.

In almost all age and sex groups, lung cancer mortality is more strongly related to  $SO_2$  density than to any single variable on smoking (Table 5). Mortality among men at working age is significantly related to the percentage of men working in the sector "transport and manufacturing industry".

Table 6 presents the outcomes of multiple regression analyses of mortality on smoking, work in transport and manufacturing industry, and air contamination. The average number of cigarettes smoked was selected because this measure had the strongest relationships with mortality. Similar results were obtained for physical-environmental variables when control was made for other smoking variables.

Significantly positive relationships with the average number of cigarettes smoked are only found for women. Work in transport and manufacturing industry is still significantly related to lung cancer mortality among men 45-64 years.  $SO_2$  density is positively associated with mortality among older birth cohorts. Sex differences in the  $SO_2$  regression estimates can be due to random fluctuations. The average regression estimate for older birth cohorts (about 0.015) implies a 24% mortality difference between less contaminated regions (one standard deviation or  $8 \mu g SO_2/m^3$  below the national average of  $22 \mu g/m^3$ ) and more contaminated regions (one standard deviation above national average).

Individual-level studies showed that only a part of the occupations in transport and manufacturing industry may be associated with an increased risk of dying from lung cancer [20]. In additional analyses, these occupations were related individually to regional mortality differences. No significantly positive relationships were found for any occupational category.

## DISCUSSION

The aim of this analysis was to find clues to the explanation of regional differences in lung cancer mortality in the Netherlands. In the first step of the analysis, some clues for the role of smoking were obtained from a detailed description of regional mortality differences.

It was found that regional mortality patterns are strongly determined by cohort effects, at least among men: regional mortality patterns vary between birth cohorts, and have been stable for over 30 years within the older birth cohorts. This cohort effect was associated to income: for most birth cohorts, mortality levels are positively related to income level. But an inverse relationship is found for younger male co-

Table 4. Relationship between lung cancer mortality in 1980-84 and smoking status in 1972, by sex- and age-group

Sex Smoking variable <sup>b</sup>	Product-moment correlation <sup>a</sup>			
	45-64	65-74	75+	all 45+
<b>Females</b>				
% smokers of cigarettes	0.46**	0.23	0.34*	0.41**
% ever-smokers of cigarettes	0.63**	0.37*	0.50**	0.58**
Average number of cigarettes smoked per respondent	0.65**	0.31	0.33	0.63**
<b>Males</b>				
% smokers of cigarettes	0.18	0.14	0.38*	0.20
% ever-smokers of cigarettes	0.05	0.18	0.41**	0.29
Average number of cigarettes smoked per respondent	-0.10	0.09	0.37*	0.13

<sup>a</sup>With ln (SMR) weighted by population size,  $N = 39$  regions. \* $P < 0.05$ ; \*\* $P < 0.01$ .

<sup>b</sup>Data are specific to sex and birth cohort.

Table 5. Relationship between lung cancer mortality in 1980-84 and physical-environmental variables, by sex and age-group

Sex Variables	Product-moment correlation <sup>a</sup>			
	45-64	65-74	75+	all 45+
<b>Females</b>				
Proportion working in transport and manufacturing industry, 1971 <sup>c</sup>	0.25	0.21	0.15	0.25
50 percentile SO <sub>2</sub> (μg/m <sup>3</sup> ), 1979-80	0.54**	0.59**	0.60**	0.69**
<b>Males</b>				
Proportion working in transport and manufacturing industry, 1971 <sup>c</sup>	0.35*	0.17	0.01	0.12
50 percentile SO <sub>2</sub> (μg/m <sup>3</sup> ), 1979-80	-0.07	0.52**	0.75**	0.65**

<sup>a</sup>With ln (SMR), weighted by population size, *N* = 39 regions. \**P* < 0.05; \*\**P* < 0.01.

<sup>b</sup>Data on work in transport and manufacturing industry are specific to sex and (for men) birth cohort.

<sup>c</sup>Major group 7-9 of the International Classification of Occupations of 1968 (Dutch version).

horts, the only cohorts that are in a later phase of the lung cancer epidemic [21]. This suggests that during the lung cancer epidemic, a diffusion takes place from regions with a high to regions with a low socio-economic level. Biersteker estimated that mortality trends among men in rural, low-income provinces in the Netherlands lag about 10 years behind the trends in urban, high-income provinces [29]. It is likely that regional differences in (quitting) cigarette smoking are responsible for the cohort effects observed here and in particular for the diffusion of the lung cancer epidemic from regions with a high to regions with a low socio-economic status.

The strong relationship with the most 'cultural' of the socio-demographic variables, percentage Roman Catholic, is another indication that regional differences in a behavioural variable like tobacco consumption, and not in physical-environmental variables, are largely responsible for the regional mortality differences in the Netherlands.

For women, mortality rates in 1980-84 were indeed correlated with smoking ten years before. But the findings for men are surprising: not only are relationships weak, for men under 65 years of age they are

even in the wrong direction. How can this anomalous finding be explained?

The possible role of data problems should first be addressed. Regional differences in the quality of declaration of causes of death may distort the observation of regional differences in cause-specific mortality. However, we have previously shown that within the Netherlands there are no large regional differences in the quality of the declaration of causes of death (as measured by the percentage of deaths with the cause ill defined) [30]. Moreover, lung cancer is relatively easy to diagnose (as compared to many other forms of cancer) and is rarely confounded with secondary causes of death. An indication that even for men 75 years and over problems with cause-of-death registration have little effect, is that regional mortality differences among men 75 and over in 1980-84 are strongly correlated with the mortality differences among them in 1950-54, when they were younger than 64 years.

A problem relating to smoking data is random error in smoking prevalence estimates because of small regional sample sizes. This data problem cannot explain, however, that weak correlations were found

Table 6. Relationship between lung cancer mortality in 1980-84 and three possible determinants: outcome of regression analyses, by sex and age-group

Sex Variable	Regression coefficient <sup>a</sup>			
	45-64	65-74	75+	all 45+
<b>Females</b>				
Average number of cigarettes smoked per respondent, 1972	0.149**	0.009	0.054	0.101*
Proportion working in transport and manufacturing industry, 1971	4.27	4.46	0.17	3.11
50 percentile SO <sub>2</sub> (μg/m <sup>3</sup> ), 1979-80	0.011	0.017**	0.020**	0.014**
% variance explained	49	36	37	57
<b>Males</b>				
Average number of cigarettes smoked per respondent, 1972	-0.015	-0.002	0.029	-0.004
Proportion working in transport and manufacturing industry, 1971	0.68**	0.25	-0.34	0.18
50 percentile SO <sub>2</sub> (μg/m <sup>3</sup> ), 1979-80	-0.001	0.008**	0.019**	0.009**
% variance explained	17	28	60	42

<sup>a</sup>From multiple regression analysis of ln (SMR) on the three mentioned variables, weighted by population size, *N* = 39 regions. \**P* < 0.05; \*\**P* < 0.01.

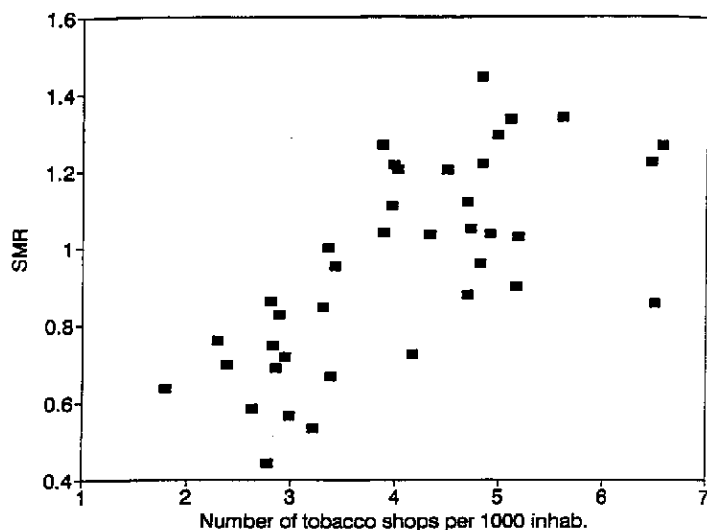


Fig. 2. Relationship of sale of tobacco in 1930 with mortality rates in 1980-84 among men 75 years and older.

for men but not for women. Another problem relating to smoking data is underreporting of the amount of cigarettes smoked, which has been estimated to be about 23% in the whole country [22]. If the degree of underreporting varied largely by region, this would have led to biased estimates of regional differences in cigarette consumption. However, this cannot explain that the same 'negative' results were obtained with the percentage of respondents currently smoking cigarettes, a measure that is much less susceptible to underreporting of tobacco consumption.

The most likely explanation is that the available information on smoking, although more extensive than the data used in other regional studies, did not cover several relevant aspects of smoking, such as life-time exposure, age at beginning, type of tobacco, number of puffs per cigarette, inhalation practise, and length of cigarette butt. It would have been particularly valuable to possess data on smoking earlier in life, since the stability in regional mortality patterns over the last 30 years suggests that regional mortality variation among men in 1980-84 reflect differences in smoking in the 1940s or before.

In order to have some indication of former regional differences in smoking, we acquired data on the number of tobacco shops per region in 1930 [31]. The number of shops per 1000 population might be a very approximate indicator of the amount of tobacco smoked by adult men around 1930. As illustrated in Fig. 2, this indicator is strongly related to death rates among these men in 1980-84 (when they were 75 years and older). The correlation is 0.67. This smoking measure suffers of course from a number of limitations, which may be expected to have caused a bias toward weaker relationships. It is therefore reasonable to expect that a more accurate measure would show even more clearly that regional mortality

variation among men in 1980-84 reflect differences in smoking various decades before.

The strong influence of tobacco consumption various decades before can be understood from several factors:

- an independent effect of the age at starting with smoking on lung cancer occurrence [13, 14];
- a higher tar-content of cigarettes in former years [13, 14];
- a high prevalence of smoking among boys in the Netherlands in the first half of the 20th century [26]. For example, in the period 1907-1916 about 75% of boys in the age group 11-12 smoked cigarettes, some of the boys smoking several cigarettes per day [32].

The dominant effect of smoking in early life on regional mortality differences in 1980-84 can explain why these differences are hardly related to smoking in 1972. The correlation between the smoking indicator for 1930 and the mean number of cigarettes smoked in 1972 (by men 75 years and over) is very weak: 0.23. This implies that the regional pattern of tobacco consumption has considerably changed. Further insights in these changes are provided by Table 7. While in 1930 smoking was positively related to income level, in 1972 the relationship was inverse. Although men in lower-than-average income regions now smoke more, lung cancer mortality among them is not higher, because of having smoked less in early life.

Regional variation in lung cancer mortality is relevant to the study of the causes of lung cancer, because this variation may be reflective of other factors than tobacco consumption. The paper has an important message for etiological studies using ecological data, in showing that controlling for current

Table 7. Relationship between socio-demographic variables around 1950 and smoking in 1930 and 1972 respectively,<sup>a</sup> birth cohort of men aged 75+ years in 1980-84

Socio-demographic variable <sup>b</sup>	Standardized regression coefficient <sup>c</sup>	
	Smoking in 1930	Smoking in 1972
Average income per capita	0.88**	-0.11
Proportion of population in large cities	-0.25	0.54*
Proportion of population Roman-catholic	0.29*	0.11
% variance explained	42**	23**

<sup>a</sup>Measure for 1930: number of tobacco shops or other shops selling tobacco per 1000 population. For 1972: average number of cigarettes smoked by men 65+ years.

<sup>b</sup>Socio-demographic data apply to the entire regional population.

<sup>c</sup>From multiple regression analysis of the respective smoking measure on three socio-demographic variables, weighted by population size,  $N = 39$  regions. \* $P < 0.05$ ; \*\* $P < 0.01$ .

levels of smoking only (which is often done) may not only be imperfect, but even superfluous.

Regional differences in work in transport and manufacturing industry contributed to a small degree towards the explanation of regional differences in lung cancer mortality among men of working age. This agrees with the associations found in French and American studies [3, 33]. No relationships could be demonstrated for more specific occupational categories. Apparently, workers in specific occupations have such a small share in regional populations, and the eventual mortality effects associated with these occupations are probably so modest, that these effects could not be discerned in this regional analyses.

The strong relationships that were found with air contamination, among men and women older than 65 years, might be attributable to inadequate control for cigarette smoking earlier in life. An approximate control for smoking history can be made by using proxies for smoking in earlier years, such as:

—the number of cigarette shops per capita in 1930;  
—socio-economic status and urbanisation. Controlling for socio-demographic variables carries however the risk of *over* control, because e.g. rural-urban mortality differences could be partly due to air contamination instead of cigarette smoking.

Notwithstanding eventual overcontrol, additional analysis showed that the relationships between air contamination and lung cancer mortality were still statistically significant after control for either proxy of smoking history.

The results of this study are in contrast to the findings of Chinn *et al.* for England and Wales [34]. Controlling for a number of socio-demographic variables,  $SO_2$  and smoke levels were found to be associated with regional lung cancer death rates in the 1950s and 1960s, but not in the period 1969-1973. A possible explanation is that the U.K. clean air legislation might have strongly diminished previous effects of air contamination on regional lung cancer death rates. An alternative explanation that is more in line with the results of our analysis refers to the diffusion of the lung cancer epidemic, which occurred in England and Wales as well as in the Netherlands [35]. Perhaps this diffusion process has been such that

it caused an artificial association with air contamination in the case of the Netherlands, or that it obscured an association in the English case.

Time-series analyses have demonstrated short-term effects of episodes of air contamination on mortality in the general population, even below established air contamination standards [36, 37]. In contrast, studies on the effects of long-term exposure to low levels of air contamination have been handicapped by several methodological problems and have come to inconclusive results [18, 19, 38, 39]. The fact, however, that in various countries lung cancer death rates are highest in heavily industrialised areas, is suggestive of a relationship with air pollution [3-6]. Our study showed that this seeming relationship cannot be easily explained away by confounding with tobacco consumption. This study thus yielded new indications of a long-term effect of air pollution on lung cancer mortality. However, the evidence is still weak and should primarily be considered as a stimulus to carry out further analysis on the excess lung cancer mortality in polluted regions.

*Acknowledgements*—This study was supported by a grant from the Ministry of Welfare, Public Health and Culture, Rijswijk, the Netherlands. The mortality data were supplied by the Dutch Central Bureau of Statistics. Survey data on smoking have been made available by the Steinmetz Archives at Amsterdam (data base no. PO351), with the generous permission of the owner of the data, the Verenigde Nederlandse Uitgeversbedrijven.

## REFERENCES

1. Blot W. J. and Fraumeni J. F. Changing patterns of lung cancer in the United States. *Am. J. Epidemiol.* **115**, 664, 1982.
2. Howe G. M. Mortality from selected malignant neoplasms in the British Isles: the spatial perspective. *Soc. Sci. Med.* **15D**, 199, 1981.
3. Caselli G. Les causes des décès en France III. Un effort d'interprétation des différences géographiques: application à la période 1974-76. *Population* **6**, 1011, 1984.
4. Gatzweiler H. P. and Stiens G. Regionale mortalitätsunterschiede in der Bundesrepublik Deutschland. Daten und hypothesen. *Jahrbuch für Regionalwissenschaft* **3**, 36, 1982.
5. Grosclaude A., Lux B., Van Houte-Minet M. and Wunsch G. Mortalité régionale et comportements différentiels. *Pop. Fam.* **48**, 1, 1979.



6. Caselli G. and Egidi V. Géographie de la mortalité en Europe: influence de l'environnement et de certains aspects du comportement. In *Proceeding of the International Population Congress*, Manilla, 1981, Vol. 2, p. 165. IUSSP, Liège, 1983.
7. Wagner R. I. and Merabishvili V. M. Geographical distribution of lung cancer in the U.S.S.R. *Int. J. Cancer* **34**, 67, 1984.
8. Minowa M., Shigematsu I., Nagai M. and Fukutomi K. Geographical distribution of lung cancer mortality and environmental factors in Japan. *Soc. Sci. Med.* **15D**, 225, 1981.
9. Cleston F. and Coebergh J. W. W. *Cancer in the Netherlands: Scenarios on Cancer*, pp. 1985-2000. Kluwer, Dordrecht, 1988.
10. Murata M., Takayama K., Fukuma S. et al. A comparative epidemiologic study on geographic distributions of cancers of the lung and the large intestine in Japan. *Jap. J. Cancer Res.* **79**, 1005, 1988.
11. Mohner M. Analyse der geographischen Verteilung von Krebsinzidenzen am Beispiel der Lungentumoren in der DDR den Zeitraum 1978-1982. *Arch. Geschwulstforsch.* **58**, 191, 1988.
12. WHO Expert Committee on Smoking Control. Controlling the smoking epidemic. Technical Report Series no. 636. WHO, Genève, 1979.
13. U.S. Department of Health and Human Services. The health consequences of smoking. Cancer. A report of the Surgeon General. USDHHS, Office on Smoking and Health. Rockville (Maryland), 1982.
14. Royal College of Physicians. *Health or smoking? Follow-up report of the Royal College of Physicians*. Pitmans, London, 1983.
15. Weinberg G. B., Kuller L. H. and Redmond C. K. The relationship between the geographic distribution of lung cancer incidence and cigarette smoking in Allegheny county, Pennsylvania. *Am. J. Epidemiol.* **115**, 40, 1982.
16. Valkonen T. and Notkola V. Influence of socioeconomic and other factors on the geographic variation of mortality in Finland, Sweden and Norway. *Yearbook of Population Research in Finland* **15**, 9, 1977.
17. Taioli E., Nicolosi A. and Wynder E. L. Possible role of diet as a host factor in the aetiology of tobacco-induced lung cancer: an ecological study in Southern and Northern Italy. *Int. J. Epidemiol.* **20**, 611, 1991.
18. Higgins I. T. T. Air pollution and lung cancer: diesel exhaust, coal combustion. *Prev. Med.* **13**, 207, 1984.
19. Speizer F. E. Overview of the risk of respiratory cancer from airborne contaminants. *Environmental Hlth Persp.* **70**, 9, 1986.
20. Dubrow R. and Wegman D. H. Setting priorities for occupational cancer research and control: synthesis of the results of occupational disease surveillance studies. *J.N.C.I.* **71**, 1123, 1983.
21. Hoff N. M. van der. Cohort analysis of lung cancer in the Netherlands. *Int. J. Epidemiol.* **8**, 41, 1979.
22. Reek J. van. Rookgedrag in Nederland van 1958-1982. *Tijdschr. Alcohol Drugs* **9**, 99, 1983.
23. Mackenbach J. P., Kunst A. E. and Looman C. W. N. Cultural and socioeconomic determinants of geographical mortality patterns in the Netherlands. *J. Epidemiol. Commun. Hlth* **45**, 231, 1991.
24. Hofstee E. W. *Short Demographic History of The Netherlands from 1980 to the Present*. Fibula-van Dishoeck, Haarlem, 1981.
25. Van Engelsdorp Gastelaars R. E., Ostenhof W. J. M. and Vos S. de. Typologieën van Nederlandse gemeenten naar stedelijkheidsgraad. Monographs Population Census 1971. Staatsuitgeverij, The Hague, 1980.
26. Gadourek I. Riskante gewoonten en zorg voor eigen welzijn (dissertation). University of Groningen. Groningen, 1963.
27. National Institute for Public Health and Environmental Hygiene. Landelijk meetnet luchtkwaliteit: verslag over de periode 1 April 1979-1 April 1980. Report no. 69/80 LMO, Series NML-RIV no. 18. National Institute for Public Health. Bilthoven, 1980.
28. Gail M. The analysis of heterogeneity for indirect standardized mortality ratios. *J. R. Stat. Soc. A* **141**, 224, 1978.
29. Biersteker K. Verontreinigde lucht (dissertation). University of Amsterdam. Amsterdam, 1966.
30. Mackenbach J. P. Mortality and medical care (dissertation). Erasmus University Rotterdam. Rotterdam, 1988.
31. Central Bureau of Statistics. Bedrijfstelling 31 december 1930. Algemene Landsdrukkerij. The Hague, 1935.
32. Proosdij C. van. Roken, een individueel- en sociaalgeeneeskundige studie (dissertation). University of Amsterdam, Amsterdam, 1957.
33. Blot W. J. and Fraumeni J. F. Geographic patterns of lung cancer: industrial correlations. *Am. J. Epidemiol.* **103**, 539, 1976.
34. Chinn S., Florey C. du V., Baldwin I. G. and Gorgol M. The relation of mortality in England and Wales to measurements of air contamination. *J. Epidemiol. Commun. Hlth* **35**, 174, 1981.
35. Swedlow A. J. and dos Santos Silva I. Geographic distribution of lung and stomach cancer in England and Wales over 50 years: changing and unchanging patterns. *Br. J. Cancer* **63**, 773, 1991.
36. Derriennic F., Richardson S., Mollie A. and Lellough J. Short-term effects of sulphur dioxide pollution on mortality in two French cities. *Int. J. Epidemiol.* **18**, 186, 1989.
37. Schwartz J. and Marcus A. Mortality and air pollution in London: a time series analysis. *Am. J. Epidemiol.* **131**, 185, 1990.
38. Krzyzanowski M. and Wojtyniak B. Ten-year mortality in a sample of an adult population in relation to air pollution. *J. Epidemiol. Commun. Hlth* **36**, 262, 1982.
39. Brown L. M., Pottern L. M. and Blot W. J. Lung cancer in relation to environmental pollutants emitted from industrial sources. *Environmental Res.* **34**, 250, 1984.