

Influence of Socioeconomic and other Factors on the Geographic Variation of Mortality in Finland, Sweden and Norway

By TAPANI VALKONEN and VEIJO NOTKOLA

University of Helsinki

Department of Sociology

The purpose of this report is to present a summary of the findings of two empirical studies which are part of a larger project on the geographic variation in mortality. The project was started in 1975 with a study of the influence of socioeconomic factors on age-specific death rates for countries in Finland, Sweden and Norway (Valkonen 1975). This was followed by a study of death rates in a sample of Finnish rural municipalities (Notkola 1976). The project is continuing with a more intensive analysis of the data on Finnish municipalities and with a study of the cause-specific death rates for counties in Finland, Sweden and Norway.

The main impetus to the project came from two related facts uncovered during several earlier studies. The first is the high mortality in Finland as compared to that of the other Scandinavian countries (see e.g. Bolander 1971). The second is the existence of clear and systematic geographic differences within countries. In Finland special attention has been paid to the high death rates in the Eastern parts of the country (Kannisto 1947, Leppo, Lindgren and Ritamies 1971).

Similarly, in Norway there exists an area of exceptionally high death rates in the North (e.g. Regional mortality 1969—1972, 1974). The variation of death rates is smaller in Sweden, but still systematic (Mortality and causes of death by regions 1964—1967, 1971).

The analysis of geographic differences in mortality is important for two reasons. Variation in mortality indicates differences in the general level of health among the population and can be considered one of the most important indicators of the general quality of life. Regional mortality data can thus be used in analysing the more general problems of regional quality of life and regional inequality. On the other hand, research on the regional variation in death rates is connected with epidemiology. The analysis of cause-specific death rates may yield important clues for the study of etiology of diseases. This may be of special importance regarding such diseases as cardiovascular diseases and cancer, the etiologies of which are still unclear.

The purpose of the present research is not to describe the geographic variation of death rates, as this has already been done in several studies. Rather, the main task here will be to determine the relative importance of various factors which lead to the existence of differential mortality. Research on these factors

has not been sufficiently extensive in the Scandinavian countries, and there exist only tentative answers to the main problems (see e.g. Härö 1966 and 1975).

In certain other countries, e.g. Great Britain and the United States, more extensive research has been carried out. The conclusions are, however, very vague. Thus, in summarizing the findings of American studies, Sauer (1974, 122—123) states:

»Both low levels of smoking and large proportions of population engaged in farming have been demonstrated to be associated with low age-specific death rates. These associations may very well explain some of the geographic variations in death rates, in view of the low rates reported for farmers of middle age and the high rates for smokers. Many other factors have been identified as associated with statistically significant (0.01 level) geographic variations in age-specific death rates; these include characteristics of the drinking water, specifically low levels of total solids, hardness, calcium, magnesium, and sodium, as well as other factors such as low elevation above sea level, relative humidity, small average daily changes in January temperature, proportion of women working, and health manpower available. The evidence pertaining to these and many other variables is, however, so inconclusive at present that it constitutes no more than leads for further study.»

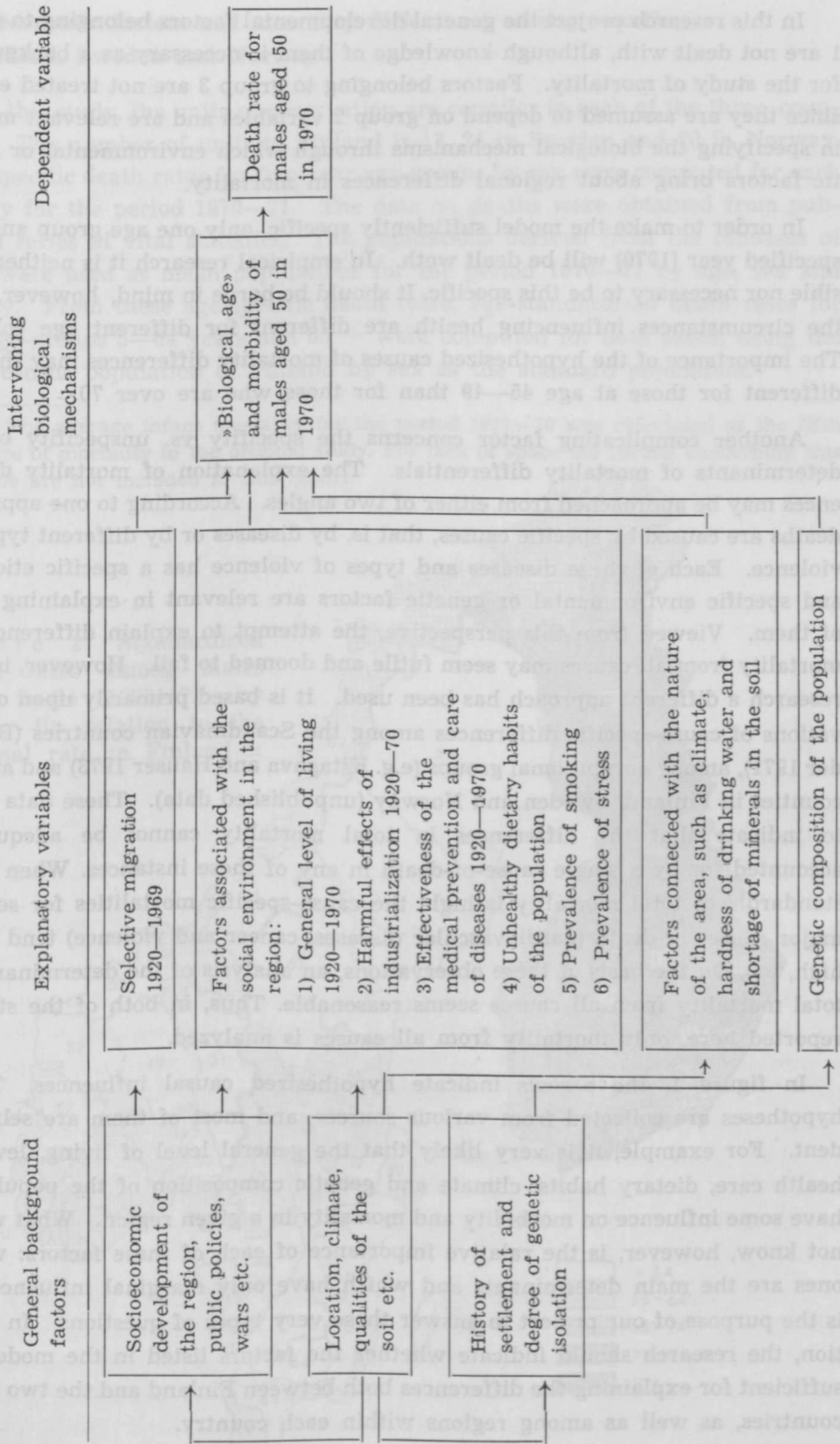
The main reason for the inconclusive results of earlier research may be that the research has not been based on any over-all frame of reference about the factors influencing the level of mortality. Correlations without a general framework are not very useful in the search for answers about the processes leading to mortality differences. In our research, therefore, we have constructed a kind of general model for the explanation of geographic differences in age-specific death rates. This model will be described briefly in the next section of this article. The main findings of the two studies will be presented thereafter.

Hypotheses about the causes of geographic mortality differences

Various hypotheses about the possible causes of mortality differences are presented in concise form in Figure 1. According to this »model», three levels of causes are distinguished:

1. Geographic and historical factors which have influenced the general development of a region and brought about conditions and ways of life directly affecting the level of health of the population in a region.
2. Those living conditions and other factors just mentioned which have a direct effect on the biological health of the population. These may be connected with the social environment, with the natural environment or with the genetic characteristics of the population.
3. Variables indicating the level of health or risk of becoming ill among the population, such as the prevalence of hypertension or elevated cholesterol level. Hypothetically this can also be called the »biological age» of the population.

Figure 1. Model for explaining regional variation in mortality: death rate for males aged 50 years in 1970 as an example



In this research project the general developmental factors belonging to group 1 are not dealt with, although knowledge of them is necessary as a background for the study of mortality. Factors belonging to group 3 are not treated either, since they are assumed to depend on group 2 variables and are relevant mainly in specifying the biological mechanisms through which environmental or genetic factors bring about regional differences in mortality.

In order to make the model sufficiently specific, only one age group and one specified year (1970) will be dealt with. In empirical research it is neither possible nor necessary to be this specific. It should be borne in mind, however, that the circumstances influencing health are different for different age cohorts. The importance of the hypothesized causes of mortality differences may thus be different for those at age 45—49 than for those who are over 70.

Another complicating factor concerns the specificity vs. unspecificity of the determinants of mortality differentials. The explanation of mortality differences may be approached from either of two angles. According to one approach deaths are caused by specific causes, that is, by diseases or by different types of violence. Each of these diseases and types of violence has a specific etiology, and specific environmental or genetic factors are relevant in explaining each of them. Viewed from this perspective, the attempt to explain differences in mortality from all causes may seem futile and doomed to fail. However, in this research a different approach has been used. It is based primarily upon observations of cause-specific differences among the Scandinavian countries (Bolan-der 1971), among occupational groups (e.g. Kitagava and Hauser 1973) and among counties in Finland, Sweden and Norway (unpublished data). These data seem to indicate that the differences in total mortality cannot be adequately accounted for by a single cause-of-death in any of these instances. When (age-standardized) total mortality is high, the cause-specific mortalities for several major causes of death (gastrovascular diseases, cancer and violence) tend to be high, too. On the basis of these observations, an analysis of the determinants of total mortality from all causes seems reasonable. Thus, in both of the studies reported here, only mortality from all causes is analyzed.

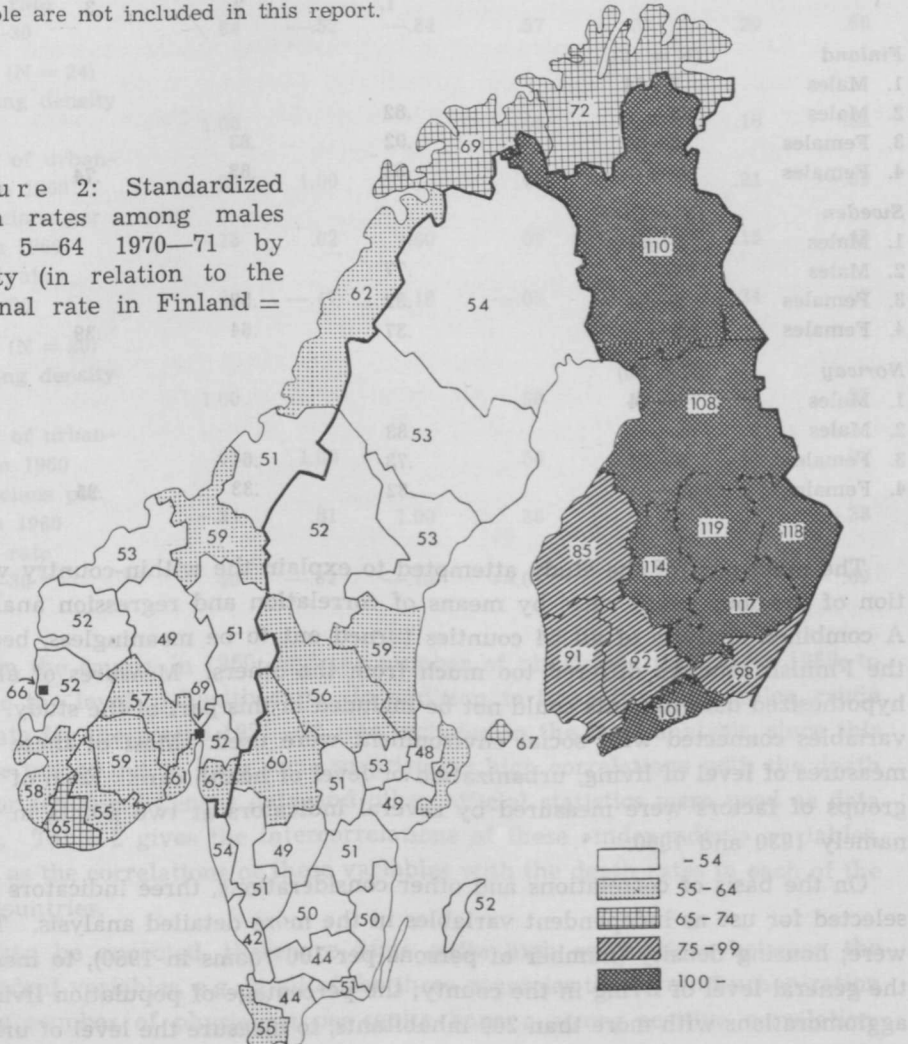
In figure 1, the arrows indicate hypothesized causal influences. These hypotheses are collected from various sources, and most of them are self-evident. For example, it is very likely that the general level of living, level of health care, dietary habits, climate and genetic composition of the population have some influence on morbidity and mortality in a given region. What we do not know, however, is the relative importance of each of these factors: which ones are the main determinants and which have only marginal influence. It is the purpose of our project to answer these very types of questions. In addition, the research should indicate whether the factors listed in the model are sufficient for explaining the differences both between Finland and the two other countries, as well as among regions within each country.

Socioeconomic factors and mortality differences among counties in Finland, Sweden and Norway

In this study, the units of observation are counties in each of the three countries. The number of units in Finland is 12, 24 in Sweden and 20 in Norway. Age-specific death rates for five year age-groups by sex were computed for each county for the period 1970—71. The data on deaths were obtained from published series of vital statistics. The populations derived from the censuses of 1970 were used as mean populations for the period 1970—71 by age, sex and county. From these age-specific death rates, age-standardized death rates for the age-groups 5—64 years and 65 + were computed for both sexes, using the 1970 census population of Finland by sex as the standard population.¹

¹ The average infant mortality for the period 1961—70 was calculated as the fifth measure of mortality in the original study. For lack of space the results concerning this variable are not included in this report.

Figure 2: Standardized death rates among males ages 5—64 1970—71 by county (in relation to the national rate in Finland = 100).



In Figure 2 the data for males at age 5—64 are presented in index numbers (Finland = 100). The high level of death rates in all of the Finnish counties except Åland is quite obvious. In addition clear regions of high and low mortality can be distinguished within each country. Maps for the other age groups are not published here, but in most respects they resemble Figure 2: death rates are higher in Finland than in the other countries and the regions of high and low mortality coincide to a large extent.

The intercorrelations of the four mortality variables were calculated for each country (Table 1). The correlations confirm the impression received from visual inspection of the maps: the age-specific death rates correlate strongly.

Table 1. The intercorrelations of age-sex-specific death rates separately for counties in Finland, Sweden and Norway

		1.	2.	3.
<i>Finland</i> (N = 12)				
1. Males	5—64			
2. Males	65—	.82		
3. Females	5—64	.92	.83	
4. Females	65—	.82	.83	.74
<i>Sweden</i> (N = 24)				
1. Males	5—64			
2. Males	65—	.77		
3. Females	5—64	.32	.52	
4. Females	65—	.37	.64	.39
<i>Norway</i> (N = 20)				
1. Males	5—64			
2. Males	65—	.83		
3. Females	5—64	.73	.64	
4. Females	65—	.32	.33	.25

The main part of the study attempted to explain the within-country variation of the four death rates by means of correlation and regression analyses. A combined analysis of all 54 counties turned out to be meaningless, because the Finnish counties differed too much from the others. Measures of all the hypothesized determinants could not be included in this part of the study. Only variables connected with social environment were used. Most of these were measures of level of living, urbanization or level of health care. Each of these groups of factors were measured by several indicators at two points in time, namely 1930 and 1960.

On the basis of correlations and other considerations, three indicators were selected for use as independent variables in the more detailed analysis. These were; housing density (number of persons per 100 rooms in 1960), to measure the general level of living in the county; the percentage of population living in agglomerations with more than 200 inhabitants, to measure the level of urbani-

Table 2. The intercorrelations of the main independent variables, and their correlations with mortality variables in Finland, Sweden and Norway

	Independent variables			Death rates			
	1	2	3	Male 5—64	Male 65—	Female 5—64	Female 65—
<i>Finland (N = 12)</i>							
1. Housing density 1960	1.00			.85	.65	.70	.79
2. Level of urban- ization 1960	-.35	1.00		-.16	.00	-.04	-.35
3. Physicians per capita 1960	-.77	.55	1.00	-.56	-.51	-.39	-.78
4. Birth rate 1921—30	.84	-.52	-.84	.57	.37	.29	.60
<i>Sweden (N = 24)</i>							
1. Housing density 1960	1.00			.44	.46	.16	.32
2. Level of urban- ization 1960	.09	1.00		.41	.25	.21	-.23
3. Physicians per capita 1960	-.18	.62	1.00	.07	-.08	-.15	-.47
4. Birth rate 1921—30	.60	-.47	-.18	-.03	-.01	-.34	.07
<i>Norway (N = 20)</i>							
1. Housing density 1960	1.00			.58	.42	.27	.37
2. Level of urban- ization 1960	.07	1.00		.59	.80	.49	-.27
3. Physicians per capita 1960	-.24	.81	1.00	.38	.51	.38	-.38
4. Birth rate 1921—30	.46	-.62	-.70	-.07	-.18	-.13	.69

zation in the county in 1960; and the number of physicians per capita 1960, to measure the level of health care. In addition to these three variables, crude birth rate for the period 1921—30 was included in the final analysis, since this variable turned out to have some surprisingly high correlations with the death rates for females. Census data and other official statistics were used as data sources. Table 2 gives the intercorrelations of these »independent» variables, as well as the correlations of these variables with the death rates in each of the three countries.

As can be expected, there are some quite high correlations between the independent variables, e.g. in all of the three countries the level of urbanization and the number of physicians per capita have a strong positive correlation.

Housing density and urbanization have a weak negative correlation in Finland but negligible positive correlations both in Sweden and Norway. Birth rate 1921—30 correlates strongly both with housing density (positive correlation) and urbanization (negative correlation) in each country.

The primary focus of Table 2, however, is upon the correlations of the independent variables with the death rates for the two age groups by sex. The following observations can be made from these correlations:

1. Housing density correlates with the mortality variables in all of the three countries, indicating a systematic association of low level of living and high mortality. The strength of this association varies considerably, however, from country to country. In Finland, the correlations are quite strong, much weaker in Norway and weakest in Sweden. The correlations tend to be lower for females than for males.

2. Level of urbanization of the county has a strong positive correlation with death rates among Norwegian males. In Sweden, a similar but weaker association of high level of urbanization and high mortality can be observed, but not in Finland. Among females the situation is different: in all countries death rates for females 65 years and older have negative correlations with the level

Table 3. Summary of the results of regression analyses by country. Dependent variables: Standardized death rates for two broad age-groups by sex. Independent variables: percentage of urban population (U) and housing density (H, number of persons per 100 rooms)

Country	Constant	Regression coefficients		Stand. regr. c.		R ²	R ² only H*
		U	H	U	H		
<i>Males 5—64</i>							
Finland	96	.097	3.30	.16	.91	.75	.72
Sweden	99	.081	1.79	.38	.40	.33	.19
Norway	—7	.103	3.72	.56	.54	.65	.34
<i>Males, 65 and over</i>							
Finland	5109	1.027	17.0	.26	.74	.48	.42
Sweden	2788	.616	26.7	.21	.44	.26	.22
Norway	2133	1.874	33.1	.77	.37	.77	.18
<i>Females, 5—64</i>							
Finland	170	.025	.49	.24	.78	.54	.49
Sweden	157	.021	.32	.19	.15	.06	.03
Norway	119	.031	.57	.48	.24	.30	.07
<i>Females, 65 and over</i>							
Finland	3503	—318	19.1	—0.08	.77	.64	.63
Sweden	3419	—516	13.9	—0.26	.34	.17	.10
Norway	3265	—338	16.4	—0.29	.39	.22	.13

* Proportion of explained variance (R²), when housing density alone is used as the independent variable.

of urbanization. This indicates that mortality among old women is lower than average in urbanized counties. Among younger females the findings are similar to the findings for men: a positive correlation in Sweden and Norway and no correlation in Finland.

Because of the intercorrelations between the independent variables, the inspection of the zero-order correlations does not show the net influence of each of the independent variables. Therefore, regression analysis was used for more detailed study. A summary of the findings is presented in Tables 3 and 4.

Table 3 gives the results of the analyses in which housing density and level of urbanization were used as the only independent variables. The findings are, of course, largely the same as those that were obtained directly from the correlations. In only one respect is there a difference: the regression coefficients for urbanization now have the same signs in Finland as in Sweden and Norway. This means that when the level of living is controlled for, urbanized counties in Finland are also characterized by somewhat higher than average mortality among men and middle-aged women but lower than average mortality among old women.

The highest multiple correlation coefficients in the regression analyses are obtained for Norwegian males of 65 and over, and for Finnish males at ages 5—64. In these groups, more than 75 per cent of the intercounty variation of the death rates can be explained by the two independent variables. The lowest percentages of explained variance are observed among both groups of Swedish females. The multiple correlations are generally high for all four groups in Finland and for Norwegian males, and low for all groups in Sweden and for Norwegian females.

Table 4 gives a summary of two different analyses. In the first, part (a), physicians per capita is used as the third independent variable in addition to housing density and urbanization. The complete regression equations are not presented. Rather, the multiple correlation coefficient squared (R^2) for the three variable analysis, the rise of R^2 as compared to the two variable model presented in Table 3, and the standardized regression coefficient for the physician variable are shown. The general conclusion from this part of Table 4 is that the number of physicians per capita has no, or at most, very slight direct influence on the level of mortality. This conclusion is, however, complicated by the fact that physician density correlates quite strongly with the level of urbanization in all countries and also with level of living in Finland.

In part (b) of Table 4, a similar analysis is carried out using crude birth rate in 1921—30 as the third independent variable, in addition to housing density and level of urbanization. In these analyses there are slight increases in the multiple correlations for males, but these can be considered insignificant. In some, but not all female groups, the increments in the percentage of the explained variance are so great that some non-random factors must obviously be operating. These groups are younger females in Finland, older females in Norway, and both younger and older females in Sweden. In all of these groups the

Table 4. Summary of the regression analyses by country, when
 a) number of physicians per capita 1960 (P) and
 b) crude birth rate 1921—30 (B)
 is used as independent variable in addition to percentage of
 urban population and housing density

	a) Physicians as the additional variable			b) Birth rate as the additional variable		
	R ² for the whole model	Increment to R ² from P	Stand. r.c. of P	R ² for the whole model	Increment to R ² from B	Stand. r.c. of B
<i>Males, 5—64</i>						
Finland	.75	.01	.17	.79	.05	— .45
Sweden	.34	.01	— .15	.36	.03	— .28
Norway	.66	.02	.24	.65	.00	.07
<i>Males, 65 and over</i>						
Finland	.50	.02	— .28	.53	.05	— .46
Sweden	.29	.03	— .23	.35	.09	— .51
Norway	.77	.00	— .08	.82	.05	.36
<i>Females, 5—64</i>						
Finland	.56	.02	.27	.78	.24	— 1.00
Sweden	.16	.10	— .42	.38	.31	— .93
Norway	.31	.01	.19	.31	.01	.16
<i>Females, 65 and over</i>						
Finland	.70	.07	— .46	.66	.03	— .34
Sweden	.28	.10	— .43	.37	.20	— .74
Norway	.23	.01	— .21	.53	.31	.92

increment in explained variance is at least 20 percentage points. The standardized regression coefficients are quite high as well. It is astonishing, however, that the signs of the regression coefficients are not the same in all cases, but different for Norway than for the other two countries. This means that, when the influence of level of living and urbanization has been controlled for, the mortality among females (especially if they are younger than 65 years old) in Sweden and Finland is *lower* than the average for counties where the birth rate was high in the 1920's. On the other hand, in Norway, the mortality of older women is *higher* than the average for counties with high birth rate in the 1920's.

The analyses as a whole can be interpreted as indicating that two of the variables (level of living and harmful consequences of industrialization) included in the original explanatory model explain a considerable part of the geographic variation of the death rates. The relative significance of these variables varies from country to country and from age-sex-group to age-sex-group. Among older women, the influence of urbanization is actually the opposite than that assumed. It has also been found that the (present) level of health care is not a highly significant factor. In addition, a new variable that was not part of the original model was found to be relevant in explaining mortality differences for females. The meaning of this new variable (earlier level of birth

rates) is not clear, however, and more research is needed to clarify the nature of its association with mortality.

The entire intercounty variation of death rates was not accounted for in the regression analyses. To get some idea of the geographic distribution of these unexplained differences in death rates, the following computations were carried out separately for each country:

1. The residuals for each of the four mortality variables were calculated from the regression analyses presented in Table 3 in which density and urbanization were used as the independent variables. The residuals for each county give the difference between the actual death rates and the death rates expected on the basis of housing density and level of urbanization in the county.

2. Absolute residuals were transformed to relative residuals by expressing them as percentage of the expected death rate.

3. The mean of the four relative residuals was computed for each county, to indicate how much higher or lower than expected the actual mortality is on the average.

The average relative residuals are presented in Figure 3. In the shaded counties the death rates are higher than expected, and in the unshaded counties lower than expected. Since a considerable part of the variation in death rates was explained by the two independent variables, the variation of these residuals around the country mean is much smaller than the variation of the original death rates was. The variation of residuals is, however, surprisingly regular and systematic. In each country clear zones of high and low mortality can be distinguished. In Finland the high mortality zone consists of five counties in Eastern Finland. In Sweden the high mortality counties are concentrated in the central area and on the East coast. In Norway there are two zones of high mortality, one in the North and one in the South.

The existence of the high and low mortality zones not explainable by regional differences in the level of living and urbanization, can be interpreted to indicate that some other social factors (or factors associated with the natural environment or the genetic composition of the population) influence mortality. On the basis of the data in this part of the study, however, it is only possible to speculate about these factors. The purpose of the second part of the study was to analyze the problem with more detailed empirical data.

Mortality in 74 rural municipalities in Finland

There are several weaknesses in the data used in the first part of the study. Counties are large and heterogeneous units, each of them containing both rural and urban populations. The number of units is very low for a multivariate analysis, especially in Finland. In addition, only a part of the factors hypothesized to influence mortality were measured.

In the second part of the study some of these weaknesses were attempted

to be overcome by data on Finnish rural municipalities. In this way it was possible to use units with greater internal homogeneity than counties, and to increase considerably the number of cases. It was also possible to obtain meaningful data on several explanatory variables included in the model presented earlier.

The main problem in using municipality data is to obtain death rates which are not too greatly influenced by random variation. To achieve this goal, a ten year period (1961—70) was chosen for collecting the data. In addition, all small

Figure 3: The average relative (%) deviations of observed death rates from expected death rates on the basis of the regression analyses, in which percentage of urban population and housing density are independent variables (computed separately for the counties in each country)

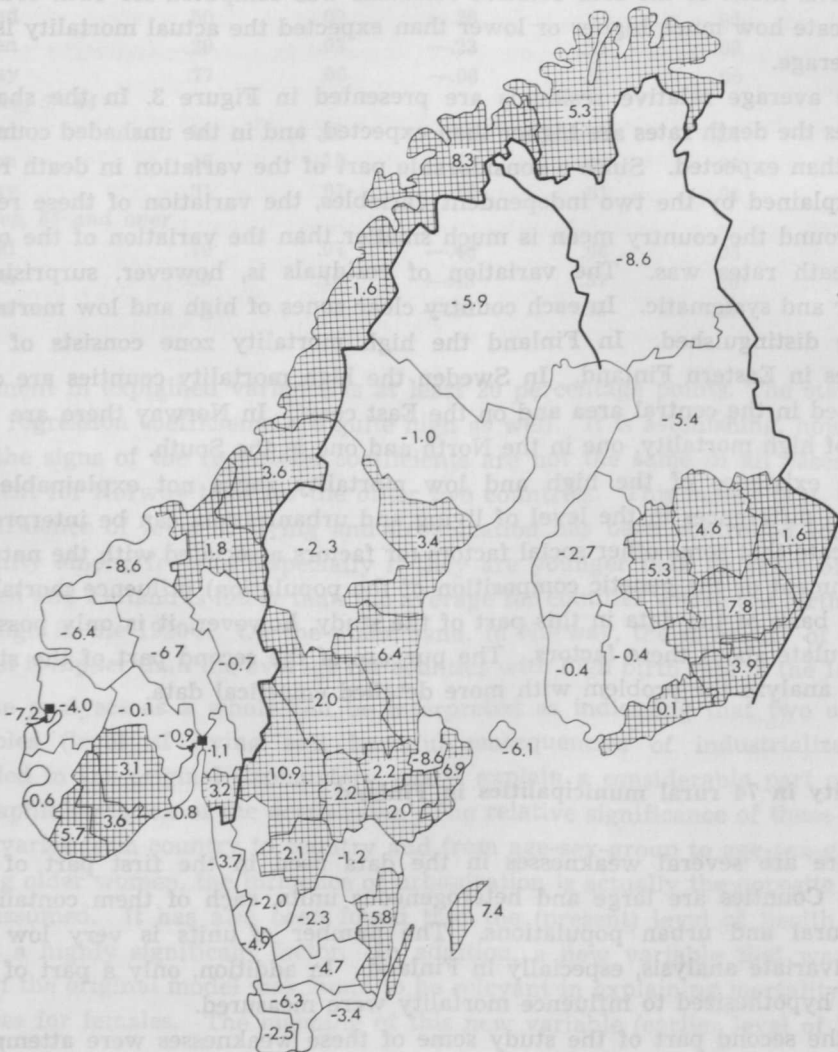


Figure 4: Standardized mortality ratios (males and females) in 74 rural municipalities in Finland 1961—1970. Indirect standardization, sex- and age-specific death rates for all rural municipalities as standard mortality



municipalities (population smaller than 7 600) were excluded. A sample of 74 rural municipalities was drawn by selecting the largest rural municipalities from each of the eleven »continental» counties. The rural municipalities of the Åland province were treated as one unit in the data.

Standardized mortality ratios¹ were calculated for the municipalities in the sample by indirect standardization using the age-sex-specific death rates for all rural municipalities as standard death rates. The averages of the census populations of 1960 and 1970 by age and sex were used as the mean populations for the period 1961—70.

Mortality ratios were calculated both for all deaths and for males and females separately. Figure 4 gives the mortality ratios for all deaths in the sample of municipalities. The highest mortality ratios can be found in municipalities in Eastern and Central Finland, that is, in the regions where death rates are known to be high.

¹ The standardized mortality ratio indicates the ratio (multiplied by 100) of the actual number of deaths in the municipality to the number expected on the basis of the age-sex composition of the population.

Part of the variation of the mortality ratios is random due to the limited number of deaths in each municipality. On the basis of the number of deaths in the municipalities, it can be roughly estimated that the standard error of the mortality ratios varies from two to four percent.

The variances of the mortality ratios calculated separately for males and females do not differ very greatly from one another. The standard deviation for both males and females is 0.11. The correlation coefficient between the male and female ratios is quite high (0.67). The real correlation between male and female mortality is still higher, since random errors in both variables make the observed correlation lower than the real correlation.

In addition to the mortality variables just described, a fourth dependent variable was constructed using data on the number of disability pension recipients in the sample of municipalities. These data are published by the National Pensions Institute. Figure 5 gives age-standardized ratios for the percentage of disability pension recipients among the total working-age population (15—64 years old) at the end of the year 1972. These ratios are analogous to the standardized mortality ratios, and indicate the ratios of the actual numbers of pension recipients to the expected numbers. The ratios were also computed separately for three types of disability-causing diseases (diseases of the circulatory system, mental disorders, and diseases of the musculoskeletal system and connective tissue). Since these ratios correlated strongly with one another, only the total disability pension ratio was used in the analysis.

Figure 5 indicates that the differences among municipalities in disability pension ratios are much greater (standard deviation 0.36) than the differences in mortality ratios. The disability pension ratios are relatively higher than the mortality ratios in many municipalities in Northern Finland. Otherwise the pattern of the variation of disability pension ratios is similar to the variation of mortality ratios. Thus generally, the highest values are in Eastern and Central Finland and the lowest in Southwestern Finland. The correlation coefficient between mortality and disability ratios is 0.67.

The independent variables for this part of the study were chosen on the basis of the model presented earlier. About 50 indicators were originally constructed to measure the explanatory factors.

Some of the factors included in the model could not be measured at the level of the municipality. Most important of these were smoking habits and dietary habits. On the basis of indirect evidence it was, however, concluded that regional differences in these factors are not likely to explain mortality and disability differences to a significant degree.

Regional differences in smoking do not seem to be great enough to explain mortality differences (see Pedersen et al. 1969, Karvonen 1972, Rimpelä and Puska 1975). In addition, if smoking habits were an important determinant of mortality differences, the variation of mortality ratios would have to be much greater for males than for females, and a high correlation between these variables could not be expected.

According to the studies by Pekkarinen (1962) and by Koskinen (1975), the regional differences in the fat content of the diet are not great and systematic. The observed differences in the blood cholesterol level are also quite small (Aromaa et al. 1973). In the light of these findings, it seems less likely that dietary habits at the present time would be a major cause for mortality differences.

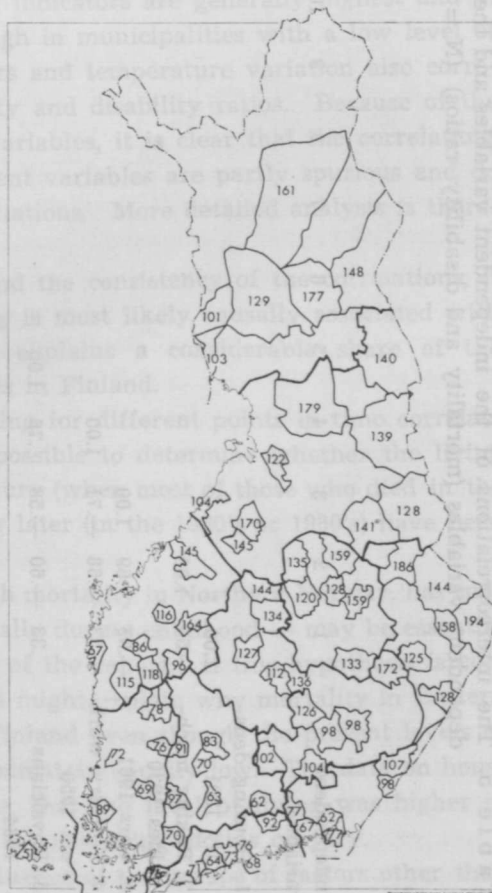
The associations of the independent variables with the dependent variables were first studied using correlations. On the basis of these analyses, ten variables representing the main groups of hypothesized determinants of mortality were selected for more detailed analysis. These variables were as follows:

I Level of living 1900—1960

1. Housing density 1901: persons per 100 rooms in 1901 (Gebhard 1916, table 2)
2. Raking machines 1930: horse drawn raking machines per 100 farms in 1930 (OSF III 26: 3, table 1)
3. Cars 1961: registered cars per 1000 inhabitants in 1961 (Palmgren 1964, table 3)

II Occupational structure

Figure 5. Standardized disability pension ratios (males and females) in 74 rural municipalities in 1971. Indirect standardization, sex- and age-specific shares of disability pension recipients per working age population in all rural municipalities as standard disability rates.



4. Forestry workers 1950: percentage of forestry workers among the total economically active population in 1950 (OSF IV C: 102, table 1)

III Level of health care

5. Physicians 1964: an index measuring the average distance of the population in the municipality to the nearest physician (Kalimo 1967, Appendix 5)

IV Natural environment

6. Calcium in the soil: average calcium content (mg/ltr) of the soil according to soil tests in the municipality (Kurki 1972, table 16). The

Table 5. The intercorrelations of the independent variables and their correlations with the dependent variables (mortality and disability ratios) (N = 74)

Independent variables	1	2	3	4	5	6	7	8	9	10	Mortality 1961-70		Disability		
											M	F		M&F	
1. Housing density 1901	1.00											.64	.56	.66	.61
2. Raking machines 1930	-.50	1.00										-.58	-.45	-.57	-.69
3. Cars 1961	-.55	.69	1.00									-.63	-.57	-.65	-.85
4. Forestry workers 1950	.53	-.68	-.71	1.00								.62	.45	.59	.75
5. Physicians 1964	.33	-.60	-.58	.76	1.00							.35	.20	.31	.65
6. Calcium in the soil	-.32	.59	.64	-.59	-.58	1.00						-.39	-.36	-.42	-.70
7. Magnesium in the soil	-.35	.60	.42	-.35	-.28	.46	1.00					-.34	-.31	-.36	-.39
8. Temperature variation	.49	-.55	-.76	.74	.65	-.66	-.25	1.00				.62	.59	.66	.65
9. Precipitation	.31	-.11	-.14	.23	.18	.14	-.03	.26	1.00			.25	.30	.29	.13
10. Percentage non-migrant	.39	-.67	-.62	.53	.58	-.55	-.40	.39	-.14	1.00		.34	.29	.35	.64
11. Level of living 1901-60	-.75	.86	.88	-.75	-.60	.61	.53	-.70	-.20	-.67	1.00	-.70	-.61	-.72	-.84

/r/ > .30 is statistically significant at the 0.01 level, when N = 74

M = Males

F = Females

M&F = Males & Females

data are based on about 1.5 million soil tests taken 1955—70 throughout the entire country.

7. Magnesium in the soil: same as the previous variable
8. Temperature variation: the average range of yearly variation in temperature 1931—60 (Kolkkki 1966, table 13). Values for municipalities were estimated on the basis of data from the nearest weather observation stations.
9. Precipitation: the average number of rainy days per year 1931—60 (Helimäki 1967)

V Genetic composition of the population

10. Percentage non-migrant: percentage of population who in 1960 lived in the same municipality where they were born (Statistical Surveys 44).

The intercorrelations of the ten independent variables are presented in Table 5. It can be seen at once that most of the correlation coefficients are high or moderately high, which makes interpretation of the study's findings difficult.

Table 5 also contains the correlations of the independent variables with the dependent variables. Almost all of the coefficients are statistically significant. The correlations of the level of living indicators are generally highest and in the expected direction: mortality is high in municipalities with a low level of living. Percentage of forestry workers and temperature variation also correlate quite strongly with both mortality and disability ratios. Because of the intercorrelations of the independent variables, it is clear that the correlations between the independent and dependent variables are partly spurious and do not necessarily imply any causal associations. More detailed analysis is therefore necessary.

On the basis of earlier research and the consistency of the correlations, it was concluded that the level of living is most likely causally associated with mortality and disability ratios, and explains a considerable share of the geographic variation of these variables in Finland.

Since the indicators of level of living for different points in time correlate strongly with one another, it is not possible to determine whether the living conditions at the beginning of the century (when most of those who died in the 1960's were children) or those existing later (in the 1930's or 1960's) have been more important.

Forsdahl (1973), in studying the high mortality in Northern Norway, has concluded that living conditions — especially during childhood — may be essential as the determinant of later mortality of the cohort. If this hypothesis is correct and relevant to Finland as well it might explain why mortality in Eastern Finland is higher than in Northern Finland even though the present levels of living in these two regions are approximately equally low. The data on housing density in 1901 seems to indicate, that the level of living was higher in Northern than in Eastern Finland at the beginning of this century.

To simplify the analysis of the influence of the groups of factors other than

level of living, a combined index of level of living was used to cover the whole period 1900—1960. The index was constructed as an average of the three indicators of level of living (housing density 1901, raking machines 1930 and cars 1961), after these variables had been standardized to give them equal weights.

To determine whether the other groups of factors have influences on mortality when the effect of level of living has been controlled for, partial correlations were calculated between the remainder of the independent variables and the dependent variables, holding the combined index of level of living constant.

Table 6. Partial correlations between the independent variables and the mortality and disability variables controlling for the level of living 1900—1960: whole country, Western region and Eastern region

	Whole country (N = 74)				Western region (N = 30) ¹				Eastern region (N = 23) ²			
	Mortality		Disability		Mortality		Disability		Mortality		Disability	
	M	F	M&F	M&F	M	F	M&F	M&F	M	F	M&F	M&F
Forestry workers												
1950	0.20	-0.01	0.12	0.35*	0.25	0.42*	0.38*	0.01	0.23	0.07	0.20	0.50*
Physicians 1964	-0.12	-0.25*	-0.21	0.34*	-0.07	0.08	0.01	0.32	-0.10	0.27	0.09	0.10
Calcium in the soil	0.03	0.01	0.04	-0.44*	0.02	-0.25	-0.14	-0.39*	0.12	0.03	0.11	-0.57*
Magnesium in the soil	0.05	0.02	0.04	0.10	0.16	-0.05	0.06	0.12	-0.41*	-0.29	-0.47*	0.27
Temperature												
variation	0.16	0.29*	0.31*	0.40*	0.39*	0.55*	0.54*	0.44*	0.22	0.41*	0.39	0.42*
Precipitation	0.25*	0.23*	0.22	-0.07	0.33	0.29	0.35	-0.07	-0.08	0.40	0.16	0.24
Percentage non-migrant	-0.25*	-0.20	-0.26*	0.18	-0.21	0.01	-0.11	0.25	-0.06	-0.46*	-0.30	-0.13

* Statistically significant at the 0.05 level.

¹ The counties: Ahvenanmaan, Uudenmaan, Turun and Porin, Hämeen and Vaasan lääni

² The counties: Mikkelin, Kuopion, Keski-Suomen and Pohjois-Karjalan lääni

To check the consistency of the findings, these partial correlations were computed separately for municipalities in the five low-mortality counties in Western Finland, and for municipalities in the four high-mortality counties in Eastern Finland (Table 6).

According to Table 6 the only variable which has consistent partial correlations with the dependent variables, holding level of living constant, is the variation of temperature. Since the regional variation in the summer temperatures is quite small, this variable in practice measures the coldness of winters, in which there are considerable regional differences. It is quite possible that the high death rates in Eastern and Northern Finland are partly caused by the very cold winter climate (cf. Dudley et al. 1969). Two reservations should, however, be taken into account. First, the climate in Finland changes systematically from south-west to north-east. The observed partial correlations may thus be caused by any other factor having the same regional pattern (e.g. the genetic composition of population). Second, there are some municipalities in Northern Finland with lower than average mortality ratios. The other climatic variable, precipitation, also shows some partial correlations which may be more than random. At this point, however, no conclusions can be drawn concerning this.

The percentage of forestry workers shows consistent positive partial correlations with the mortality ratios of males, and with the disability ratios. This can be interpreted as supporting the opinion commonly held about the harmful effect of forest work on health. This variable, however, explains only a very small portion of the total variance of mortality ratios.

Two of the variables, average distance to the nearest physician and percentage of the population born in the commune, have mostly low and inconsistent partial correlations with the dependent variables and it seems impossible to interpret them in a meaningful way. This is also partly the case with both variables measuring the mineral content of the soil. Two observations, however, seem worth mentioning. First, the shortage of calcium in the soil is associated consistently, and moderately strongly, with high disability pension ratio, but not with mortality ratios. Second, the shortage of magnesium is associated with high mortality ratios in Eastern Finland, but not in Western Finland.

The latter observation could be considered an artifact, due to the small number of cases or spurious correlations. A more detailed inspection of the association shows, however, that in those six municipalities (five of which are in Eastern Finland) in which the magnesium content of the soil is smallest, the mortality ratios are quite high. This may indicate that the association of the magnesium content of the soil and mortality is not linear, but that there is a threshold below which the shortage of magnesium begins to have harmful effects on health. In most areas the magnesium content is above this threshold, but there may be some areas, especially in Eastern Finland, in which the magnesium content (or some other quality of the soil correlated with magnesium

content) is of relevance. Research on this hypothesis would be closely connected to the much discussed role of the magnesium content and hardness of drinking water in the etiology of heart diseases.

Conclusions and hypotheses for further research

The main purpose of the present research has been to discover what types of factors are most important as determinants of geographic differences in mortality. The research has used data partly on counties in Finland, Sweden and Norway and partly on municipalities in Finland. The findings are preliminary and the study is being continued with more detailed data. At this phase we can complement the hypothetical model for the explanation of mortality differences presented at the beginning of this paper with the following observations and additional hypotheses:

1. General level of living is an important determinant of mortality. Its importance is, however, smaller in more developed areas (e.g. Sweden, Norway) than in less developed areas (Finland). It can also be hypothesized that the present living conditions are not as important as those conditions which prevailed during the time when the cohorts now dying were young.
2. The harmful effects of urbanization and industrialization on health are mixed with the beneficial effects of the rising level of living associated with industrialization. Therefore, high mortality is more clearly connected with urbanization in areas, where the general level of living has been relatively high for so long that it is no longer important in explaining mortality differences. The mechanisms through which urbanization is linked with higher mortality are not known, e.g. the role of selective migration has not been studied. Also, it should be taken into account that mortality in urbanized areas is not high among all groups, but death rates for older women are lower than average.
3. The quality of health care is associated both with level of living and level of urbanization. Therefore it is very difficult to separate its influence on mortality from the influence of these general factors.
4. Low winter temperatures may have some negative influence on mortality. This relationship is, however, probably conditioned by other factors such as level of living and occupational structure: climatic factors can be assumed to be more important when the level of living is low and when outdoor work is common.
5. This study has given slight support to the hypothesis that some characteristics of the soil, notably the shortage of magnesium or other trace elements, influence mortality. The association, however, is not likely to be linear.
6. The influence of dietary and smoking habits on mortality has not been analyzed in this study. On the basis of data from other studies, it seems unlikely that regional differences in these habits are great enough to satisfactorily explain the geographic variation of death rates in Finland. This does not

imply, however, that these factors are irrelevant to the explanation of the mortality differences among countries.

7. When the influence of the groups of factors mentioned above has been taken into account, some systematic regional variation is likely to remain. To explain this remainder, the genetic composition of the population and various cultural factors associated with, for example, level of social integration may be relevant.

References

- Aromaa, A., Björkstén, F., Maatela, J., Tamminen, M.: Seerumin kolesteroli ja triglyseridit kuudessa suomalaisessa väestöryhmässä. Helsinki 1973. Kansaneläkelaitoksen julkaisuja ML: 1/1973.
- Bolander, A.-M.: A comparative study of mortality by cause in four Nordic countries, 1966—1968, with special reference to male excess mortality, Statistiska centralbyrån, Statistiska meddelanden Be 1971:9.
- Dudley, E. F., Beldin, R. A. & Johnson, B. C.: Climate, water hardness and coronary heart disease. *J. Chron. Dis.* 1969, 22: 25—48.
- Forsdahl, A.: Momenter til belysning av den høye dødelighet i Finnmark fylke. *Tidsskrift for den norske lægeforening* 1973, 96: 661—667.
- Härö, A. S.: Kuolleisuus Suomessa ja muissa pohjoismaissa. *Duodecim* 1966, 82: 1136—1151.
- Härö, A. S.: Health care in Scandinavia, *Scandinavian Review*, 3, 1975, 10—19.
- Kannisto, V.: Kuolemansyyt väestöllisinä tekijöinä Suomessa. Helsinki 1947.
- Karvonen, M. I.: Tupakanpolto ja koronaaritauti. *Duodecim* 1972, 88: 1283—1289.
- Kitagawa, E. M. & Hauser, P. M.: *Differential Mortality in the United States: A study in socioeconomic epidemiology*. Cambridge, Mass. 1973.
- Koskinen, E. H.: Mitä suomalaiset syövät. *Sosiaalilääketieteellinen aikakauslehti* 1975, 13: 433—437.
- Leppo, K., Lindgren, J. & Ritamies, M.: Mortality trends in Finland in the 1960's. *Yearbook of Population Research in Finland* 1971, XII.
- Mortality and causes of death by regions 1964—1967 (SOS). The National Central Bureau of Statistics, Stockholm 1971.
- Notkola, V.: Sosiaalinen ympäristö, luonnonympäristö ja kuolleisuus. Tutkimus Suomen 1960-luvun alueellisista kuolleisuuseroista. Unpublished *pro gradu* thesis in sociology, University of Helsinki 1976.
- Pedersen, E., Magnus, K., Hougen, A., Bjelke, E., Hakama, M. & Saxen, E.: Lung cancer in Finland and Norway. Copenhagen 1969.
- Pekkarinen, M.: Tutkimuksia maalaisväestön ravinnosta eräissä Itä- ja Länsi-Suomen pitäjissä. Suomen maataloustieteellisen seuran julkaisuja 99.5. Hämeenlinna 1962.
- Regional mortality 1969—72, NOS A 672, Oslo 1974.
- Rimpelä, M. & Puska, P.: Miesten veren lipidit, verenpaine ja tupakointi Pohjois-Karjalassa ja Varsinais-Suomessa. *Sosiaalilääketieteellinen aikakauslehti* 1975: 1: 366—372.
- Sauer, H. I.: Geographic variation in mortality and morbidity, in Erhardt and Berlin (eds.). *Mortality and Morbidity in the United States*. Cambridge, Mass., 1974.
- Valkonen, T.: Elintaso, kaupungistuminen ja alueelliset kuolleisuuserot Suomessa, Ruotsissa ja Norjassa. Helsingin Yliopiston sosiologian laitoksen tutkimuksia No. 204. 1975.

Statistical sources

- Gebhard, A.: Asunto-olot. Tilattoman väestön alakomitea. Tilastollinen tutkimus yhteiskuntataloudellisista oloista Suomen maalaiskunnissa v. 1901. II. Helsinki 1916.
- Helimäki, U. I.: Taulukoita ja karttoja Suomen sateoloista kaudelta 1931—1960. Helsinki 1967.
- Kalimo, E.: Lääkintäpalvelusten ekologiset ominaisuudet. Kansaneläkelaitoksen julkaisusarja M: 4/1967.
- Kolkki, O.: Taulukoita ja karttoja Suomen lämpöoloista kaudelta 1931—1960. Helsinki 1966.
- Kurki, M.: Suomen peltojen viljavuudesta, vuosina 1955—1970. Viljavuuspalvelu Oy:ssä tehtyjen tutkimusten perusteella. Helsinki 1972.
- Official Statistics of Finland III, Agriculture 26:3. Yleinen maataloustiedustelu 1920—30 III.
- Official Statistics of Finland IV C:102. Population Census 1950 II.
- Official Statistics of Finland VI C:103. Population Census 1960 I, table 1.
- Official Statistics of Finland VI C:103. Population Census 1960 II, table 1.
- Official Statistics of Finland VI C:104. Population Census 1970 I, table C.
- Official Statistics of Finland VI C:104. Population Census 1970 I, tables 1 and 2.
- Official Statistics of Finland XI 62. Public Health and Medical Care 1960, table 42.
- Official Statistics of Finland. Statistical Surveys 44. Statistical Maps prepared on the basis of the 1960 Census.
- Official Statistics of Finland VI A. Vital Statistics 1961—1970, table 1.
- Official Statistics of Finland VI A:132. Vital Statistics 1970, table 40.
- Official Statistics of Finland VI A:133. Vital Statistics 1971, table 37.
- Official Statistics of Norway XII 157. Population Census 1960, vol. VI, table IV.
- Official Statistics of Norway. Population by age and marital status 31. 12. 1970, table 1.
- Official Statistics of Norway XII 91. Public Health and Medical care 1960, table 2.
- Official Statistics of Norway. Statistical Yearbook of Norway 1964, table 31.
- Official Statistics of Norway. Statistical Yearbook of Norway 1970, table 7.
- Official Statistics of Norway. Statistical Yearbook 1973, table 31.
- Official Statistics of Norway. Vital Statistics 1970, table 8.
- Official Statistics of Norway. Vital Statistics 1971, table 8.
- Official Statistics of Sweden. Housing census 1. 11. 1960 II, table 4.
- Official Statistics of Sweden. Population 31. 12. 1970, table 2.
- Official Statistics of Sweden. Public Health and Medical Care 1960, table 2.2.
- Official Statistics of Sweden. Vital Statistics, Summary for the decade 1921—30, table 4.
- Official Statistics of Sweden. Vital Statistics 1965, table 5.3. and 1966 table 5.2.
- Official Statistics of Sweden, Vital Statistics 1969, table 99 and 1970, table 105.
- Official Statistics of Sweden. Vital Statistics 1970. Part 3, table 4.1.
- Official Statistics of Sweden. Vital Statistics 1971. Part 3, table 4.1.
- Palmgren, K.: Kehittyneisyyden alueittaisista eroavuuksista Suomessa. Valtakunnan-suunnittelutoimiston julkaisusarja A: 15. Helsinki 1964.
- Statistical Yearbook of Finland 1969, table 35.
- Statistical Yearbook of Finland 1973, table 30.
- Statistical Yearbook of the National Pensions Institute of Finland 1970, tables 50 and 54.
- Statistical Yearbook of Sweden 1966, table 20.