A Comparative Epidemiologic Study on Geographic Distributions of Cancers of the Lung and the Large Intestine in Japan

Motoi MURATA,^{*1} Kimiko TAKAYAMA,^{*1} Seigo FUKUMA,^{*2} Naoyuki OKAMOTO,^{*3} Ikuko KATO,^{*4} Aya HANAI,^{*5} Hideaki NAKAYAMA,^{*6} Kouichi FUJIWARA,^{*7} Takayoshi IKEDA^{*8} and Isaburo FUJIMOTO^{*5}

^{*1}*Division of Epidemiology and* ^{*2}*Emeritus President, Chiba Cancer Center, Nitona 666-2, Chiba 280,* ^{*3}*Department of Epidemiology, Clinical Research Institute, Kanagawa Cancer Center, Yokohama 241,* ^{*4}*Division of Epidemiology, Aichi Cancer Center Research Institute, Nagoya 464,* ^{*5}*Department of Field Research, Osaka Center for Adult Diseases, Osaka 535,* ^{*6}*Department of Hygiene, Tottori University, Tottori 683,* ^{*7}*Ministry of Health and Environment, Saga Prefecture, Saga 840 and* ^{*8}*Department of Pathology, Nagasaki University School of Medicine, Nagasaki 852*

To examine what kinds of factors could have caused the geographic variation observed in lung cancer morbidity in Japan, a correlation study was performed comparing various regional traits. The same study was also conducted on large intestinal cancer, aiming to distinguish the possible urban factors associated with both cancers. Lung cancer was highly correlated with industrialization-related factors such as localization of manufacturing industries, automobile traffic and air pollution, whereas colon cancer was correlated with the population density of workers in the tertiary industries such as services, trade and government. A multiple regression analysis could not detect any single factor with an exceptionally strong influence on either cancer. The present findings suggest that the hazardous environmental condition of urban areas has, to some extent, contributed to the recent increase of lung cancer cases in this country.

Key words: Lung cancer — Colon cancer — Cancer registry — Geographic correlation — Air pollution

In Japan lung cancer has been rapidly increasing in recent years. The age-adjusted mortality rates were 27.4 in males and 7.7 in females per 10⁵ population in 1985, nearly five times and four times larger, respectively, than those in 1955. It is generally suspected that the growing tobacco consumption in these decades has made a significant contribution to the past increase of lung cancer.¹⁾ However, the incidence rate also shows a notable geographic variation which mainly corresponds to the urban-rural difference. If all geographic districts are classified by their population size, the lung cancer mortality is highest in the largest cities and lowest in the rural areas. This urban-rural difference may be, as has been speculated from epidemiologic findings reported in other countries,²⁾ ascribable to differential tobacco consumption rates. Nevertheless, it is quite improbable that other kinds of hazardous factors in the urban environment have nothing to do with this disease.

Several authors in other countries have reported that, even after age and smoking rate

were standardized, a significant effect of residence still appeared in the geographic variation of lung cancer mortality.³⁻⁵⁾ The effect of residence could be due to either some feature(s) of urban life style or environmental pollution. Regarding the association of these latter factors with lung cancer, many epidemiologic studies⁶⁻¹¹⁾ have been carried out and provided apparently positive but inconclusive evidence because other factors such as occupation and smoking rate were not thoroughly standardized.¹²⁻¹³⁾

The present study attempted to obtain some clues concerning the geographic correlation between lung cancer morbidity and those factors inherent to urbanization, especially environmental pollution or localization of its sources. In general, this type of correlation study is rather inferior to the more advanced analytical epidemiologic studies, because it can only demonstrate a correlation between the disease and regional traits. Still it should be useful if the factors to be studied are of area-specific rather than person-specific kinds.

It is especially advantageous when sources of pollution are multiple. In order to check the adequacy of the correlation study, large intestinal cancer was also studied in parallel with lung cancer. It is reasonable to say that large intestinal cancer, which has also been rapidly increasing in recent years and is highly prevalent in urban areas, should at least not be associated with air pollution. If factors correlated with the morbidity of both cancers are not similar, it would provide some basis for speculation about possible causes of the increase of lung cancer in urban areas.

MATERIALS AND METHODS

This study was accomplished with the cooperation of the cancer registries of 11 prefectures and one city as follows; Miyagi, Yamagata, Chiba, Kanagawa, Aichi, Osaka, Hyogo, Okayama, Tottori, Fukuoka and Saga prefectures and Nagasaki city. Each prefecture was divided into small districts, such as cities and counties or cities and towns. A large city was further divided into wards. The total number of districts is thus 583, and their population size is quite variable, ranging from less than 1,000 (only one district) to over one million (only one district). Data on their cancer morbidity and mortality rates and on various epidemiologic traits were provided from individual registries.

1. Morbidity and mortality data: Age-adjusted incidence and death rates of cancers of the large intestine (colon and rectum) and the lung, standardized by Segi's world population, in 1975-1979.

2. Epidemiologic traits: a. Demographic traits (population density, etc.). b. Medical services (number of hospital beds, etc.). c. Socio-economic factors (proportion of employees by industry in a population, number of establishments (plants and agencies) of manufacturing industries per unit area, proportion of a population living on welfare aid, automobile traffic density, amount of municipal tax revenue per capita, etc.). d. Air pollution (SO_2 , NO_2 , HC, floating dust, etc.). e. Climatologic factors (temperature, rainfall, etc.). f. Food consumption (expenditure on food per household).

A summary of these 63 epidemiologic traits is given in Table I. They were all adopted for study because they were obtainable from official publications, such as annual reports on statistics of the prefectures or published reports,¹⁴⁾ and also because of their potential epidemiologic significance in relation to lung cancer. Population density was chosen because it usually reflects fairly well the urbanization level of the district. Proportion of a

population living on municipal welfare aid was adopted as an index of the socio-economic level of the district. Some of the past epidemiologic studies reported that people of a lower socioeconomic class tend to be more susceptible to lung cancer.¹⁵⁾ Indices of medical services were chosen with the aim of seeing whether the geographic variation of cancer morbidity rate might be ascribed to the level of medical service of the districts. Climatologic factors were included based on the speculation that air pollution might be influenced by the amount of rainfall. Information on the automobile traffic density was not available in any official publication. Thus, a traffic density index was tentatively calculated by multiplying the length of national roads by the amount of automobile traffic on the roads and dividing by the area. Data on other important factors such as regional smoking rate and dietary pattern were not available. The year of measurement for most of the epidemiologic traits was 1975.

It should be pointed out that the number of districts (n) was not constant in examining all traits since not all municipalities have a complete set of data (Table I). In particular, information on the expenditure on food was limited to only 4 prefectures. Moreover, as for the indices of air pollution, the survey locations are usually distributed more densely in industrialized districts. In Table I, the 583 districts were divided into three classes by their population density, namely, less than 1,000, 1,000 to 3,000 and over 3,000 per square kilometer. We tentatively call them rural, suburban and urban areas, respectively. Average values of morbidity and mortality rates for lung and large intestinal cancers are shown in Table II.

The simple correlation coefficient (r) was calculated between the morbidity rate and various epidemiologic traits for the total districts and also separately for the three groups of districts with different urbanization levels. The correlation was also calculated for the mortality rate. Morbidity statistics in Japan are not necessarily reliable in some districts because completeness of registration may not be fully attained in some cancer registries. Only if the correlation coefficients for both morbidity and mortality rates are statistically significant, do we consider a relationship to be definitely indicated. The influence of incomplete registration will be discussed in a later section in more detail. Measurements for an epidemiologic trait per unit population or per unit area usually do not follow the normal distribution. In such a case, logarithmic transformation was performed before the calculation of the correlation. The morbidity and mortality rates also underwent the same transformation.

Subsequently we tried a multivariate analysis in order to see if any one of the epidemiologic traits shows a distinctively strong association with the

Table I. Number of Districts in which Information on Each Trait Is Available

Study items	Total	Urban	Suburban	Rural
1 Population density	583	115	116	352
2 Number of households	582	115	116	351
3 Living on welfare aid	422	105	85	232
4 Municipal tax revenue	409	102	85	222
5 Tobacco tax revenue	409	102	85	222
6 Number of hospitals	556	114	115	327
7 Number of hospital beds	556	114	115	327
8 Agriculture & fisheries	583	115	116	352
9 Mining	585	115	116	352
10 Construction	583	115	116	352
11 Manufacturing	583	115	116	352
12 Wholesale & retail trade	583	115	116	352
13 Finance & insurance	577	115	115	347
14 Real estate	583	115	116	352
15 Transport & communication	582	115	115	352
16 Electricity, gas and water	574	115	115	344
17 Services	582	115	115	352
18 Government	581	114	115	352
19 Foods	499	105	106	288
20 Textiles	459	105	102	252
21 Fabrics	473	105	103	265
22 Wood	495	105	103	287
23 Furniture	482	105	104	273
24 Pulp & paper	448	105	99	244
25 Printing & publishing	473	105	106	262
26 Chemicals	447	104	101	242
27 Petroleum & coal	397	99	82	216
28 Rubber	396	99	81	216
29 Leather	397	99	82	216
30 Glass & cement	502	105	106	291
31 Iron & steel	502	105	106	291
32 Other metals	397	99	82	216
33 Metal working	502	105	106	291
34 General machinery	502	105	106	291
35 Electric machinery	502	105	106	291
36 Transport machinery	502	105	106	291
37 Precision machinery	397	99	82	216
38 Other manufacturing	398	99	82	217
39 SO ₂	216	86	68	62
40 NO ₂	179	76	58	45
41 HC	98	36	43	19
42 Floating dust	182	74	59	49
43 Falling soot particles	147	63	42	42
44 Traffic density index	347	79	75	193
45 Rainfall	268	94	41	133
46 Mean temperature	243	93	38	112
47 Duration of sunshine	243	93	38	112
48 Natural radioactivity	238	92	59	87
49 Rice	136	25	37	74
50 Bread	136	25	37	74
51 Other main foods	136	25	37	74
52 Fresh fishes	136	25	37	74
53 Salted & dried fishes	136	25	37	74
54 Meats	136	25	37	74
55 Milk & eggs	136	25	37	74
56 Vegetables	136	25	37	74
57 Seaweeds & dry foods	136	25	37	74
58 Preserved foods	136	25	37	74
59 Condiments	136	25	37	74
60 Cakes & candies	136	25	37	74
61 Fruits	136	25	37	74
62 Alcoholic beverages	136	25	37	74
63 Other beverages	136	25	37	74

1-2: Density per km². 3: Number in a unit population. 4-7: Number or amount per capita. 8-18: Proportion of workers by industry classes. 19-38: Number of establishments per km² for manufacturing industries. 39-43: Air pollution index. 44: Automobile traffic per km². 45-48: Climatologic & geographic traits. 49-63: Monthly expenditure for food per household.

Table II. Age-adjusted Incidence (I) and Death Rates (D) of Lung Cancer and Large Intestinal Cancer in Areas with Different Levels of Urbanization (/10⁵)

		Total		Urban		Suburban		Rural	
		I	D	I	D	I	D	I	D
Lung	M	22.9	21.5	27.5	25.0	21.8	19.7	21.7	21.0
	F	6.8	6.6	8.2	7.7	6.7	6.9	6.4	6.2
Colon	M	7.8	6.0	8.6	6.4	7.8	6.2	7.6	5.8
	F	6.7	5.2	6.6	5.1	6.7	5.4	6.8	5.1
Rectum	M	8.2	6.7	8.8	6.5	7.6	6.7	8.1	6.9
	F	5.6	4.3	5.0	3.5	5.0	4.5	6.2	4.6
Large int.	M	16.0	12.7	17.4	12.9	15.4	12.9	15.7	12.7
	F	12.3	9.5	11.6	8.6	11.7	9.9	13.0	9.7

lung cancer morbidity. After choosing a certain number of traits which seemed important from the simple correlation coefficient values, a multiple regression analysis and a factor analysis were both performed.¹⁶⁾ Both analyses were applied only in the districts where information on all of the selected traits was available, especially in the urban and suburban areas. For the multiple regression analysis, the following linear relation is assumed;

$$y = b_0 + b_1x_1 + b_2x_2 + \dots + b_jx_j$$

where y is the cancer morbidity or mortality rate, x_i is a measurement of the i -th factor, and b_i is a partial regression coefficient of y on x_i . The regression coefficients are estimated by the least-squares method. The procedures for the factor analysis were briefly described by Yanai *et al.*¹⁷⁾ It assumes the following linear relation between individual basic factors, x and y inclusively, and the j -th hypothetical general factor, z :

$$x = a_1z_1 + a_2z_2 + \dots + a_jz_j$$

where the coefficient, a_j , stands for a factor loading or a regression coefficient of individual x on z_j and represents a contributing rate of z_j on normalized x . The z_j may be interpreted as a hypothetical factor that represents the overall variation observed in the basic factors.

RESULTS

Correlation Coefficient The results of the statistical test of the simple correlation coefficients obtained for the total population and for the populations with the three classes of urbanization are summarized in Tables III and IV. When the morbidity and mortality rates are both significantly correlated at the 1% level with the epidemiologic trait, then a mark, * or #, is assigned. The former is for

positive and the latter is for negative correlation. For reference, borderline values of correlation coefficient for the 1% significance level are given in a statistical table for different values of n .¹⁸⁾ They are, e.g. 0.449 ($n=30$), 0.354 ($n=50$), 0.254 ($n=100$), 0.181 ($n=200$) and 0.115 ($n=500$). Though the actual values of coefficients (r) are not given in the tables, they are rather small simply because of the large variances in both variables, i.e. in the morbidity and mortality rates and in the epidemiologic traits. For instance, the morbidity rate of male lung cancer showed a correlation coefficient of 0.176 ($n=583$) with the population density, 0.322 ($n=502$) with the number of establishments of the iron and steel industry per unit area, and 0.286 ($n=179$) with the concentration of NO_2 , respectively, in the total population. Several examples of scatter diagrams are shown in Fig. 1.

The results in Tables III and IV may be summarized as follows.

- 1) As expected, colon and lung cancers are both correlated positively with the population density and negatively with the number of agricultural workers.
- 2) Lung cancer is prevalent in those districts where the establishments of any kind of manufacturing industry are densely located, with heavier automobile traffic and with more polluted air. The proportion of a population living on welfare aid is also positively correlated with lung cancer. These tendencies are stronger in the urban as compared with

Table III. Result of Significance Test for Correlation Coefficients^{a)} between the Epidemiologic Factors and Male Cancer Rates

Study items ^{b)}	Lung				Colon				Rectum				Large intestine			
	T	U	S	R	T	U	S	R	T	U	S	R	T	U	S	R
1 Population density		*			*											
2 Number of households		*			*											
3 Living on welfare aid	*		*													
4 Municipal tax revenue																
5 Tobacco tax revenue																
6 Number of hospitals									*	*			*	*		
7 Number of hospital beds		*			*											
8 Agriculture & fisheries	#	#			#											
9 Mining																
10 Construction							*									
11 Manufacturing			#		#	#	#		#				#	#	#	
12 Wholesale & retail trade	*				*		*								*	
13 Finance & insurance																
14 Real estate	*															
15 Transport & communication																
16 Electricity, gas and water																
17 Services			*		*	*	*						*	*		
18 Government		#					*								*	
19 Foods	*															
20 Textiles	*						#									
21 Fabrics	*	*														
22 Wood	*	*														
23 Furniture	*				*		*									*
24 Pulp & paper	*	*			*											
25 Printing & publishing	*	*			*											
26 Chemicals	*	*			*											
27 Petroleum & coal	*	*			*											
28 Rubber	*	*			*											
29 Leather	*				*											
30 Glass & cement	*															
31 Iron & steel	*	*														
32 Other metals	*	*			*											
33 Metal working	*	*														
34 General machinery	*	*	#													
35 Electric machinery	*	*														
36 Transport machinery	*	*					#									
37 Precision machinery	*	*			*											
38 Other manufacturing	*	*	#		*											
39 SO ₂	*															
40 NO ₂	*															
41 HC																
42 Floating dust	*	*														
43 Falling soot particles					*	*			#				*		*	
44 Traffic density index		*			*								*			
45 Rainfall		#														
46 Mean temperature	*				*	*									*	
47 Duration of sunshine																
48 Natural radioactivity													*			
49 Rice																
50 Bread																
51 Other main foods																
52 Fresh fishes																
53 Salted & dried fishes																
54 Meats					*											
55 Milk & eggs																
56 Vegetables																
57 Seaweeds & dry foods																
58 Preserved foods				#												
59 Condiments																
60 Cakes & candies	#		#													
61 Fruits																
62 Alcoholic beverages																
63 Other beverages																

a) Correlations with morbidity and mortality rates are both statistically significant at the 1% level, in the positive (*) or negative (#) direction.

b) See footnotes of Table I.

Abbreviations; total (T), urban (U), suburban (S) and rural (R) districts.

Table IV. Result of Significance Test for Correlation Coefficients^{a)} between the Epidemiologic Factors and Female Cancer Rates

Study items ^{b)}	Lung				Colon				Rectum				Large intestine			
	T	U	S	R	T	U	S	R	T	U	S	R	T	U	S	R
1 Population density	*								#							
2 Number of households	*															
3 Living on welfare aid																
4 Municipal tax revenue																
5 Tobacco tax revenue																
6 Number of hospitals	*			*	*				*							
7 Number of hospital beds		*														
8 Agriculture & fisheries	#	#							*							
9 Mining							*		*				*		*	
10 Construction							*						*		*	
11 Manufacturing					#	#	#		#	#			#	#	#	
12 Wholesale & retail trade									#							
13 Finance & insurance									#				#			
14 Real estate									#				#			
15 Transport & communication																
16 Electricity, gas and water									#							
17 Services																
18 Government				*	*	*	*	*	*	*	*	*	*	*	*	*
19 Foods		*														
20 Textiles																
21 Fabrics	*															
22 Wood									#							
23 Furniture	*															
24 Pulp & paper	*	*														
25 Printing & publishing	*	*														
26 Chemicals	*															
27 Petroleum & coal	*															
28 Rubber	*															
29 Leather	*															
30 Glass & steel	*															
31 Iron & steel	*															
32 Other metals	*															
33 Metal working					#				#				#			
34 General machinery									#						#	
35 Electric machinery									#				#			
36 Transport machinery									#				#			
37 Precision machinery	*															
38 Other manufacturing		*							#							
39 SO ₂																
40 NO ₂	*															
41 HC																
42 Floating dust	*															
43 Falling soot particles									*						*	
44 Traffic density index	*															
45 Rainfall					*								*	*		
46 Mean temperature	*															
47 Duration of sunshine																
48 Natural radioactivity																
49 Rice																
50 Bread																
51 Other main foods																
52 Fresh fishes																
53 Salted & dried fishes																
54 Meats																
55 Milk & eggs																
56 Vegetables																
57 Seaweeds & dry foods																
58 Preserved foods																
59 Condiments																
60 Cakes & candies																
61 Fruits																
62 Alcoholic beverages																
63 Other beverages																

a) Correlations with morbidity and mortality rates are both statistically significant at the 1% level, in the positive (*) or negative (#) direction.

b) See footnotes of Table I.

Abbreviations; total (T), urban (U), suburban (S) and rural (R) districts.

Fig. 1. Correlation of male lung cancer morbidity rate with epidemiologic traits of the districts. Upper: number of establishments of the chemical industry ($n=447$, $r=0.230$). Lower: air pollution of NO_2 ($n=179$, $r=0.286$).

rural areas and also in males than in females. Tobacco expenditure tax revenue showed no correlation with the lung cancer rate.

3) The population density of workers in most industries appears to be irrelevant to lung cancer morbidity. However, if the analysis is confined to the urban area, workers, in trade

and realty are both positively correlated to it. 4) The yearly amount of rainfall is correlated negatively with lung cancer but positively with female colon cancers.

5) Colon cancer is prevalent in those districts where there are more workers in the tertiary industries such as sales, services and govern-

Table V. Partial Correlation Coefficients Estimated by a Multiple Regression Analysis Using 13 Selected Factors for Male Lung and Colon Cancer Incidence (I) and Death Rates (D)

Study items	Lung cancer		Colon cancer	
	I	D	I	D
1 Population density	0.235	0.166	0.155	-0.075
2 Living on welfare aid	0.214	0.137	0.117	0.021
3 Workers in manufacturing	-0.076	-0.029	0.044	0.015
4 Workers in trading	0.200	0.272*	-0.037	-0.014
5 Workers in service industries	-0.003	0.056	0.310*	0.234
6 Workers in government	0.003	0.007	-0.053	0.096
7 Establishments for wood manuf.	-0.203	-0.118	-0.023	-0.147
8 Establishments for pulp & paper	-0.350**	-0.244	0.168	0.156
9 Establishments for chemicals	0.403**	0.199	-0.059	0.020
10 Establishments for iron & steel	0.332**	0.218	-0.060	0.200
11 Air pollution of NO ₂	-0.149	0.024	-0.254*	-0.052
12 Air pollution of floating dust	0.170	0.095	0.087	-0.058
13 Automobile traffic density	-0.237	-0.263*	0.152	0.051

Number of districts (n) is 77. Correlations statistically significant at the 1% (**) or 5% (*) level are indicated.

ment, and less in the manufacturing industries. This tends to be true in both urban and suburban areas and in both males and females.

6) For almost all kinds of manufacturing industries, the number of establishments (density) seems to be irrelevant to colon cancer.

7) The mean expenditure for meat per household is positively correlated with male colon cancer.

8) There were few factors observed to be correlated with male rectal cancer, either positively or negatively, whereas for female rectal cancer, the correlated factors are rather similar to those of colon cancer.

Multivariate Analysis The result of a multiple regression analysis is shown in Table V. In view of the results shown in Tables III and IV, it was applied only for the male lung and male colon cancers. Districts in the rural area were excluded from this analysis. For this analysis the following 13 factors were chosen: the population density, the proportion living on welfare aid, the population density of employees in manufacturing, trading, services and government, number of establishments per unit area for woodworking, pulp and paper, chemicals, steel and metal working, air pollution indices such as NO₂ and floating dust, and automobile traffic density.

The multiple regression analysis did not reveal any factor having a significant partial correlation coefficient with both morbidity and mortality rates. For lung cancer, the partial correlation was in the positive direction with the number of establishments of the chemical industry and iron and steel industry and workers in trade, but in the negative direction with the number of establishments of the pulp and paper industry and automobile traffic density. For colon cancer, the positive correlation of service workers, along with the negative correlation of NO₂ pollution, confirms the above-mentioned observation 5) on simple correlation.

Table VI and Fig. 2 show, with factor loading values, the contributing relationships of the first, second and third general factors detected by the factor analysis to the variation observed in each of the 13 basic factors referred to above as well as in the morbidity and mortality of male colon and lung cancers in the urban and suburban areas. The first general factor probably represents industrialization in general. The second and third factors are individually ascribable to the variation of the male colon and male lung cancer rates. Thus, as is clearly seen in Fig. 2, male lung cancer rates are more prone to be grouped with most of the industrialization

Table VI. Factor Loading Values for the First (z_1), Second (z_2) and Third (z_3) General Factors Obtained by a Factor Analysis with Varimax Rotation

Study items	z_1	z_2	z_3
1 Population density (PD)	0.851	0.332	-0.408
2 Living on welfare aid (LA)	0.886	0.189	-0.423
3 Workers in manufacturing (Mo)	0.218	-0.688	0.692
4 Workers in trading (Sa)	0.543	0.606	-0.582
5 Workers in service industries (Se)	-0.130	0.796	-0.591
6 Workers in government (Go)	-0.924	0.380	-0.044
7 Establishments for wood manuf. (Wo)	0.992	0.037	-0.119
8 Establishments for pulp & paper (PP)	0.904	0.334	-0.267
9 Establishments for chemicals (Ch)	0.900	0.186	-0.394
10 Establishments for iron & steel (IS)	0.977	-0.192	-0.098
11 Air pollution of NO_2 (NO_2)	0.904	0.248	-0.348
12 Air pollution of floating dust (FD)	0.848	-0.520	0.102
13 Automobile traffic density (TD)	0.918	0.346	-0.192
14 Male lung cancer morbidity (LI)	0.510	-0.120	-0.852
15 Male lung cancer mortality (LD)	0.409	0.005	-0.913
16 Male colon cancer morbidity (CI)	0.266	0.962	0.054
17 Male colon cancer mortality (CD)	0.089	0.993	0.076

Fig. 2. Scatter diagram of factor loading values of the 1st and 2nd general factors for various epidemiologic traits. See Table VI for abbreviations.

factors as compared to male colon cancer rates. But there still are differences between the industrialization factors and lung cancer rates, suggesting that other factors may be

associated with the geographic variation of this cancer.

DISCUSSION

Prior to any interpretation of the present results, the quality of the data utilized should be considered. First of all, as previously stated, the completeness of registration was variable among different prefectures. Morbidity statistics which depend on registry data ought to be very much influenced by the level of its completeness. Although presently no good way of examining this influence is known, the ratio of morbidity to mortality rates might be an index of completeness. This ratio for lung cancer showed absolutely no correlation with any one of the epidemiologic traits which, as presented in Table II, were highly correlated with the lung cancer incidence rate. This indicates that the variable completeness of registration had little effect on the present results.

Secondly we made no distinction between all 583 districts according to their differential population size. The reliability of morbidity or mortality statistics should depend on the population size. It must be greatly reduced when the population size is very small. So the influence of this factor was examined by repeating

the calculation of simple correlation coefficients after excluding districts with a population size of less than 10,000. The result was quite similar to the previously obtained one. Thus we may conclude that inclusion of those districts with a small population size scarcely influenced the quality of the current results.

Another question about the current method of data collection is the problem of so-called incubation time, that is, the time interval between exposure to some suspected causal factors and cancer incidence. Generally, an incubation time of 10 to 20 years is conjectured. In spite of this, a sufficient time lag was not placed between the years of measurement for epidemiologic traits and for cancer incidence. In Japan, the population flow between different districts, mainly from rural to urban areas, is significantly large every year. If a long time interval had been placed between the measurements of epidemiologic traits and the cancer rate, then the observed populations would have been composed differently. If a short time interval is used, as was done in the present case, a part of the population who have recently immigrated ought to be unexposed to the relevant factor. Also the sensitivity of the correlation analysis will be reduced and the possible effect of exposure to some urban factors might be concealed. We intend to conduct a similar correlation analysis, when morbidity and mortality statistics become available in sufficiently later years.

In Japan quite a few epidemiologic studies have reported on the association between environmental pollution and lung cancer. Shimizu¹⁹⁾ compared lung cancer mortality rates among the four districts of Osaka prefecture with different levels of air pollution. The mortality rate was correlated with the air pollution level but not with the regional smoking rate. This correlation with air pollution seemed to be due to mobile rather than stationary sources. Shimizu *et al.*²⁰⁾ studied the correlation between the level of air pollution (SO_2 and dust fall) and the standardized mortality ratio of lung cancer in 14 wards of Nagoya city. They were mutually highly correlated and a multiple regression analysis presented a significant partial correlation coefficient even after standardizing for the regional smoking rate, municipal tax revenue per capita and welfare-aid rate.

Minowa *et al.*²¹⁾ calculated the standardized mortality ratio (SMR) of lung cancer in all of more than 3,000 autonomous units of the whole country. Two groups of districts with either the highest or lowest SMR were chosen from them and various regional traits were compared between both groups. The districts of highest SMR included significantly more air-polluted areas than those of lowest SMR. Other indices of industrialization also showed a higher value in the former group. Makino and Mizoguchi²²⁾ performed a Spearman rank correlation test by using age-standardized mortality rates of lung cancer by ward in metropolitan Tokyo and obtained a significant correlation with various indices of air pollution as well as with automobile traffic density. A principal component analysis also demonstrated that lung cancer mortality was closely associated with the second component, which was representative of the industrial activity of the district.

The present study yielded more or less the same results as those mentioned above. Moreover, it disclosed contrasting features in the correlation of the lung cancer incidence rate with various urban-associated factors to those of colon cancer, despite the fact that both of them are correlated with the population density or with the level of urbanization in general. Lung cancer morbidity or mortality is, especially in males, positively correlated with the density of establishments per unit area for almost all kinds of manufacturing industries, with the automobile traffic density, and with air pollution indices such as SO_2 , NO_2 and floating dust. This is consistently observed even when the calculation is restricted to the urban area alone. Multiple regression analysis also resulted in approximate agreement with the result of simple correlation. On the other hand, colon cancer morbidity and mortality rates are correlated positively with the proportion of employees in tertiary industries such as sales, services and government among all workers and negatively with that of manufacturing industry, but not with any of the factors referred to above for lung cancer. In general, workers in the tertiary industries belong to a relatively higher socio-economic class than those in manufacturing industry. From these findings it is conjectured that, although both cancers have increased in

urban areas, lung cancer is more likely to be prevalent in areas with a more hazardous environment, i.e. with air pollution from industrial plants and/or automobiles. A negative partial correlation found for the automobile traffic density in the multiple regression analysis suggests that stationary rather than mobile sources of air pollution might be more important.

From the present result, lung cancer morbidity and mortality appear to be scarcely influenced by the number of manufacturing workers living in respective districts. Thus the above-mentioned relation between this disease and industrialization should be related not to any occupational feature, but to some aspect(s) of the residential environment. The result of multivariate analyses demonstrated that there was, at least among the presently studied traits, no single factor with a particularly strong influence on the lung cancer morbidity. Many hazardous factors in the urban environment, even though the individual effects are not so marked, may accumulate to give rise to a significant influence. Furthermore, the fact that, in Fig. 2, the male lung cancer rate is not necessarily plotted very close to the group of various industrialization-related factors is suggestive of possible involvement of other factors in the geographic variation of its incidence. Life style of the urban residents, including dietary factor and tobacco smoking, can not be ruled out as possible contributors.

It is already quite evident that tobacco smoking is the most important etiological factor for lung cancer. Nevertheless, this study does not include it as a study subject. Tobacco expenditure tax revenue showed no correlation with lung cancer. The per capita amount of tobacco expenditure tax revenue of prefectures is larger in urban than in rural areas in Japan.²³⁾ However, this does not reflect the actual smoking rate because people do not necessarily buy cigarettes at their place of residence. As mentioned previously, we have insufficient information about the smoking rate by districts. According to a national survey on family income and expenditure,²⁴⁾ per capita amount of tobacco expenditure can roughly be estimated for 49 cities in Japan. It shows a small positive (but not statistically significant) correlation with the population

density of those cities after logarithmic transformation ($r=0.251$, $d.f.=47$, $0.05 < P < 0.10$).

As far as we know, the data presented by Shimizu *et al.*²⁰⁾ are the only published ones obtained by a population-based survey on smoking rate in Japan. They presented smoking rate in different wards of Nagoya city in 1973–1974. According to their data, wards of Nagoya city can be divided into two groups by the male smoking rate, namely 8 wards with less than and 8 wards with more than or equal to the city average. By comparing the two areas with respect to various epidemiologic traits included in this study, we found that the area of higher smoking rate was also characterized by a higher proportion of workers in secondary industries, a lower one of workers in tertiary industries, a denser localization of establishments of various manufacturing industries, higher automobile traffic density and more population living with welfare aid. These are all the traits correlated with lung cancer, as stated above. Hence there still remains a possibility that the presently found correlation between the industrialization-related factors and lung cancer might be a superficial one and the true cause could be the effect of tobacco smoking. However, since the data were restricted to Nagoya city, a conclusive interpretation is not possible.

ACKNOWLEDGMENTS

This work was supported in part by a Grant-in-Aid for Cancer Research from Ministry of Health and Welfare of Japan. We are indebted to all the staff of population-based cancer registries who kindly provided their statistical data for this study, and especially to Dr. A. Takano of Miyagi Tumor Registry, Dr. Y. Satoh of Yamagata Cancer Registry, Dr. T. Ishida of Hyogo Medical Center for Adults, Dr. K. Takagi of Okayama Prefecture Ministry of Environment and Health and Dr. T. Shigematsu of Fukuoka Cancer Registry. We are also grateful to Dr. S. Tominaga and Dr. N. Tokudome, and to Mrs. R. Inoue for helpful discussions.

(Received April 11, 1988/Accepted July 23, 1988)

REFERENCES

- 1) Tominaga, S. Epidemiology of respiratory cancer in Japan. In "Current Status of Cancer Research in Asia, The Middle East and Other Countries," ed. T. Wada, K. Aoki and A. Yachi, pp. 173-179 (1987). The University of Nagoya Press, Nagoya.
- 2) 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. *Am. J. Epidemiol.*, **115**, 40-58 (1982).
- 3) Hammond, E. and Horn, D. Smoking and death rates. *J. Am. Med. Assoc.*, **166**, 1294-1308 (1958).
- 4) Haenszel, W., Loveland, D. and Sirken, M. Lung cancer mortality as related to residence and smoking histories. I. White males. *J. Natl. Cancer Inst.*, **28**, 947-1001 (1962).
- 5) Buell, P. Relative impact of smoking and air pollution on lung cancer. *Arch. Environ. Health*, **15**, 291-297 (1967).
- 6) Carnow, B. W. and Meier, P. Air pollution and pulmonary cancer. *Arch. Environ. Health*, **27**, 207-218 (1973).
- 7) Blot, W. J., Brinton, L. A., Fraumeni, J. F. and Stone, B. J. Cancer mortality in U.S. counties with petroleum industries. *Science*, **198**, 51-53 (1977).
- 8) Weiss, W. Lung cancer mortality and urban air pollution. *Am. J. Public Health*, **68**, 773-775 (1978).
- 9) Gottlieb, M. S., Pickle, L. W., Blot, W. J. and Fraumeni, J. F. Lung cancer in Louisiana: death certificate analysis. *J. Natl. Cancer Inst.*, **63**, 1131-1137 (1979).
- 10) Vena, J. E. Air pollution as a risk factor in lung cancer. *Am. J. Epidemiol.*, **116**, 42-56 (1982).
- 11) Dayal, H., Chiu, C. Y., Sharrar, R., Mangan, J., Rosenwaike, I., Shapiro, S., Henley, A. J., Goldberg-Alberts, R. and Kinman, J. Ecologic correlates of cancer mortality patterns in an industrialized urban population. *J. Natl. Cancer Inst.*, **73**, 565-574 (1984).
- 12) Doll, R. and Peto, R. The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. *J. Natl. Cancer Inst.*, **66**, 1191-1308 (1981).
- 13) Greenberg, M. R. "Urbanization and Cancer Mortality," pp. 5-13 (1983). Oxford University Press, New York.
- 14) Abe, S., Fujitaka, K., Abe, M. and Fujimoto, K. Extensive field survey of natural radiation in Japan. *J. Nucl. Sci. Tech.*, **18**, 21-45 (1981).
- 15) Williams, R. R. and Horm, M. S. Association of cancer sites with tobacco and alcohol consumption and socioeconomic status of patients: interview study from the third national cancer survey. *J. Natl. Cancer Inst.*, **58**, 525-547 (1977).
- 16) Okuno, T., Kume, H., Haga, T. and Yoshizawa, T. "Multivariate Analysis" (1975). Nikkagiren Press, Tokyo (in Japanese).
- 17) Yanai, H., Inaba, Y., Takagi, H. and Yamamoto, S. Multivariate analysis of cancer mortalities for selected sites in 24 countries. *Environ. Health Perspect.*, **32**, 83-101 (1979).
- 18) Snedecor, G. W. and Cochran, W. G. "Statistical Methods" (1967). The Iowa State University Press, Ames.
- 19) Shimizu, T. Lung cancer mortality and air pollution in Osaka. *J. Air Pollut. Res.*, **10**, 779-783 (1976) (in Japanese).
- 20) Shimizu, H., Tominaga, S., Nakagawa, N. and Kuroishi, T. An epidemiological study on lung cancer — the relationship between the lung cancer mortality and air pollution by ward in a large city. *J. Air Pollut. Res.*, **14**, 324-333 (1979) (in Japanese).
- 21) Minowa, M., Shigematsu, I., Nagai, M. and Fukutomi, K. Geographical distribution of lung cancer mortality and environmental factors in Japan. *Soc. Sci. Med.*, **15D**, 225-231 (1981).
- 22) Makino, K. and Mizoguchi, I. Relationship between respiratory disease mortality and air pollution or social indices in the Ku-area of Tokyo (I). *Jpn. J. Public Health*, **31**, 297-303 (1984) (in Japanese).
- 23) Japan Tobacco and Salt Corporation. Report of Survey on Smoking (1975) (in Japanese).
- 24) Prime Minister's Office Statistics Bureau. Annual Report on the Family Income and Expenditure Survey (1981) (in Japanese).