Explanations for the Age, Sex, Spatial, and Temporal Structure of Czech Mortality for the Period 1987 – 97

Environment and Planning A.

Explanations for the age, sex, spatial and temporal structure of Czech mortality for the period 1987-97.

Jeroen Spijker

Abstract. This article disentangles the age, sex, spatial, and temporal structure of mortality in the Czech Republic during the period 1987-97 for a selection of cause-of-death categories and investigates possible socioeconomic and other causes for the uncovered mortality differences. The clarification of the major effects, as well as possible interactions between them, verifies whether causes of death should be modelled separately and which interactions should be tested in the explanatory analysis. For instance, it was considered inappropriate to incorporate district-level time-series data in the analysis when spatial mortality differences did not show significant variation between 1987 and 1997, as occurred with cancer and digestive system diseases. In this case the exogenous variables would take on the average for the study period.

1 Introduction

In this article, regional mortality in the Czech Republic (CR) during the period 1987-97 is studied. One reason for choosing this specific period was because it coincides with many economic, social and political changes that took place throughout Eastern Europe, even though it would appear that the transition had less influence on the health of the Czech population than for most other Eastern European countries as life expectancy in the CR actually increased during these 10 years by about 2.5 years among both sexes (Czech Statistical Office, 1999b), a view that is shared by Blažek and Dzúrová (2000).

For decades, the Czech Republic exhibited a stable pattern of regional mortality differentials, with very high levels in north-west Bohemia and north-Moravia, and low levels in south Moravia (Dzúrová, 2000). During the transition period, spatial differences in life expectancy in the CR persisted. According to Dzúrová (2000), mortality changes that took place between 1990/91 and 1995/96 were uneven and regionally highly differentiated, as changes were most intense in regions with traditionally extremely low and extremely high levels of mortality. Mortality in several north-Bohemian districts became more favourable, while the south-Moravian region experienced a moderate worsening of mortality conditions. The regional pattern of the mortality decline during this period was significantly associated with changes in mortality from the three most frequent causes of death: circulatory diseases, cancer and external causes.

As adult mortality from natural (i.e. non-external) causes is the result of many years of exposure to risk factors, it suggests that regional differentiation in such factors must exist in order to have produced these spatial mortality differences.

As the description and interpretation of regional patterns of Czech mortality for the period studied are well documented (e.g. Rychtaříková and Dzúrová, 1992; Blažek and Dzúrová, 2000; Burcin and Kučera, 2000; Dzúrová, 2000; Rychtaříková, 2004), this article focuses on exploring a new methodological approach to study mortality patterns at the aggregated level, whereby other studies will serve to compare and interpret the results.

The analysis is split up into two parts. Firstly, it sets out the conditions for studying the underlying causes behind mortality changes and differences by developing a model for the most important cause-of-death categories that adequately describes the data in terms

of the age, sex, temporal and spatial dimensions, maintaining a meaningful balance between model fit and parsimony. The model specifications are derived from log-linear regression analyses. While it is obvious that age- and sex-specific differences exist in cause-specific mortality, the main question to be answered is if the spatial and/or temporal differences in the age- and sex-mortality structure are statistically significant and whether the mortality structures are the same for the studied cause-of-death categories.

In the second part it is determined to what degree these observed spatial and temporal differences in mortality from the selected causes can be explained by exogenous variables. These include socioeconomic indicators, as well as demographic, environmental and health care variables, all of which were selected on the basis of established associations with mortality (see Table 1). Possible interactions with age and sex are also considered in order to improve the validity of the analysis, because the explanatory value of ecological mortality studies involving multivariate analyses has been questioned in the past due to the impossibility to impute relationships pertaining to individuals from statistical models applied to aggregated data as one cannot establish causal factors that lead to mortality changes (Jones and Moon, 1987). Others say, however, that ecological analysis can be justified in its own right by considering the effect of contextual factors such as environmental pollution that can only be measured at the population level. It has also been suggested that the identification of statistical associations between mortality and economic factors at the small area level (which the Czech districts are considered here) may be interpreted in conjunction with epidemiological investigations at the micro level (Higgs et al., 1998). This is why variables in this study have only been selected on known associations at the individual

level and why it is also proposed that inferences made on observed associations between dependent and independent variables have a stronger basis if the structure of the dependent variable is analysed before a multivariate analysis is performed. By discarding or including certain effects or interactions, what one is effectively doing is statistically specifying the data requirements of the covariates that will enable a more precise interpretation of any established association observed from the multivariate analysis. For instance, age- and/or sex-specific time-series data should be collected if the mortality data shows that the time trend also differs significantly by age and/or sex. Subsequently, the multivariate analysis should establish if the model coefficients of the chosen covariates actually significantly differ by age and/or sex.

2 Data

Yearly aggregate counts of age-, sex- and cause-specific mortality data for the period 1987-97 were obtained from the Czech Institute of Health Information and Statistics (UZIS) and population data from the Czech Statistical Office (CSO) for the 76 *okresy* (districts) that existed at that time. The five most common cause-of-death chapters and all remaining deaths were analysed (Table 2). Data on specific diseases were also available, but were only used to help to explain some of the model results.

Analysing cause-of-death groups is preferred to all-cause mortality because disease trends can be linked to trends in their risk factors that are sensitive to age- and cohort patterns (Tabeau et al., 2001). This becomes especially relevant when a particular cause of death dominates the total mortality pattern due to its high prevalence, as an all-cause model will be structurally biased towards it and other important developments may remain unobserved.

For the second analysis, data on socioeconomic and other characteristics of each district were obtained from several issues of the Czech Statistical Office publication "Okresy České Republiky" and several other sources (see Table 1). The three employment-structure indicators, wages and unemployment only apply to the working age population (15-64); educational attainment only include those older than 25; while the ethnicity, divorce, smoking, pollution, health and urban variables concerns the entire population. In order to capture the relationship between changes in socioeconomic and other factors and mortality, the intention was to collect and analyse time series data for each of the indicators. Unfortunately, this was only possible for unemployment, wages, divorce and SO_2 pollution. The other variables were only averages that covered several years.

As no reliable district-level smoking data could be found, the impact of smoking was estimated using the available age- and sex-specific lung cancer data. Given that up to 95% of lung cancer deaths in developed countries is caused by smoking (Peto et al., 1992) lung cancer is considered a useful proxy for smoking when there is no better alternative (Barendrecht et al., 2002). Indeed, after remodelling the data for respiratory system diseases, the results demonstrated a visible improvement (see results section). Due to the small number of annual fatalities in many districts, the calculated standardised death rates (SDR) for lung cancer pertain to the entire study period, rather than for each year separately.

The two health indicators hospital physicians and hospital beds were highly correlated (0.888). It was therefore decided to use only the one that showed the highest explanatory value when it was first entered into the cause-specific model.

Three national time-series indicators of national economic development were included in the analysis, namely: GDP, average wages and unemployment. Only the correlation

between wages and unemployment proved to be quite high (0.761). It should be noted that the possible simultaneous effect of regional and national wage pattern on mortality differences and changes was never tested simultaneously due to their high correlation (0.980). A three-year moving average for the annual level of GDP growth was used, as it exhibited a rather erratic trend, while mortality patterns change more slowly.

To facilitate the later interpretation of the results, each covariate was first individually regressed against standardised all-cause death rates in order to see if a non-linear transformation of the covariate was desirable (see Table 1).

3 Method

The first part of the analysis aims to detect time (T) and spatial (R) variations in the six cause-of-death categories to ascertain whether the causes of death should be combined or analysed separately. The resulting multidimensional demographic model describes mortality in terms of events that are simultaneously classified using five dimensions:

C: Cause of death

A: Age

S: Sex

R: District

T: Time (year)

Denoting the number of events (deaths) in a period t by X_{casrt}, then these events are linked to the mid-year population that is exposed to the risk of experiencing death (i.e. P_{asrt}) by the mortality rate M_{casrt}. Then, the model of interest will describe each mortality rate M_{casrt} as the product of a limited number of parameters such that the fitted number of deaths, $Y_{casrt} = M_{casrt} * P_{asrt}$, approximates to the observed number X_{casrt}. Given a model

specification, the parameter values can be estimated. Further, if one assumes that the X's are generated by a Poisson process, unbiased parameters of the model can be estimated using iterative proportional fitting (IPF), even if the data show large variability (over-dispersion) (Davies and Guy, 1987). The main advantage of using Poisson regression techniques over traditional methods of modelling mortality data (such as the ordinary least squares approach) occurs when there are few deaths (Lovett et al., 1986). It was thus appropriate here, as there were districts where no deaths occurred in certain age groups for particular cause-of-death categories.

The goodness of the model's fit is assessed by means of a log-likelihood ratio. This is equal to the 'deviance' D (a term used in the context of GLIM models). For each cause of death it has the form:

$$D = 2 \times \sum_{casrt} X_{casrt} \times \log\left(\frac{X_{casrt}}{Y_{casrt}}\right), \text{ where}$$

X_{casrt}. and Y_{casrt} are the observed and fitted number of cause-, age- , sex- , district- and time-specific deaths respectively.

Having assumed that the data were generated by a Poisson process, D may be compared to a χ^2 distribution with the same number of degrees of freedom as in the model (Francis et al., 1993), i.e. model B will be selected over model A when the difference in their deviances is less than the critical value U, i.e.

$D_B - D_A < U$, where $U = 0.5 \times \left(1.645 \times \sqrt{2N - 1}\right)^2$, $D_B \leq D_A$, in which N = the difference in the degrees of freedom between models A and B.

In this study, the resulting loglinear models always included the AS term (age + sex + age*sex) because age and sex patterns are dominant in the observed causes of death, while the sex mortality ratio also differs with age. The model was first parameterised by

age-specific rates, followed by relative mortality factors (ratios) for all other single effects and interactions whose average equals unity, i.e. the age- , sex- , district- and time-specific mortality rate for cause of death *c* according to the AS+AR+RT model equals:

$m_{casrt} = r_{as}^{AS} \times \tau_r^R \times \tau_t^T \times \tau_{ar}^{AR} \times \tau_{rt}^{RT}$, where

r_{as}^{AS} = age- and sex-specific mortality rate

τ_r^R = district-specific relative mortality factor

τ_t^T = time-specific relative mortality factor

τ_{ar}^{AR} = age- and district-specific relative mortality factor

τ_{rt}^{RT} = district- and time-specific relative mortality factor

The above model is just one from 114 possible ways in which the four variables A (age), S (sex), R (district) and T (time) can be combined for each cause of death; ranging from A+S+R+T (no interactions) to ASRT (the saturated model). Given the known age dependency in mortality rates and the fact that, with a few exceptions, each age-specific death rate is higher for men than for women regarding the causes of death selected, the model that only included the AS term was selected as the benchmark model. The object of this analysis was thus to ascertain whether the spatial and time variables significantly influenced the Czech mortality pattern, i.e. to what extent R and T reduced the deviance (D) and improved the fit when compared to the AS model.

The modelling procedure followed a particular sequence. Firstly, the five possible 2 * second-order models that included the AS term, as well as AS+R and AS+T were tested,

from which the one with the best fit was selected[1]. This model (e.g. AS+AR) was subsequently tested against the possible 3 * second-order models (e.g. AS+AR+RT) until the best model was found that was then used to test against 4 * second-order models, etc. until either no improvement could be made or all six second-order effects were included. Preference went to the most parsimonious models. Although third-order interactions were also analysed, the large increase in parameters only led to a slight improvement in the fit.

In the subsequent analysis, the GLIM4 statistical package (Francis et al., 1993) was used to estimate the proportion of deviance that could be explained by the selected exogenous variables in relation to the total deviance not accounted for by the AS term, i.e. the exogenous variables become substitutes for the district and time factors. The following specification was used in the modelling procedure, with the log of the population as the offset:

$$\ln M_{carst} = \ln P_{arst} + \mu_a^A + \mu_s^S + \mu_{as}^{AS} + \mu_{ar}^{AR} + \ldots$$

Subsequently, in order to obtain $\ln m_{carst}$:

$$\ln M_{carst} - \ln P_{arst} = \ln(M_{carst} / P_{arst}) = \ln m_{carst},$$

i.e. the explanatory model that was applied to each cause of death, i.e.:

$$\ln m_{carst} = \mu + \mu_a^A + \mu_s^S + \mu_{as}^{AS} + f^{AS}[X_{ir}^R, X_{it}^T, X_{irt}^{RT}, \beta], \text{ where}$$

m_{casrt} = the age- and sex-specific mortality rate for district r, time t and cause c;

μ = log rate of mortality for men aged 25-44 (the constant);

[1] In a model with four variables, a maximum of 6 second-order interactions is possible: i.e. AS, AT, AR, RT, RS and ST. A "2* second-order model" refers to a model that consists of two second-order effects and their single effects, e.g. AS+AR = A+S+R+AS+AR.

a = the additive log rate for the age groups 45-54, 55-64, 65-69, 70-74, 75-79, 80-84 or 85+[2];

s = the additive log rate for women;

f = linear function which may be age- and sex-specific

X_i = covariate X_i;

r,t = index of district r, time t;

β = the coefficient of the covariates.

Each explanatory variable was individually tested to see if its inclusion in the basic AS model would improve the explanation of the deviance. Subsequently, both forward and backward regression techniques were applied to test if the covariates should be included in the model. The decrease/increase in deviance that occurred when a variable was included/excluded in the model as well as the resulting parameter estimates were tested for statistical significance (95% level). Since the association between a particular health-related factor and mortality might differ between men and women and different age groups, interactions were also analysed. As the calculated deviance of the model usually showed over-dispersion, a scalar was used to estimate the parameters and correct the fitted values (the scalar equalling the deviance divided by the degrees of freedom; see Davies and Guy, 1987). The results were then used to determine which of the major effects and interactions needed further investigation. For example, if a particular cause of death showed no significant S*R interaction, then the exogenous variables at district level were not tested for a possible interaction with sex.

However, a coefficient on its own reveals little about the magnitude of the effect of the covariate while, at the same time, the various coefficients cannot be compared. This

[2] To produce more-robust results fewer age groups were used than in the first analysis.

dilemma was resolved by calculating the elasticity of a determinant using the formula $\varepsilon = \frac{dm/\bar{m}}{dx/\bar{x}} = \frac{dm}{dx}\frac{\bar{x}}{\bar{m}}$. This is equivalent to the relative change in mortality rate m in relation to the relative change in the covariate x, with dx and dm representing very small changes in x and m respectively and $\bar{x}$ and $\bar{m}$ district averages. Because in linear models the derivative $\frac{dm}{dx}$ translates to the regression coefficient β, the formula becomes $\varepsilon = \beta\frac{\bar{x}}{\bar{m}}$. When the covariate was transformed, the elasticity formula was adjusted: i.e. the elasticity simply equalled the coefficient of the covariate ($\varepsilon = \beta$) when the log was taken, $\varepsilon = \frac{-\beta}{\bar{x}}$ for the reciprocal transformation and $\varepsilon = \frac{\beta}{2}\sqrt{\bar{x}}$ for the square root transformation.

4 Results I: Age, sex, temporal, and spatial structures

After determining the best models (Table 3), results showed that the A*S and A*R interactions were important in terms of each cause of death over and above the effects from the four isolated factors: age, sex, district and time, i.e. both the sex ratio and spatial mortality differences varied significantly with age. The other interactions were important on fewer occasions: time did not interact with district in the models for cancer and digestive system diseases, and time did not interact with age in respiratory system diseases and remaining causes. Sex and district interacted significantly on three occasions (cancer, respiratory system diseases and external causes of death) and only once did time interact with sex (respiratory system diseases), implying that here the sex ratio changed

over time. Respiratory system diseases and external causes of death showed the greatest number of second order interactions (five in total).

As each model included the AS and AR terms, and one or both of RT and AT, it was investigated whether the mortality structure was the same for each cause of death, or if only a cause-specific factor was necessary for the varying levels of mortality. However, results showed that the six cause-of-death categories could not be analysed simultaneously. Furthermore, the structure of an aggregated model would be largely determined by the pattern of circulatory system diseases, as this group recorded by far the largest deviance in the data.

To describe the age, sex, district, time and the two-order interaction effects, the model parameters for each cause of death were calculated and plotted. The main results are described below. The age effect is expressed in rates, and the other significant effects and interactions in terms of relative mortality figures (RMF). Results are not described in detail for each district, but in case of clear geographic patterns, reference is made to the larger *kraj* or 'region' (see Figure 1).

4.1 Single effects

As shown in Figure 2, most causes of death exhibit very clear age-specific patterns: a peak at age 0, followed by a sharp decline, reaching the lowest point at ages 5-9 or 10-14, and subsequently rising in a general log-linear fashion. There are two exceptions to this rule: cancer and external causes of death. Cancer shows little variation across ages 0-29, after which the rate increases log-linearly, until the age of 55 after which the rate of increase declines. The other exception, external causes, shows a sharp increase in the

mortality rate between age groups 10-14 and 15-19. The rate then increases very slowly until age group 65-69, after which it accelerates. The pattern for the population of working age is not surprising given that this age group is particularly exposed to the risk of dying from unnatural causes related to working conditions and lifestyle.

Male-female mortality differences are illustrated in Figure 3. Each cause shows excess male mortality, with the largest differences being found for respiratory system diseases and the smallest for remaining causes of death.

Clear spatial RMF patterns were distinguished for each cause of death (see Figure 4). The largest differences between districts were recorded for respiratory system diseases (RMF ranging between 0.53 and 1.76), while the smallest differences were, as expected, observed for the most frequently occurring cause of death circulatory system diseases (between 0.86 and 1.23). Particularly for cancer, external causes and remaining causes of death, the highest levels of mortality were recorded in the west part of the country. Regarding circulatory system diseases, levels were relatively high in some districts in the Central-Bohemia region, although in Prague (an enclave in the centre of this region) levels were very low. Similarly, Prague and several of its neighbouring rural counties showed very low levels of respiratory system diseases, particularly when compared to mining districts in the west and in the far east of the country. The digestive system disease pattern was quite specific, with the highest levels being found in North-Bohemia and North-Moravia.

The last of the four main effects, time, showed a general decline in mortality from 1990 onwards (1989 for respiratory system diseases). This was most spectacular with remaining causes of death and respiratory system diseases, whereas cancer mortality only declined marginally (Figure 5), reasons for which will be provided in the discussion.

4.2 Interactions

Not surprisingly, sex differences were not uniform across age. In general, when compared to the overall sex mortality ratios, the female advantage was lowest among the youngest and oldest age groups, while male excess mortality was particularly high at middle age. The exception was cancer, because there was a female peak between ages 30 and 54 that could largely be attributed to breast cancer, while males observed particularly high mortality from lung cancer between ages 55 and 79. Only in the case of external causes above the age of 85 was the mortality rate higher among women (as a result of accidental falls).

The age structure of mortality was also statistically different across districts. In some cases, the regional pattern for an age-group was very similar to the all-age pattern, especially when it pertained to an age group with high absolute levels of mortality (e.g. compare total respiratory system disease mortality pattern (Figure 4c) with that for age-group 75-84 (Figure 6)).

In terms of the interaction between district and time, there were significantly different spatial differences in changes over time for circulatory system diseases, respiratory system diseases, external causes and the remaining deaths category. For example, Prague and several districts in East-Bohemia observed the largest declines in circulatory system diseases between 1987 and 1997, and mining districts in external causes (results not shown).

The A*T two-order interaction effect was significant for cancer, circulatory system diseases, digestive system diseases and external causes. If we disregard mortality under the age of 25 as the age-time pattern is rather erratic due to the small number of deaths, we see in Figure 7 that although cancer mortality was virtually stable over the study

period (as shown earlier in Figure 5), it declined among those below the age of 50 and increased slightly for those aged 80+. Circulatory system diseases showed a disproportionately large decline in mortality among the middle aged, while the 80+ category saw a below average decline. In terms of digestive system disease mortality, smaller decreases were seen among the 35-54 age group than in the population as a whole, while the 65-79 year olds demonstrated the largest relative improvement. Although external causes saw an overall decline between 1987 and 1997, a sharp increase occurred among 25-64 year olds, which can partly be ascribed to the increase in the number of murders after 1989 (WHO, 2001).

Spatial variations in the sex ratios of mortality (the S*R interaction) were significant for cancer, respiratory system diseases and external causes. In general, sex differences were not related to the overall spatial pattern. The differences between the sexes with respect to cancer were particularly small in the cities of Prague, Brno and Plzen, and in East-Bohemia, and were largest in the rural districts situated in Central-, South- and West-Bohemia, although the range in relative differences was quite small. High ratios were recorded in the mining areas in North-Bohemia and North-Moravia for respiratory system diseases (related to coalworker's pneumoconiosis (ICD10: J60)), while lower than average sex mortality differences were observed in West-Bohemia. With regard to external causes of death, the largest sex mortality ratios (M:F) were observed in districts with low female external cause of death mortality (with the correlation coefficient R equalling –0.65), especially in the two Moravian provinces.

The last two-order interaction effect investigated was between sex and time (S*T) and only proved significant for respiratory system diseases, as mortality rates declined more rapidly for men than for women during the study period (results not shown).

5 Results II: explanatory analysis

The objective of the second analysis is to provide an explanation for the spatial and temporal variation in the Czech mortality data that was found in the previous section using a selection of socioeconomic and other variables. As external and remaining causes of death are two quite heterogeneous groups they were excluded from further analysis. The method that was used was partly influenced by the outcomes from the previous section. As few deaths occurred before the age of 25, only mortality after this age was analysed. Given that working age mortality was also fairly limited, some of the remaining age groups were subsequently combined to leave the following eight age categories: 25-44, 45-54, 55-64, 65-69, 70-74, 75-79, 80-84, and 85+. Another consequence of the outcome of the previous analysis was that the available time-series data were not used in the explanatory analysis for cancer and digestive system diseases because the earlier analysis showed no significant R*T interaction. However, because these variables may still be responsible for some of the spatial variations in mortality from these two causes of death, they were incorporated as cross-sectional data by applying the average for the study period. In addition, the smoking proxy was not used in the analysis of cancer to avoid a circular argument, as 32% of all male and 8% of all female cancer deaths were due to lung cancer.

The explanatory analysis focused on the geographic and time variations in mortality data after controlling for age-, sex- and age*sex-effects. In the previous section, the best second-order interaction models for each cause of death were determined, explaining between 15% and 58% of the deviance that remained after excluding the AS term (Table 3) and between 17% and 50% when recalculated for the fewer age groups (Table 4 – column 2). Although the analysis in the previous section was conducted to determine the

type of data required for this analysis, in practice not all the desired types of data could be obtained. For example, there were no data on age-specific regional or national factors and only limited sex-specific data were available. One way to get around this problem was to test the interaction effect between either the age- or the sex-specific death rates and the covariate under consideration. Although this produced additional parameters (seven extra for the age interaction and one extra for sex), the advantage of this method was the gain in explanatory value. For example, if age- or sex-specific mortality interacted significantly with one of the covariates, the parameter values indicated which age groups or sex were the most or least affected by this variable, or if the association was in opposing directions at different ages or between men and women. For secondary sector employment, sex-specific data were obtained because the nature and associated risk factors of secondary sector employment often differ substantially between men and women (e.g. 'heavy' industrial employment versus 'light' assembly work), thus making them distinct indicators. Concomitantly, this provided the opportunity to test whether districts with high levels of either male or female secondary sector employment could be associated with mortality levels associated with the opposite sex. This is plausible, as social class not only affects the behaviour of the individual, but also that of their family. Column 3 in Table 4 displays the models that were selected once both the previous results and the data collected for the explanatory analysis were considered.

After replacing the spatial and time effects with the socioeconomic and other variables, the covariates explained between 6% and 29% of the deviance remaining after the benchmark model had been taken into account (column 6 in Table 4), which amounts to between 34% and 71% of the total deviance due to the spatial and temporal effects of the

tested models. In other words, the selected variables captured, for example in the case of cancer, 58% of the deviation that was explained by the R+T+SR tested model. Depending on the best model, interactions with age and sex were tested for each cause of death but only helped to explain a few additional percentage points. The largest improvement was observed for cancer. While the spatial and temporal covariates explained 11% of the proportion of deviance not attributed to the AS term, the additional interactions increased this to 15% as certain variables showed different effects by age and between men and women. In the case of the respiratory system disease model, the third-order interaction of age with the sex-specific lung cancer data was also tested. Although this model is therefore slightly different from the earlier-established best model, it made theoretical sense to do so as smoking is known to affect respiratory system disease mortality less at young and very old ages. The R*T covariates were also tested for a possible significant effect across age groups or for differences in the effect on men and women when the R*T interaction effect was significant. The main results of the GLIM analysis are described below. The elasticities that are referred to are given in the Appendix.

Although one should consider that there were more socioeconomic indicators than other types, the covariates that were found to be principally responsible for the region and time specificity in the data were socioeconomic ones. Particularly regional wage levels were important in reducing the deviance left unexplained after the benchmark model, being associated with each death-of-death category except cancer. The largest contributions were observed in the models for circulatory system diseases. The association was negative, except for digestive system diseases. The latter may be explained through

secondary sector employment as the socialist legacy allowed manual labourers employed in key industries such as mining to earn more than most other workers and as is known (e.g. Davey Smith et al., 1998, Balarajan and McDowall, 1988) manual workers tend to live less healthy.

Another variable was unemployment. It was the major contributor towards explaining the regional differences in digestive system diseases. Stress and unhealthy lifestyles, particularly excess alcohol consumption, are known risk factors of digestive system diseases that are also more prevalent among the unemployed. The effect of unemployment was largest among the working age population and did not differ much between men and women. Since unemployment is still a new phenomenon in the CR – there was no unemployment in 1987 and the highest district level was 12.4% in 1997 – it was not sufficiently prevalent to have a marked effect on the health of the working population with regard to the other causes, even though individuals are likely to have succumbed to the stress of being unemployed, while in order to ascertain an effect on the retired population, a time lag will need to be incorporated in the data.

All three employment sector indicators, i.e. primary sector employment, male and female secondary sector employment, had substantial associations with mortality and generally in the expected direction. However, there was a positive association between primary sector employment and respiratory system diseases, although this may be because healthier lifestyles, better housing conditions and cleaner air, i.e. characteristics of rural districts, do not sufficiently compensate for a less accessible and inferior health care. The observed negative association between primary sector employment and digestive system diseases coincides with previous observations that both alcohol consumption and chronic liver disease (LDC) mortality is lower in rural than in urban settings

(Mackenbach and Verkleij, 1997; OPCS, 1978). LDC is the most important sub-category of digestive system diseases (43% of its total during the study period), whereby alcohol is known to contribute up to 85% of LDC deaths (Nizard and Nuñoz Perez, 1994).

In a number of instances, clear age-specific differences were found in the association between a covariate and a cause-of-death category. For instance, the association between male secondary sector employment and circulatory system diseases first increased with age before declining and was strongest among 45-54 year olds, while the coefficient estimates of primary and male secondary sector employment sector indicators clearly increased with age in the respiratory system disease model. Especially in the latter example it suggests that negative effects may accumulate over time. As this finding also applied to female mortality, it suggests that industrial employment does not only affect mortality through occupational hazards but is also associated with lifestyle and other factors conducive to respiratory system diseases than those included in the analysis. However, the third employment variable, female secondary sector employment, did not show an interaction with age, although its association with mortality was conform expectation (negative with cancer, circulatory and respiratory system diseases). Due to the higher elasticity recorded for men in the cancer model and a lack of interaction with sex regarding the other two causes of death it suggests that female employment in (usually light) industries also has an impact on men's health, i.e. the health effects of material and non-material rewards received by women for participating in the labour force also extend to (their) men. On the other hand, male secondary sector employment only seemed to have had an effect on male cancer.

The last of the socioeconomic variables, education, was most important in explaining cancer mortality differences, whose expected negative effect decreased with age and was

greatest for men, but was not as important in explaining spatial mortality differences as one might expect. This was in part because the addition of the smoking proxy removed much or all of the significant effect of education in the circulatory and respiratory system disease models[3], a plausible result given that highly educated people are better informed about smoking and are also less likely to smoke due to other psychosocial and material factors associated with education. A likely reason why some independent effect of education in the circulatory system disease model still remained is because education may lead to a reduction in other health damaging behaviours associated with heart disease, including dietary factors. However, due to the different political situation, labour and remuneration systems and regional development policies in the East before the socio-political transformation of the 1990s a materialist explanation was not considered for the then observed educational mortality differences (unlike in western European countries), as education was not rewarded by higher income or other material privileges and the distribution of income was more equitable than in western countries (Bobak et al., 1999).

The other variables (divorce, smoking, the percentage of Gypsies in the population, pollution, urbanisation and the number of physicians) were relatively unimportant.
Divorce, significant in both cancer and circulatory system disease models, showed a positive association although without an interaction with age and sex.
As to smoking, lung cancer proved to be a useful proxy, although it was never the most important variable in terms of deviance explained. For circulatory and respiratory system disease mortality, its effect declined by age after reaching retirement age.
In the case of the Gypsy variable – significant in three of the four causes of death analysed – stronger associations would have observed if the smoking proxy was

[3] The lung cancer smoking proxy was not tested in the cancer model due to circular referencing.

excluded. No differences were found between men and women and only in the case of digestive system diseases did any interaction take place with age.

SO2 pollution was the only environmental indicator investigated. As its association with mortality was negative on three occasions, it seemed that its health damaging effect was confounded by other factors. For example, the respiratory system diseases model results showed that after excluding wages its association with SO2 pollution turned positive. As we know that until the mid-1990s districts with high average wages were often mining districts with high levels of pollution, the effect of pollution seems to be underestimated.

The association between urbanisation and mortality is rather complex, because urban influences on health can be both positive and negative, being related to cultural factors (e.g. more divorcees), the environment (e.g. pollution) and access to health care (Mackenbach et al 1991; Van Poppel, 1991). In the analysis, a strong positive association was found with respiratory system diseases, with elasticities increasing with age and being higher for men than for women. This seems to indicate that there are factors associated with urbanisation that are detrimental to a population's health which become manifested at older ages due to an accumulative effect.

Regarding the health care indicators, the number of hospital beds per capita did not contribute to explaining spatial mortality differences. Conversely, the average number of in-patient physicians did demonstrate an influence on the level of circulatory and respiratory system diseases. The effect was small and negative without an interaction with age or sex.

Finally, economic contextual covariates were included to see if they would explain part of the national trend over time. However, results showed that the elasticities that corresponded to the covariates were usually fairly insignificant, although average Czech

wages played some role in explaining time changes in cancer mortality among young and middle-aged adults. The average national employment rate appeared to be more an indicator of economic development than economic uncertainty as its association with mortality was, although weak, consistently negative.

6 Discussion

In earlier research it has been suggested that recent economic changes in the CR resulted in discernible shifts in economic variables both in relation to each other as well as to other variables, in which especially wages has appeared as an instigator of socioeconomic health differences in the CR, particularly among men (Blažek and Dzúrová, 2000). The importance of wages was also established in this study, whereby when sex differences were established, the effect was generally greater for men. With regard to male and female secondary sector employment it was interesting to note that both variables affected mortality levels of the opposite sex. This suggests that certain lifestyle aspects surrounding gender-specific occupations have an influence on the health of partners. As the analytical method also permitted to analyse interactions between the exogenous variables and age, when covariates recorded the highest associations among older age categories, it was possible to make a cautious reference to the life course, for instance that the positive or adverse effects of the socioeconomic environment are due to cumulative exposures over the entire life course (Kuh *et al.*, 1997). This was particularly the case with respiratory system disease mortality, as significant age differences in the coefficients were observed for urbanisation, primary and secondary (i.e. industrial) sector employment. With regard to the other cause-of-death categories, the effects of labour related factors generally declined with age, such as with circulatory system diseases

(wages and male secondary sector employment) and digestive system diseases (unemployment and primary sector employment), implying that its largest effect occurred while still in the labour force. Several of the covariates also showed a quadratic form of association with mortality. For example, the effect of smoking on respiratory system diseases mortality was greatest between ages 55 and 74.

This type of analysis may particularly benefit those working in public policy, as in addition to regional mortality patterns it provides information on the impact of health damaging or promoting factors, which, as we have seen, may differ depending on age and sex. For instance, if the highest impact of a factor is among a specific age-group and mortality (or morbidity) rates of the associated disease are most elevated in a specific part of the country, it may be more effective to only direct policy at this specific sub-group of a country's population (for instance, anti-alcohol campaigns aimed at young unemployed men living in mining regions). Estimating the magnitude of the effect of health damaging or promoting factors on mortality should also benefit those who incorporate exogenous factors in their mortality forecasts (which is, incidentally, rarely done). For example, a continued decline in secondary sector employment should result in further declines mortality even if other factors remain constant.

Regarding several of the other results, the first analysis showed that respiratory system disease mortality in the CR declined by a third during the 11-year study period. The literature suggests that declining respiratory system disease mortality in the CR and other Eastern European countries are partly due to improvements in the quality of the natural environment resulting from economic restructuring provoked by the general recession within the industrial and mining sectors that caused both a huge outflow of workers,

reducing the population at (high) risk, and reducing emissions through the substantial fall in industrial outputs and environmental measures at both the local (e.g. the installation of more environmentally sensitive technology) and national (e.g. environmental legislation) level (UNECE, 2001; Blazek and Dzúrová, 2000; OECD, 1999). While in this study the effects of environmental measures were not tested and pollution explained only a small part of respiratory system mortality differences, the male and female secondary sector employment indicators were important. High elasticities were observed for each sex with both employment indicators affecting mortality of the opposite sex. This suggests that non-working conditions such as lifestyle and housing conditions are also important in shaping regional mortality patterns.

A smaller decline, though not less in importance given its share of total mortality, was observed for circulatory system diseases. One important reason for this decline and considered more important than behavioural changes are the changes that were introduced to the health care system that improved the quality of health services, in particular the increase in public expenditure allocated to health care, the upgrading of medical equipment and sharp increases in invasive operations (Blazek and Dzúrová, 2000). Also the regional health indicator tested here in the second analysis (number of in-patient physicians) demonstrated a small negative association with circulatory system disease mortality, while the smoking proxy observed lower elasticities than most economic covariates. Placed in an international context, we know that circulatory system disease mortality has declined for some time in market economies and that most improvements, especially among women, now occur at older ages. It appears that the CR has also entered this process, known as the "age of delayed degenerative diseases" within the context of the epidemiological transition (Olshansky and Ault, 1986) and partly

explains why cancer mortality only decreased marginally during the study period as survivors were left to be exposed to cancer at later ages.

If we consider the differences between the observed and modelled age-specific death rates, then each final model was slightly underspecified (see Figures 8a-8f), suggesting that certain district-level health-related factors still need to be taken into account. A likely candidate for whom no district-level data could be obtained was an indicator of alcohol consumption, which would undoubtedly have altered the model for digestive system diseases. Other potential explanatory variables worth investigating in the future include material deprivation (e.g. housing conditions), psychosocial factors (e.g. employment related stress), medical technologies and the efficacy of health care (e.g. mortality from conditions amenable to medical intervention; c.f. Mackenbach et al., 1990) and smoking data to analyse cancer mortality. In addition, for some of the tested variables regional data could only be obtained for one or several years and given the rapid economic changes that occurred during the period studied, completed time series would undoubtedly have improved the explanatory analysis of this study.

Another issue concerns the operationalisation and measurement of the pollution indicator. As only district-level SO_2 *emission* data could be obtained and districts responsible for large amounts of pollution are likely to affect the air quality of neighbouring districts, the health damaging effect of their own emissions are in reality less. Neither did the indicator take other pollutants into account.

The final methodological issue relates to the problem of confounding and stresses the general predicament of ecological studies, i.e. that one cannot explicitly ascertain causal relationships (Gravelle, 1998). There were two results that prompted the suggestion of

confounding given their implausibility theoretically and from scientific evidence: i) The positive association between wages and digestive system diseases. Both individual based (e.g. Singh and Hoyert, 2000) and small area (e.g. Middelkoop et al., 2001[4]) research indicate a negative association concerning its largest subcategory LDC. One reason for the observed anomaly is the strong positive correlation between wages and male secondary sector employment (indeed, when wages was replaced by male secondary sector employment, the association between the employment variable and digestive system disease mortality was both positive and significant). ii) The associations between pollution and mortality linked to circulatory, respiratory and digestive system diseases were negative. It is possible that the variable represents not only an environmental indicator but also a type of economic development not covered by the wage variable.

7 Conclusion

In attempting to explain spatial mortality patterns, regions may be compared in several different ways. For example, a mortality measurement can be calculated, and by grouping small administrative districts based on various socioeconomic or other characteristics, they can be compared. These characteristics may either be from individual (e.g. Fox et al., 1984) or surrogate measures derived from knowledge of the area (e.g. Spijker, 2000). Another method is to apply regression analyses to measurements of mortality and a number of explanatory variables (e.g. Mackenbach et al., 1991). In this article a different approach was used that consisted of two parts: disentangling the age, sex, spatial and time structure of mortality for six cause-of-death categories before investigating possible socioeconomic and other causes for the

[4] Rather than wages on its own, a deprivation score was used in this study which comprised of two interrelated factors: family income and unemployment.

uncovered differences in time and space. It was considered that the clarification of the major effects and possible interactions between them would not only verify whether causes of death should be modelled separately rather than combined (i.e. all-cause mortality), but also if a complete spatial and temporal analysis would make any sense. For example, if spatial mortality differences from a particular cause of death did not show a significant change over the period studied, incorporating longitudinal district-level data would be unnecessary (this was the case with cancer and digestive system diseases). Similarly, it would not be logical to estimate age- or sex-specific coefficients for (the spatial- or temporal-specific) explanatory variables if no interaction had been established in the mortality data between on the one hand age or sex and on the other hand district or time. In addition, analysing age-specific mortality rather than an age-standardised measure allows for a more precise analysis of how certain exogenous variables affect mortality, especially when causes of death are analysed.

This study has shown that, in a former socialist country, regional mortality differences do exist and are, by and large, related to socioeconomic factors, which are in part due to the rigorous economic and social planning of the past. Although death is, in most instances, the result of the behaviour of individuals, it seems that in the CR certain socioeconomic groups are more susceptible to adverse behaviour than others, and that more than 40 years of state planning has resulted in some of such groups being concentrated in certain areas where new industries were created. It seems though that in the CR, as elsewhere, the wealthy and well-educated live longer, those employed in industry do not, and neither do those who live in very urbanised areas or come from an ethnic minority. However, the cause-specific analysis has revealed that for the studied period not all of the established western associations were applicable in the CR as the socialist remuneration legacy led to

a positive association between digestive system diseases and wages. Given the growing regional differentiation of the mechanisms of economic transformation, economic conditions already contribute to the explanation of regional mortality differences (Dzúrová, 2000), this association is unlikely to have remained so. The next step, therefore, is to analyse the 10 years after the end of the socio-political transformation to ascertain if and how these and other associations between socioeconomic and other indicators and cause-specific mortality have changed.

References

Balarajan R, McDowall M E, 1988, "Regional socioeconomic differences in mortality among men in Great Britain today" *Public Health* **102**(1): 33 – 43

Barendregt J J, Looman C W N, Brom-Hansen H, 2002, "Comparison of cohort smoking in Denmark and The Netherlands" *Bulletin of the WHO* **80**(1): 26 – 32

Blažek J, Dzúrová D, 2000, "The decline of mortality in the Czech Republic during the transition period: A counterfactual case study", in *The mortality crisis in transitional economies* Eds GA Cornia, P Paniccià (Oxford University Press) pp 303 – 327

Bobak M, Feachem R G A, 1995, "Air pollution and mortality in Central and Eastern Europe: An estimate of the impact" *European Journal of Public Health* **5** 82 – 86

Bobak M, Hertzman C, Skodova Z, Marmot M (1999), "Socioeconomic status and cardiovascular risk factors in the Czech Republic". *International Journal of Epidemiology* **28** 46 – 52

Burcin B, Kučera T, 2000, "Changes in fertility and mortality in the Czech Republic: An attempt of regional demographic analysis", in *New demographic faces of Europe* Eds T Kučera, O Kučerová, O Opara, E Schaich (Spring, Berlin) pp 371 – 417

Czech Statistical Office, 1992-98, *Okresy České Republiky v roce 1992-98* (Districts of the Czech Republic, 1992-98) (CSO, Prague)

Czech Statistical Office, 1999a, Time-series of basic indicators of labour statistics (1948-1999) (CSO, Prague)

Czech Statistical Office, 1999b, *Statistical Yearbook of the Czech Republic 1999* (CSO, Prague)

Davey Smith G, Hart C, Hole D, MacKinnon P, Gillis C, Watt G, Blane D, Hawthorne V, 1998, "Education and occupational social class: which is the more important indicator of mortality risk?" *Journal of Epidemiology and Community Health* **52** 153 – 190

Davies R, Guy C, 1987, "The statistical modelling of flow data when the Poisson assumption is violated" *Geographical Analysis* **19** 300 – 314

Dzúrová D, 2000, "Mortality differentials in the Czech Republic during the post-1989 socio-political transformation" *Health & Place* **6** 351 – 362

Fox A J, Jones D R, Goldblatt P O, 1984, "Approaches to studying the effect of socioeconomic circumstances on geographic differences in mortality in England and Wales" *British Medical Bulletin*, **40**(4) 309 – 314

Francis B, Green M, Payne C (Eds), 1993, *The GLIM system: Generalized linear interactive modelling* (Oxford University Press, Oxford)

Gravelle H, 1998, "How much of the relation between population mortality and unequal distribution of income is a statistical artefact*?" British Medical Journal* **316** 382 – 385

Higgs G, Senior M L, Williams H C W L, 1998, "Spatial and temporal variation of mortality and deprivation 1: Widening health inequalities" *Environment and planning A* **30** 1661 – 1682

Jones K, Moon G, 1993, "Medical geography: taking space seriously" *Progress in Human Geography* **17**(4) 515 – 524

Kardaun J W P F, Glerum J J, 1995, *Sterfte, doodsoorzaak en inkomen*. (CBS, Voorburg)

Kuh D, Power C, Blane D, Bartley M, 1997, "Social pathways between childhood and adult health", in: Kuh D, Ben-Shlomo Y (Eds) *A life course approach to chronic disease epidemiology* (Oxford University Press, Oxford)

Kunst A E, Groenhof F, Mackenbach J P, EU working group on socioeconomic inequalities in health, 1998, Occupational class and cause-specific mortality in middle aged men in 11 European countries: comparison of population bases studies. *British Medical Journal* **316** 1636 – 1642

Lovett A A, Bentham C G, Flowerdew R, 1986, "Analytical geographic variations in mortality using poisson regression: the example of ischemic heart disease in England and Wales 1969-1973" *Social Science and Medicine* **23**(10) 935 – 943

Mackenbach J P, Bouvier-Colle M H, Jougla E, 1990, ""Avoidable" mortality and health services: a review of aggregate data studies" *Journal of Epidemiology and Community Health* **44**(2) 106 – 111

Mackenbach J P, Kunst A E, Looman C W N, 1991, "Cultural and economic determinants of geographical mortality patterns in The Netherlands" *Journal of Epidemiology and Community Health* **45** 231 – 237

Mackenbach J P, Looman C W N, 1994, "Living standards and mortality in the European Community" *Journal of Epidemiology and Community Health* **48** 140 – 145

Mackenbach J P, Verkleij H (Eds) (1997), *Volksgezondheid Toekomst Verkenning: II. Gezondheidsverschillen* (RIVM, Bilthoven, the Netherlands)

Middelkoop B J C, Struben H W A, Burger I, Vroom-Jongerden J M, 2001, "Urban cause-specific socioeconomic mortality differences. Which causes of death contribute most?" *International Journal of Epidemiology* **30** 240 – 47

Nizard A, Muñoz-Perez F, 1994, "Alcohol, smoking and mortality in France since 1950" *Population* **6** 159 – 194 English edition

Nolte E, McKee M, 2004, Does health care save lives? Avoidable mortality revisited (Nuffield Trust, London)

OECD, 1999, Environmental performance reviews: Czech Republic (OECD, Paris)

Osler M, Christensen U, Lund R, Gamborg M, Godtfredsen N, Prescott E, 2003, High local unemployment and increased mortality in Danish adults; results from a prospective multilevel study. *Occupational and Environmental Medicine* **60**(e16)

Olshansky S J, Ault A B (1986), The fourth stage of the epidemiologic transition: The age of delayed degenerative diseases. Milbank Quarterly **64**(3) 353 – 391

OPCS, 1978, Occupational mortality 1970-1972 (HMSO, London)

Peto R, Lopez A D, Boreham J, Thun M, Heath Jr C, 1992, Mortality from tobacco in developed countries: indirect estimation from national vital statistics *Lancet* **339** 1268 – 1278

Rychtaříková J, Dzúrová D, 1992, "Les disparités géographiques de la mortalité en Tchécoslavaquie" (Geographic inequalities in mortality in Czechoslovakia) *Population* **47**(3) 617 – 644

Rychtaříková J, 2004, "The Case of the Czech Republic. Determinants of the Recent Favourable Turnover in Mortality" *Demographic Research* **S2** 105 – 138

Singh G K, Hoyert D L, 2000, "Social epidemiology of chronic liver disease and cirrhosis mortality in the United States, 1935-97: trends and differentials by ethnicity, socioeconomic status, and alcohol consumption" *Human biology* **72**(5) 801 – 20

Spijker J J A, 2000, "A regional study of socioeconomic differences in life expectancy in the Czech Republic" *NIDI Working Paper No. 2000/3* (NIDI, The Hague)

Srb V, 1990, "Územní rozptyl československých Romů v letech 1968-1988" (Spatial dispersion of Czechoslovak Roma in the years 1968-1988) *Demografie* **32**(2) 169 – 174

Tabeau E, Van den Berg Jeths A, Heathcote C (Eds), 2001, Forecasting mortality in developed countries: Insights from a statistical, demographic and epidemiological perspective (Kluwer Academic Publishers, Dordrecht)

UNECE, 2001, "Assessment of progress in sustainable development since Rio 1992 for Member States of the United Nations Economic Commission for Europe" *CEP/AC.12/3* (UNECE, Geneva)

Valkonen T, Martelin T, Rimpelä A, Notkola V, Savela S, 1993), "Socioeconomic mortality differences in Finland, 1981-90" *Population* **1** (Statistics Finland, Helsinki)

Van Poppel F W A, 1991, "Roman Catholicism and regional mortality differences in the Netherlands, 1850-1933" Paper presented at the European Population Conference, 21-25 October 1991, Paris. Copy available from NIDI Library, Lange Houtstraat 19, 2511 CV The Hague

WHO, 2001, *WHO Mortality Database* (30 January update) http://www3.who.int/whosis/whsa/ftp/download.htm.

Appendix Elasticities of covariates (including interactions) for each cause of death

Covariate	Sex	Age	Cancer	Circ.	Resp.	Digestive
Unemployment	-	-				
	-	25-44				0.333
	-	45-54				0.343
	-	55-64				0.242
	-	65-69				0.158
	-	70-74				0.053
	-	75-79				0.111
	-	80-84				-0.024
	-	85+				-0.015
	men	25-44				
	men	45-54				
	men	55-64				
	men	65-69				
	men	70-74				
	men	75-79				
	men	80-84				
	men	85+				
	women	25-44				
	women	45-54				
	women	55-64				
	women	65-69				
	women	70-74				
	women	75-79				
	women	80-84				
	women	85+				
Wages	-	-				0.720
	men	-			-0.250	
	women	-			-0.042	
	-	25-44		-0.303		
	-	45-54		-0.256		
	-	55-64		-0.262		
	-	65-69		-0.232		
	-	70-74		-0.230		
	-	75-79		-0.213		
	-	80-84		-0.161		
	-	85+		-0.112		
	men	25-44				
	men	45-54				
	men	55-64				
	men	65-69				
	men	70-74				
	men	75-79				
	men	80-84				
	men	85+				
	women	25-44				
	women	45-54				
	women	55-64				
	women	65-69				
	women	70-74				
	women	75-79				
	women	80-84				
	women	85+				
Primary	-	-		-0.020		
sector	-	25-44			0.190	-0.255
employment	-	45-54			0.095	-0.217
	-	55-64			0.152	-0.129
	-	65-69			0.226	-0.104
	-	70-74			0.304	-0.064
	-	75-79			0.430	-0.058
	-	80-84			0.460	-0.024
	-	85+			0.476	-0.032
	men	25-44	0.020			
	men	45-54	0.039			
	men	55-64	0.044			
	men	65-69	0.041			
	men	70-74	0.041			
	men	75-79	0.067			
	men	80-84	0.068			
	men	85+	0.062			
	women	25-44	-0.032			
	women	45-54	-0.014			
	women	55-64	-0.009			
	women	65-69	-0.012			
	women	70-74	-0.011			
	women	75-79	0.014			
	women	80-84	0.016			
	women	85+	0.010			

Appendix Continued

Covariate	Sex	Age	Cancer	Circ.	Resp.	Digestive
Male	men	-	0.272			
	women	-	0.022			
secondary sector employment	-	25-44		0.298		
	-	45-54		0.437		
	-	55-64		0.384		
	-	65-69		0.351		
	-	70-74		0.283		
	-	75-79		0.213		
	-	80-84		0.153		
	-	85+		0.111		
	men	25-44			0.797	
	men	45-54			0.868	
	men	55-64			1.088	
	men	65-69			1.257	
	men	70-74			1.294	
	men	75-79			1.540	
	men	80-84			1.638	
	men	85+			1.778	
	women	25-44			0.400	
	women	45-54			0.471	
	women	55-64			0.690	
	women	65-69			0.860	
	women	70-74			0.897	
	women	75-79			1.142	
	women	80-84			1.240	
	women	85+			1.381	
Female	-	-		-0.112	-0.261	
secondary sector employment	men	-	-0.267			
	women	-	-0.142			
Education	-	-		-0.104		
	men	-				
	women	-				
	men	25-44	-0.245			
	men	45-54	-0.253			
	men	55-64	-0.227			
	men	65-69	-0.216			
	men	70-74	-0.166			
	men	75-79	-0.108			
	men	80-84	-0.086			
	men	85+	-0.053			
	women	25-44	-0.152			
	women	45-54	-0.160			
	women	55-64	-0.134			
	women	65-69	-0.123			
	women	70-74	-0.073			
	women	75-79	-0.015			
	women	80-84	0.007			
	women	85+	0.040			
Divorce	-	-	0.273	0.038		
	-	25-44				
	-	45-54				
	-	55-64				
	-	65-69				
	-	70-74				
	-	75-79				
	-	80-84				
	-	85+				
	men	25-44				
	men	45-54				
	men	55-64				
	men	65-69				
	men	70-74				
	men	75-79				
	men	80-84				
	men	85+				
	women	25-44				
	women	45-54				
	women	55-64				
	women	65-69				
	women	70-74				
	women	75-79				
	women	80-84				
	women	85+				

Appendix Continued

Covariate	Sex	Age	Cancer	Circ.	Resp.	Digestive
Gypsies	-	-	0.023		0.079	
	-	25-44				-0.053
	-	45-54				-0.053
	-	55-64				0.050
	-	65-69				0.064
	-	70-74				0.086
	-	75-79				0.036
	-	80-84				0.142
	-	85+				0.112
Lung cancer	-	-				
	-	25-44		0.203		
	-	45-54		0.201		
	-	55-64		0.186		
	-	65-69		0.143		
	-	70-74		0.099		
	-	75-79		0.061		
	-	80-84		0.016		
	-	85+		-0.086		
	men	25-44			-0.193	
	men	45-54			0.238	
	men	55-64			0.518	
	men	65-69			0.492	
	men	70-74			0.514	
	men	75-79			0.424	
	men	80-84			0.387	
	men	85+			-0.073	
	women	25-44			-0.032	
	women	45-54			0.040	
	women	55-64			0.086	
	women	65-69			0.082	
	women	70-74			0.086	
	women	75-79			0.071	
	women	80-84			0.065	
	women	85+			-0.012	
SO2 pollution	-	-		-0.004	-0.017	-0.030
Urbanisation	-	-	0.032			-0.068
	men	25-44			0.290	
	men	45-54			0.277	
	men	55-64			0.410	
	men	65-69			0.421	
	men	70-74			0.565	
	men	75-79			0.586	
	men	80-84			0.589	
	men	85+			0.612	
	women	25-44			0.190	
	women	45-54			0.177	
	women	55-64			0.311	
	women	65-69			0.322	
	women	70-74			0.466	
	women	75-79			0.487	
	women	80-84			0.490	
	women	85+			0.513	
Physicians	-	-		-0.017	-0.105	
Unemployment (CR)	-	-			-0.120	-0.063
Wages (CR)	-	25-44	-0.171			0.095
		45-54	-0.078			0.079
		55-64	-0.071			-0.081
		65-69	-0.059			-0.205
		70-74	0.002			-0.192
		75-79	-0.046			-0.232
		80-84	0.021			-0.169
		85+	0.047			-0.105
GDP (CR)	-	-				
	-	25-44		-0.003		-0.015
	-	45-54		-0.001		-0.010
	-	55-64		-0.001		-0.009
	-	65-69		-0.001		-0.002
	-	70-74		0.005		0.000
	-	75-79		-0.001		0.002
	-	80-84		0.002		0.007
	-	85+		0.000		0.000

Table 1 Covariates used in the study: Association with mortality according to literature, definitions, sources and summary statistics before transformation

Covariate	Abbrev.	Association with mortality	Operationalisation and measurement	Period	Ave	Min.	Max	S.D.
Unemployment	UNEMP	+ [a]	% of working population registered as unemployed	1987-97 [1]	2.5	0.0	12.4	2.4
Wages	WAG	– [b]	Log of average monthly wages in Czech Crowns of registered workers, excl. the armed forces, small companies (< 25 workers), private entrepreneurs not registered in the Company Register *	1987-97 [1, 2]	5353	2652	14073	2515
Employment sector			% of employees in the national economy employed in the:					
– Primary sector	PRI	+ and – [c]	Agriculture, hunting and fishing	ave. 1993-96 [1]	8.7	0.3	19.8	4.6
– Secondary sector	SECM	+ [c,d,e]	Construction and industry (men)	ave. 1993-96 [1]	53.3	37.1	70.3	6.5
– Secondary sector	SECF	+ [c]	Construction and industry (women)	ave. 1993-96 [1]	33.9	15.7	51.0	6.1
Education	EDU	– [d]	Reciprocal of % of resident population with tertiary education	1991 [1]	5.6	3.5	16.3	2.1
Gross Domestic Product	GDP	– [f]	% of real change in GDP from previous year **	1987-97 [3]	0.5	-11.6	5.9	4.3
Urbanisation	URB	+ and – [g]	% of the population living in towns > 10,000 inhabitants.	1994 [1]	43.5	0.0	100.0	22.5
Gypsies	GYPS	+ [h]	Square root of the % of population registered as Gypsy	1988 [4]	13.9	0.4	51.4	11.7
Divorce	DIVOR	+ [i]	Number of divorces per 100 marriages	1987-97 [1]	43.9	15.3	114.3	13.8
Smoking	LUNG	+ [j]	SDR of lung cancer per pop'n of 100,000 (European Standard)	ave. 1987-97 [5]				
			– Men		176.4	120.9	261.0	35.0
			– Women		25.8	9.8	56.5	9.2
			– Total		95.2	65.5	150.0	19.5
SO_2 Pollution	SO2	+ [k]	Log of emission of SO_2 pollutants (tonnes/km^2/year) ***	1987-97 [1, 6]	25.1	0.04	540.1	64.3
Physicians	DOC	– [l]	Number of hospital in-patient physicians per 10,000 inhabitants	ave. 1996-97 [1]	7.0	0.0	18.0	2.9
Hospital beds	BED	– [l]	Number of hospital beds per 10,000 inhabitants	ave. 1996-97 [1]	63.8	0.0	123.8	20.4

Association between covariate and mortality: Signs generally relate to total mortality, as anomalies exist for specific causes (see also main text): [a] Olser (2003); [b] Kardaun and Glerum (1995); [c] Valkonen et al. (1993) – farmers tend to have mortality levels between manual and non-manual workers; [d] Davey Smth et al. (1998); [e] Kunst et al. (1998); [f] Mackenbach and Looman (1994) – the association was considered "weak" as urbanisation and industrialisation obscured the mortality lowering effect; [g] Mackenbach et al. (1991) and Van Poppel (1991) – e.g. as higher pollution levels increases but better access to health care decreases mortality levels in urban areas; [h] Burcin and Kučera (2000); [i] Blažek and Dzúrová (2000); [j] Peto et al. (1992); [k] Bobak and Feachem (1995); [l] See Nolte and McKee (2004) for an overview of mortality studies that used health care indicators.
Notes: * Years 1987-88, linear estimation using 1985 and 1989 figures. Data prior to 1992 exclude employees of cooperative farms; ** Until 1989 computed from the measure 'National income created', thereafter from 'GDP in constant prices'; *** Figures for 1987 and 1988 were interpolated using average levels for 1981-85 (taken as 1983) and 1989.
Data sources: [1] Czech Statistical Office (CSO) (1992-98); [2] CSO (1999a); [3] CSO (1999b); [4] Srb (1990); [5] See main text; [6] Rychtaříková and Dzúrová (1992).

Table 2 Causes of death used in this study, their ICD-codes and their prevalence in the CR (1987-1997)

Cause of death	ICD-10	ICD-9	Prevalence
Cancer	C00-D48	140-239	23%
Diseases of the circulatory system	I00-I99	390-459	56%
Diseases of the respiratory system	J00-J99	460-519	7%
Diseases of the digestive system	K00-K93	520-579	4%
External causes of injury and poisoning	V01-Y89	E800-E999	4%
Remaining causes of death	-	-	6%

Table 3 Best models and the percentage of decline in deviance (D) over the AS benchmark model

Cause of death	Best model	% decline in D
Cancer	AS+AR+AT+ SR	17%
Circulatory system diseases	AS+AR+AT+RT	58%
Respiratory system diseases	AS+AR+ RT+SR+TS	35%
Digestive system diseases	AS+AR+AT	15%
External causes of death	AS+AR+AT+RT+SR	18%
Remaining causes of death	AS+AR+ RT	23%

Table 4 Percentage improvements in deviance of models relative to benchmark model AS: (i) Best models, (ii) Tested models based on data availability, (iii) Models where the spatial and temporal effects are replaced by covariates, and (iv) as in (iii) but includes interactions with age and/or sex

	1 (i) Best model	2 % explained	3 (ii) Tested model	4 % explained (as % of (2))	5 Significant covariates[#]	6 % explained (as % of (4))	7 Signif cov, incl inter. w. age/sex	8 % explained (as % of (2))
Cancer	AR+AT+SR	25.31	R+T+SR	18.26 (72.13)	**R:** PRIM, SECM, SECF EDU, GYP, URB, DIVOR **T:** WAGT	10.53 (57.66)	**R:** URB, GYPS, DIVOR **AR**: PRIM, EDU **AT**: WAGT **SR**: PRIM, EDU, SECM, SECF	15.05 (59.44)
Circulatory system diseases	AR+AT+RT	49.79	R+T+RT	40.31 (80.95)	**R:** SECM, SECF, EDU, PRIM, LUNG, DOC **T:** GDP **RT:** WAG, SO2, DIVOR	28.51 (70.73)	**R**: PRIM, SECF, EDU, DOC **AR**: LUNG[1], SECM **AT**: GDP **RT**: SO2, DIVOR ART: WAG	32.78 (65.84*)
Respiratory system diseases	AR+RT+SR+TS	45.16	R+T+RT+SR	38.77 (85.83)	**R:** PRIM, SECM, SECF, URB, GYPS, DOC **T:** UNEMPT **RT:** WAG, SO2 **SR:** LUNG	13.04 (33.64)	**R**: SECF, GYPS, DOC **T**: UNEMPT **AR**: PRIM, SECM, URB **RT**: SO2 **SR**: SECM, URB SRT: WAG ASR: LUNG **TS**: -	15.32 (33.92*)
Digestive system diseases	AR+AT	17.06	R+T	8.91 (52.25)	**R:** UNEMP, WAG, PRIM, URB, GYPS, SO2 **T:** UNEMPT, WAGT, GDP	5.47 (61.41)	**R**: WAG, URB, SO2 **T**: UNEMPT **AR**: UNEMP, PRIM, GYPS **AT**: WAGT, GDP	7.64 (44.81)
External causes	AR+AT+RT+SR	29.27						
Remaining causes	AR+RT	34.46						

* Indicates that figure includes one or more interactions (i.e. those not printed in bold) that were not specified in column (1)

[#] For circulatory system diseases and remaining causes the lung cancer proxy is not sex-specific, as in the best model it showed no interaction with sex. WAGT and UNEMPT refer to national level wages and unemployment level.

Figure 1. The 8 kraj (region) and 76 okresy (district) boundaries of the Czech Republic as they were in 1987.

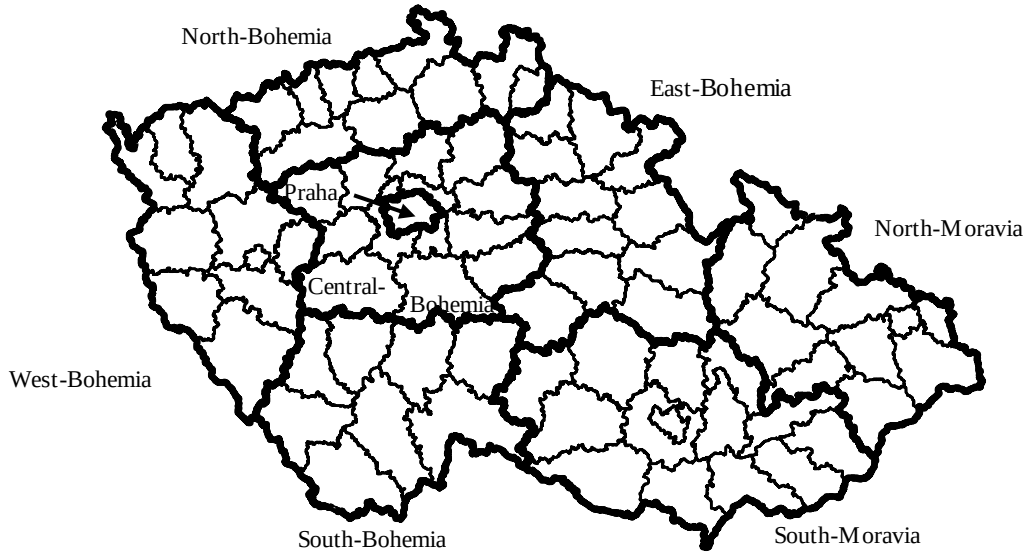


Figure 2 Age-specific mortality rates for the six categories of causes of death, CR, 1987-97

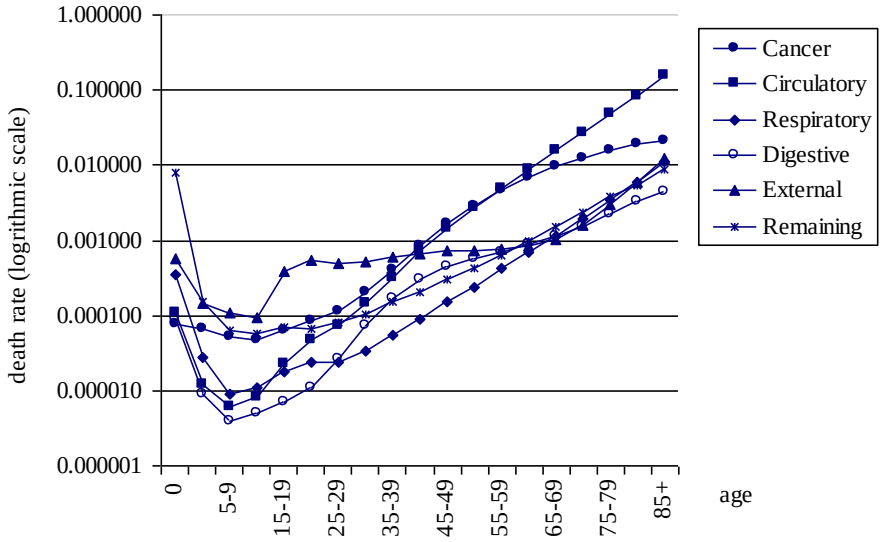


Figure 3 Male:female Relative Mortality Figures for six major categories of causes of death, CR, 1987-97

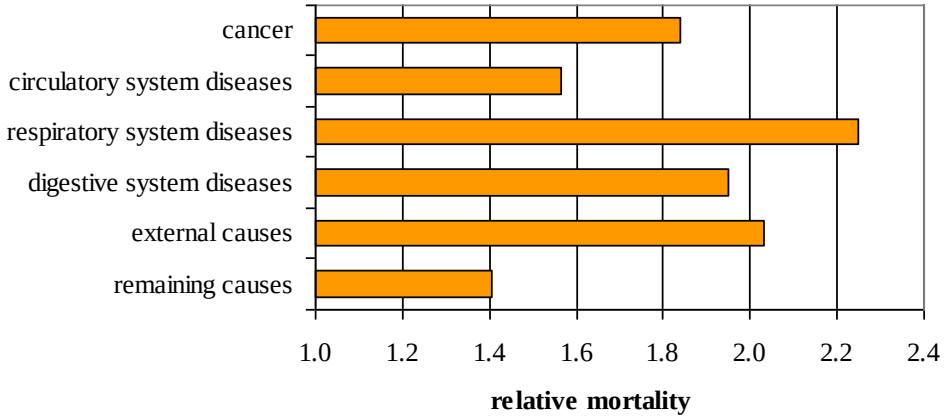


Figure 4 District effects in terms of Relative Mortality Figures for the six categories of causes of death, 1987-97, CR.

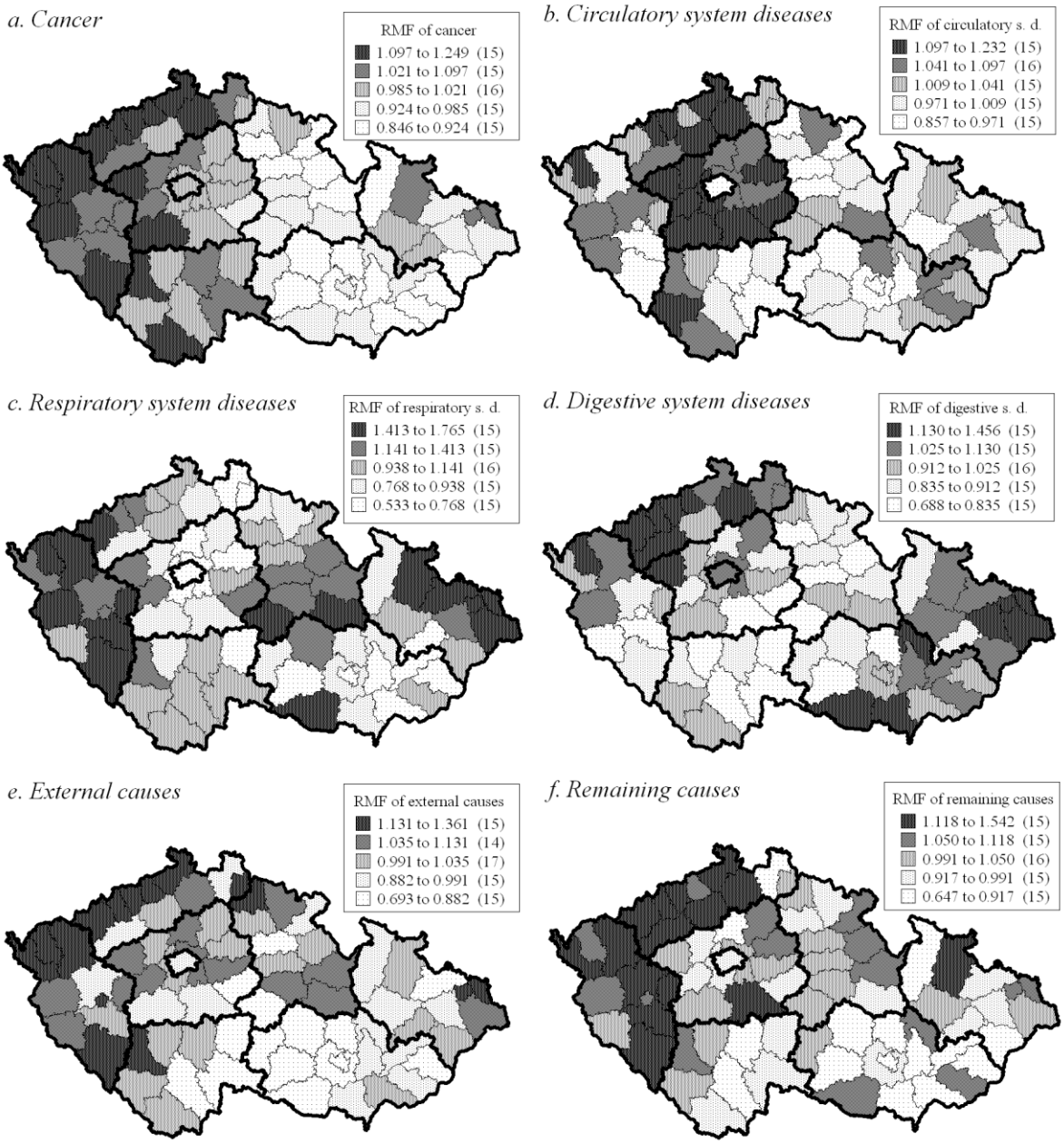


Figure 5 Annual Relative Mortality Figures for the six major categories of causes of death, CR, 1987-97.

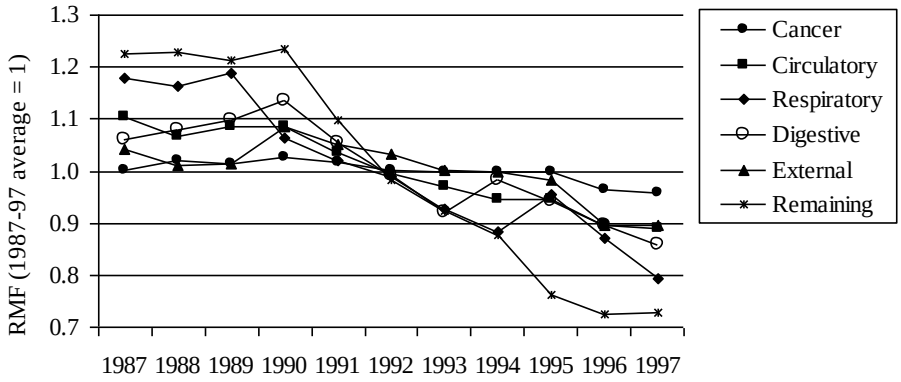


Figure 6 Relative Mortality Figures for Age*District interaction, CR, 1987-97.

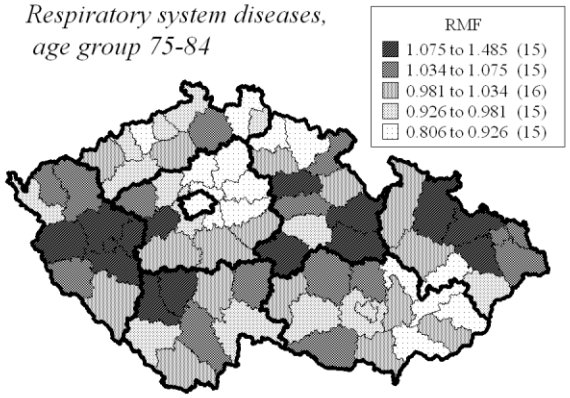
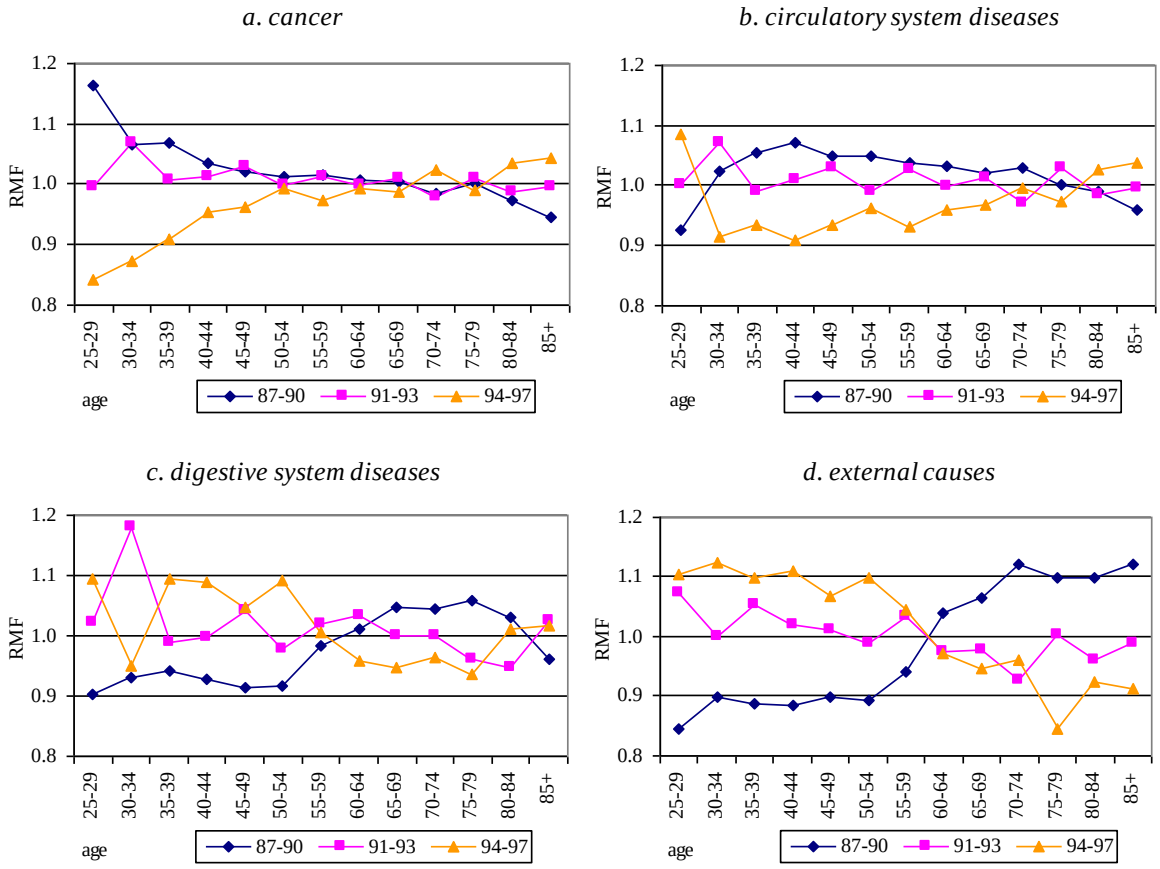
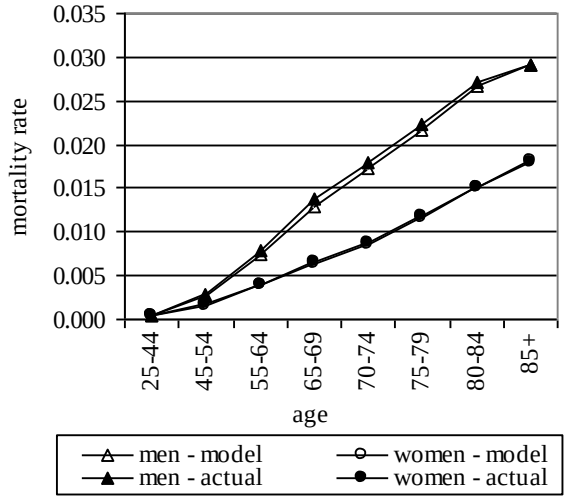


Figure 7 Relative Mortality Figures for Age*Time interaction for four causes of death, 1987-97, CR

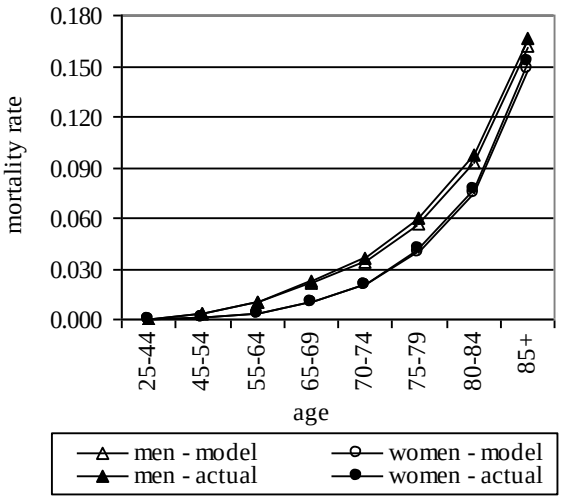


Note: For presentation purposes, the results are presented for the three periods 1987-90, 1991-93 and 1994-97 instead of single years.

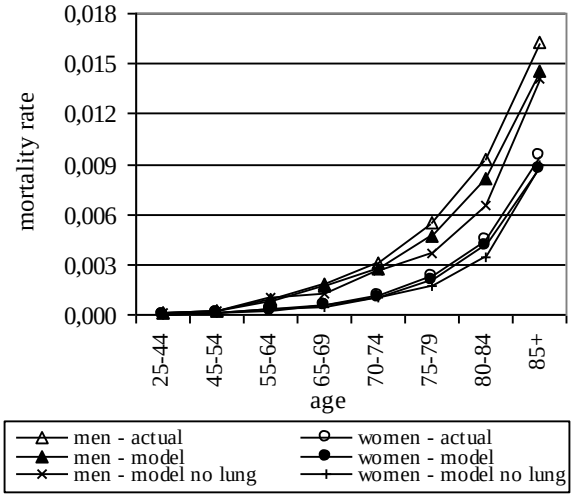
Figure 8 Age and sex-specific mortality, CR: actual and modelled



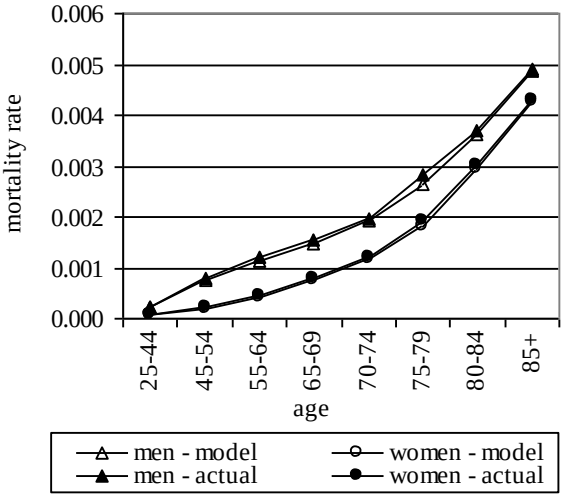
a. Cancer



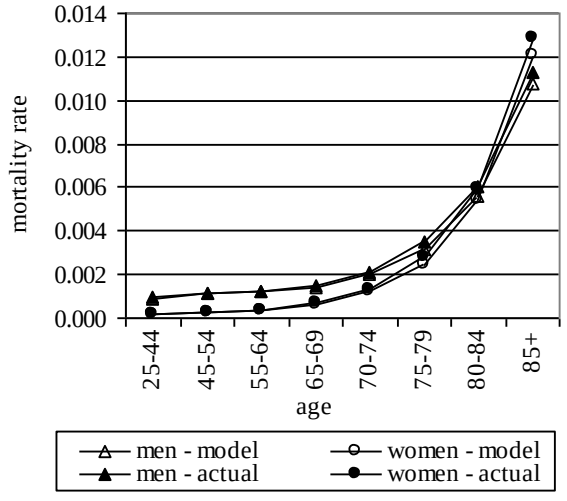
b. Circulatory system diseases



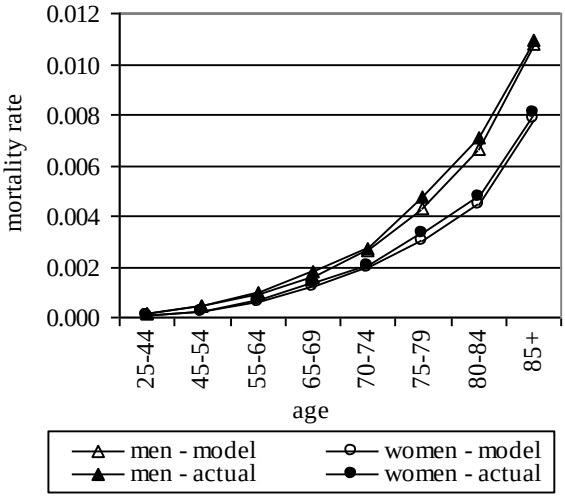
c. Respiratory system diseases



d. Digestive system diseases



e. External causes of death



f. Remaining causes of death