GENUS, **LVII** (n. 3-4) 2001, 33-62

VLADIMIR M. SHKOLNIKOV – TAPANI VALKONEN
ALEXANDER BEGUN – EVGUENI M. ANDREEV

Measuring inter-group inequalities in length of life

1. INTRODUCTION

There is a long tradition of the use of the length of life for expressing social inequalities in mortality. A number of studies on socio-economic gradients in length of life were completed beginning from the 19th century. Aaron Antonovsky (1967) reviewed many of these early studies in a classic paper. However, in the second half of 20th century this tradition weakened. Antonovsky notes: "very few later investigators have dealt with class differences in life expectancy, preferring to concentrate on differences in mortality rates" (Antonovsky, 1967, p. 35).

Certainly, simplicity of mortality rates and of their linear aggregates (e.g. standardized mortality rates and ratios) as well as their dominance in epidemiology (relative risks and odds ratios) are important advantages. However, these sorts of indicators reflect only frequency (or intensity) of death and do not tell directly about its *prematureness*.

Indeed, in the long run all people will die. Only age at death makes a principal difference. It is intuitively clear that the most favorable mortality pattern can be defined as one corresponding to the highest expected average length of life while the worst mortality pattern should correspond to the lowest expected average length of life. That is why life expectancy is so easily understandable and so widely used as a summary measure of mortality. There are, however, some technical difficulties in operating with the life expectancy, as its link to mortality rates is a complex one.

The present article describes measurement procedures for inter-group inequalities in life expectancy. They are applicable both for period and cohort data. The former case is more difficult than the latter one since the changes in the size of contemporary population with age substantially differ from what is predicted by the life table survival.

The first section presents an agenda for measurement and analysis of inter-group inequalities in life expectancy. The existing formulae for the decomposition of a difference between two life expectancies are slightly modified to be applicable to a difference between temporary life expectancies. In respect to decomposition by cause-of-death, a problematic case is considered, in which some of the age-specific death rates are equal in two populations to be compared. Then a method is suggested to split the period life table cohort into fractions corresponding to population groups. It gives a way to develop simple indices of inequality in age at death, similar to those based on mortality rates.

The second section illustrates an application and graphical presentation of these procedures to mortality data from the USA, Finland, and Russia.

2. MEASUREMENT PROCEDURES

Appropriate measurement and analytical procedures should provide a quantitative basis for:

- measuring the magnitude of inter-group differences in length of life;
- decomposing these differences by age and causes of death;
- estimating the overall degree of inequality in length of life.

Let us consider a population which consists of n groups classified according to income, occupation, ethnicity, education or other categories. Age-specific or age- and cause-specific mortality rates are known for the overall population and for each group.

2.1 Life expectancy

Life expectancy at age x depends on mortality rates at age x and older ages, while the temporary life expectancy (Arriaga, 1984) or life expectancy for the range of ages from x to X depends on mortality rates within this age interval.

$$e_x = \frac{T_x}{l_x} \qquad [1a]$$

$$e_{x,X} = \frac{T_x - T_X}{l_x} \qquad [1b]$$

where l_x is the survival rate from birth to exact age x, T_x and T_X are the numbers of person-years lived after ages x and X respectively by the survivors to these ages.

The formulae [1a] and [1b] are equally applicable to period and cohort life tables and can be used to estimate lengths of life in each population group. Obviously, these estimates permit one to learn about a magnitude of inter-group differences.

In a cohort life table (unlike a period life table) the life expectancy reflects a real mortality experience and is equal to the average number of years lived by the cohort during the follow-up period. The individuals withdrawn, lost from or added to observation can be taken into account by means of simple modification of the denominator in the formula of the probability of death (see Chiang, 1984, Kahn *et al.*, 1989, Lee, 1992).

2.2 Decomposing differences between life expectancies by age and cause of death

The method for decomposing of differences between two life expectancies by age was developed independently in the 1980s by three researchers from Russia, the USA, and France (Andreev, 1982, Arriaga, 1984, Pressat, 1985)[1]. In fact, the final formula by Andreev is exactly equivalent to that by Pressat. The final formula by Arriaga is written in a somewhat different form, but can be easily transformed into the same form (Appendix 1). In addition, the method by Andreev and Pressat is symmetrical in terms of populations to be compared. This means that the results of decomposition do not depend on whether or not changes are made in population 1 to bring it to the mortality level of population 2 or vice versa. Andreev and Pressat did this simply by averaging the components of the difference $e_0^1 - e_0^2$ with the respective components of the difference $e_0^2 - e_0^1$. Arriaga (1984) did not complete this step.

Overall, the method allows splitting a difference between two life expectancies into contributions of elementary differences in age-specific mortality rates as follows:

$$e_x^2 - e_x^1 = \sum_{y=x}^{W} {}_n\varepsilon_y$$

[1] The UN Population Division used a simplified version of the same formula (UN, 1988). Another method for decomposition of a difference between the life expectancies was developed by J.H.Pollard for a continuous case (Pollard, 1982). This method is based on a somewhat different (approximate) formula and returns slightly different numeric results.

$$
{}_n\varepsilon_y = \frac{1}{2l_x^2}\left[l_y^2\left(e_y^2 - e_y^1\right) - l_{y+n}^2\left(e_{y+n}^2 - e_{y+n}^1\right)\right] - \frac{1}{2l_x^1}\left[l_y^1\left(e_y^1 - e_y^2\right) - l_{y+n}^1\left(e_{y+n}^1 - e_{y+n}^2\right)\right] \quad [2]
$$

where 1,2 – two population groups to be compared, x,y – exact ages, n – length of elementary age interval, W – the oldest age group (usually 85+, 95+ or 100+), ${}_n\varepsilon_y$ – contribution to the overall difference between the life expectancies produced by the difference in mortality in the age group y, $y+n$. If $y = W$ then $l_{y+n}^i = 0$, $e_{y+n}^i = 0$ in [2].

If one has to decompose a difference between two temporary life expectancies $e_{x,X}^1 - e_{x,X}^2$ then life expectancies e_y in the equation [2] should be replaced by temporary life expectancies $e_{y,X}$.

If cause-specific data are available a further decomposition according to causes of death (j) can be performed (Andreev, 1982, Arriaga, 1989):

$$
{}_n\varepsilon_y = \sum_j {}_n\varepsilon_{y,j},
\qquad
{}_n\varepsilon_{y,j} = \frac{{}_nM_{y,j}^1 - {}_nM_{y,j}^2}{{}_nM_y^1 - {}_nM_y^2} \cdot {}_n\varepsilon_y, \quad [3a]
$$

where ${}_nM_{y,j}^1$, ${}_nM_{y,j}^2$ are central death rates in population groups 1 and 2 for age group $y,y+n$ and the cause of death j, ${}_nM_y^1, {}_nM_y^2$ are central death rates for the same age in population groups 1 and 2 for all causes of death combined. This formula does not work if in one of the age groups y ${}_nM_y^1 = {}_nM_y^2$ and for a cause of death j ${}_nM_{y,j}^1 \neq {}_nM_{y,j}^2$. Usually this happens when numbers of deaths in population groups 1 and 2 are small. In this case a more general formula (Appendix 2) can be applied:

$$
{}_n\varepsilon_{y,j} = \frac{1}{2} \cdot \left({}_nM_{y,j}^1 - {}_nM_{y,j}^2\right) \cdot \left(\frac{1}{l_x^2} \cdot \int_y^{y+n} l_t^2 e_t^1 dt + \frac{1}{l_x^1} \cdot \int_y^{y+n} l_t^1 e_t^2 dt\right). \quad [3b]
$$

The integrals in [3b] should be computed numerically. Therefore, for a given age group $y,y+n$ the sum of the cause-specific components ${}_n\varepsilon_{y,j}$ can be computed approximately. Fortunately, the formula [3b] is needed in relatively rare cases.

The expressions [1]-[3] permit one to examine the reasons for observed differences between group-specific life expectancies in terms of ages and causes of death.

2.3 *The weights of population groups in a life table cohort*

If a closed birth cohort consists of several sub-groups then the dynamics of its size is fully determined by mortality schedules in the sub-groups and their proportions in the overall birth cohort at the beginning of the follow up. In this case the life expectancy for the whole birth cohort is the sum of the group-specific life expectancies weighted by population-weights of the sub-groups. However, a period life table is based on a hypothetical ("synthetic") cohort. This makes a link between the longevity of the overall population and group-specific lengths of life more complicated.

Indeed, changes in the composition of the overall population across ages can be induced by factors other than mortality within population groups. For example, if the weight of an ethnic group declines with age it can be explained not only by relatively high mortality in this group, but also by differential trends in ethnic migration. Age-differentials in population composition by marital status depend on trends in marriages and divorces[2]; changes in population composition by social class depend on trends in vertical social mobility, *etc.*

For the forces of mortality the following simple equation determines the relationship between mortality in the overall population μ_x and mortality in its N sub-groups μ_x^i

$$\mu_x = \sum_{i=1}^{N} \mu_x^i p_x^i,$$

where p_x^i are the proportions of the groups i in real population at exact age

[2] If flow intensities forming the population sub-groups (for example, marriages, divorces, re-marriages, deaths by marital status) are registered then it is possible to construct a multi-status life table (Willekens *et al.*, 1983).

x. The survival function l_x is defined as $\exp\left(-\int_0^x \mu_t dt\right)$. Consequently, the link between life expectancy at age x for the whole population and group-specific mortality schedules can be written as:

$$e_x = \frac{1}{l_x}\int_x^\infty l_t dt = \frac{1}{l_x}\int_x^\infty \left[\exp\left(-\int_0^t \sum_i p_\tau^i \mu_\tau^i d\tau\right)\right] dt.$$

An analytical relationship between this expression and the life expectancies in specific population groups

$$e_x^i = \frac{1}{l_x^i}\int_x^\infty \left[\exp\left(-\int_0^t \mu_\tau^i d\tau\right)\right] dt$$

would be a complex one.

However, we can try to make a simpler linear decomposition for the overall life expectancy. This could be achieved by a division of the overall life table cohort into fractions corresponding to specific population groups. Desirable properties of these fractions are the following: at every age exact age x the sum of the fractions should be equal to the number of survivors in the overall cohort, and the sum of person-years lived by all fractions after age x should be equal to the total number of person-years lived by the whole cohort.

$$\begin{aligned} l_x &= \sum_{i=1}^N l_x^i, \\ e_x l_x &= \sum_{i=1}^N l_x^i e_x^i. \end{aligned} \qquad [4]$$

Variables e_x^i are known from the group-specific life tables. A task is to estimate the values of l_x^i (or the weights $\theta_x^i = l_x^i / l_x$). Obviously, if there are only two population groups (N=2) then the system of equations [4] has only one solution:

$$l_x^1 = \frac{l_x\left(e_x - e_x^2\right)}{e_x^1 - e_x^2}, \qquad [5]$$
$$l_x^2 = l_x - l_x^1.$$

There are, however, multiple solutions if $N>2$. This means, some additional constraints should be formulated. For a closed cohort with sub-groups (i) followed up for mortality starting from age x it is clear that: $\theta_x^i = P_x^i / P_x$, where P_x, P_x^i are the sizes of the overall cohort and its sub-groups (i) at the beginning of the follow-up. For period data a reasonable approach would be to choose the weights θ_x^i characterized by a minimum distance from average proportions of groups i in the overall population at age x and older ages P_{x+}^i / P_{x+}, with

$$P_{x+}^i = \sum_{y=x}^{W} P_y^i, P_{x+} = \sum_{y=x}^{W} P_y.$$

Appendix 3 shows that it can be formulated as a problem of minimization with constraints, which leads to the system of linear equations $\mathbf{A}\cdot\mathbf{z} = \mathbf{b}$. In this expression matrix $\mathbf{A}$ has the dimension $(N+2)\text{x}(N+2)$ and vector $\mathbf{b}$ has the dimension $(N+2)\text{x}1$:

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 0 & .. & .. & .. & 0 & 1 & e_x^1 \\ 0 & 2 & 0 & .. & .. & .. & 0 & 1 & e_x^2 \\ 0 & 0 & 2 & .. & .. & .. & 0 & 1 & e_x^3 \\ .. & .. & .. & .. & .. & .. & .. & .. & .. \\ .. & .. & .. & .. & .. & 2 & 0 & .. & .. \\ 0 & 0 & 0 & .. & .. & 0 & 2 & 1 & e_x^N \\ 1 & 1 & 1 & .. & .. & .. & 1 & 0 & 0 \\ e_x^1 & e_x^2 & e_x^3 & .. & .. & .. & e_x^N & 0 & 0 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 2(P_{x+}^1 / P_{x+}) \\ 2(P_{x+}^2 / P_{x+}) \\ 2(P_{x+}^3 / P_{x+}) \\ .. \\ .. \\ 2(P_{x+}^N / P_{x+}) \\ 1 \\ e_x \end{bmatrix}, \qquad [6]$$

Vector $\mathbf{z}$ has dimension $(N+2)\text{x}1$ and its first N elements are the unknown values θ_x^i. This vector can be calculated from:

$$\mathbf{z} = \mathbf{A}^{-1}\cdot\mathbf{b}. \qquad [7]$$

The resulting weights θ_x^i permit to present the life expectancy of the overall population as a weighted average of the group-specific life expectancies.

The above method for developing the fraction weights provides estimates which take into account only group-specific proportions in the whole population aged x+, without taking account of group-proportions at every age group (x, x+1, x+2,..., W). A more sophisticated approach would be to develop such estimates of weights for the starting age group θ_x^i, which provide the best fit to the proportions of groups in the real population P_y^i / P_y for all ages $y \geq x$.

Suppose that we know for each population group the independent survival function $S_y^i, y \geq x$ (similar to l_y^i, but unlinked to each other in respect to i). If the change in population size with age were due to mortality then for each age y the population weights of groups would be

$$\xi_y^i = \theta_x^i S_y^i / \sum_i \theta_x^i S_y^i, y = x, x+1, \ldots, W.$$

The probability L that the population counts by age are exactly equal to real populations is $\prod_{y=x}^{W} (P_y!) \prod_{i=1}^{n} \frac{(\xi_y^i)^{P_y^i}}{P_y^i!}$, where $P_y^i!$ is the number of permutations in a set of P_y^i elements. In accordance with the principle of maximum likelihood we must maximize the log(L) to estimate the values of θ_x^i. Since P_y^i and P_y do not depend on θ_x^i the log(L) can be transformed in a simpler form

$$\log(L) = \sum_{y=x}^{W} \sum_{i=1}^{N} P_y^i \log(\xi_y^i) \rightarrow \max \qquad [8]$$

For empirical calculations we have applied the MATLAB Optimization Toolbox (The MathWorks, 1999) in order to find a solution for the optimization problem [8] with constraints [4]. In most cases it appears that the estimates of θ_x^i produced by this method, do not seriously differ from those returned by the easier procedure [6]-[7].

2.4 Summary indices of inequality

Summary indices of inequality measure the general level of inter-group inequality in a given population. A number of such indices have been constructed with respect to mortality or incidence rates (Kunst *et al.*, 1994). Similar indices can be determined for life expectancies.

For example, the *PAR* (population attributable risk) index can be transformed into *PALL* (population attributable life loss) index. In relative terms it is a proportion of increase in overall life expectancy that would occur if all groups had the life expectancy of the best group. In absolute terms it is the increase in the overall life expectancy that would be achieved if all groups had the life expectancy of the best group.

$$PALL = \sum_i \frac{(e_x^{highest} - e_x^i)\theta_x^i}{e_x} \text{ (in relative terms)} \qquad [9a]$$

$$PALL_{abs} = PALL \cdot e_x = e_x^{highest} - e_x \text{ (in absolute terms).} \qquad [9b]$$

Index of dissimilarity (*ID*) is also a well-known summary measure of inequality based on mortality rates. A similar index can be built for life expectancies. It can be named index of dissimilarity in length of life (*IDLL*):

$$IDLL = \frac{\sum_i | e_x - e_x^i | \theta_x^i}{e_x} \text{ (in relative terms)} \qquad [10a]$$

$$IDLL_{abs} = IDLL \cdot e_x \text{ (in absolute terms).} \qquad [10b]$$

In relative terms it shows the proportion of person-years of the length of life that should be redistributed among the groups if all groups would achieve the average level of life expectancy. In absolute terms it shows how many years of length of life should be redistributed in the overall population to reach the situation of complete equality.

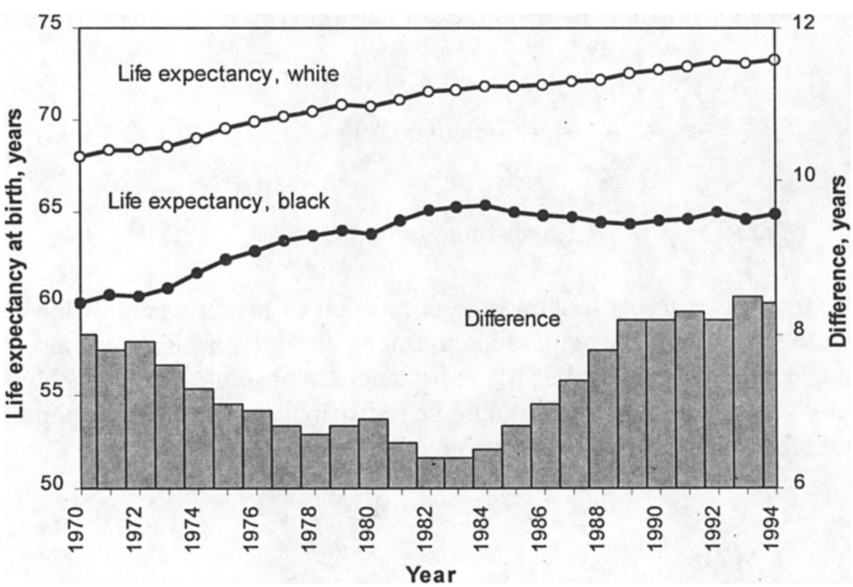
3. EMPIRICAL EXAMPLES

Empirical examples show how the measurement agenda of the previous section can be utilized practically. In three examples we analyze inequalities in male length of life: by race in the USA, by socio-occupational class in Finland, and by occupational category in Russia. The first example is described in greater detail below. The first and the second examples are based on national-level period data, while the third example employs logngitudinal data for a relatively small epidemiological cohort.

3.1 Differences between racial groups in the USA

Substantial inequalities in mortality of the US population constitute a serious public health issue. Special attention should be given to a low life expectancy of African Americans. Figure 1 suggests that a declining trend in the difference in the life expectancy at birth between white and black Americans was reversed in the early 1980s (for details see Kochanek *et al.*, 1994). As a result of this unprecedented change, the difference has increased from about 6 years in 1984 to about 8 years in 1994.

Figure 1 – *Trends in life expectancy at birth for white and black men in the USA in 1970-1994*



Routine health statistics of the USA permit a calculation of age-specific (5-year age groups) and cause-specific mortality rates for three major ethnic groups: white, black, and "other"[3]. According to the census of 1990 the percentages of these groups in the overall male population of the USA were 0.8413, 0.1190, and 0.0387, respectively.

The race-specific life tables return the following values of male life expectancy at birth for the year 1990: 71.9 years for all men, 72.7 years for white men, 64.6 for black men, and 78.1 years for "other" men[4]. The three population groups represent an impressive range of mortality levels. The life expectancy of "other" men is slightly higher than the life expectancy in Japan or Iceland (the world's highest levels for national populations)[5], the life expectancy of white men is close to the Western European average, and the life expectancy of black men is close to low levels characteristic of the countries of Eastern Europe.

In 1990 the difference in male life expectancy at birth between the two biggest racial groups (white and black Americans) was 8.2 years. This gap can be decomposed by age and leading classes of causes of death according to the formulae [2] and [3a].

Although all ages except the oldest age group contribute to the total difference between white and black men, the biggest contributions are made by differences in mortality rates at middle ages (between 40 and 65) and in infancy (Table 1, Figure 2). Cardiovascular diseases, external causes of death (accidents and violence), and cancers are the greatest contributors to the overall difference among major causes of death. Excess cardiovascular mortality of black males has a maximum weight at ages between 55 and 65, while accidents and violence are more important among young adults aged 15 to 35. A significant role of infectious diseases at ages between 25 to 45 is remarkable (Figure 2). The latter reflects a disproportional concentration of HIV deaths among young African Americans.

Formula [6] with N=3 is used to estimate the weights of racial groups in the overall life table cohort. According to [7] Matrix **A** and vector **b** are as follows:

[3] A majority of people in the residual category are Asian Americans.

[4] Our estimates are very close to those produced by the National Center for Health Statistics: 71.8 for all men, 72.7 for white men, 64.5 for black men (NCHS, 1994).

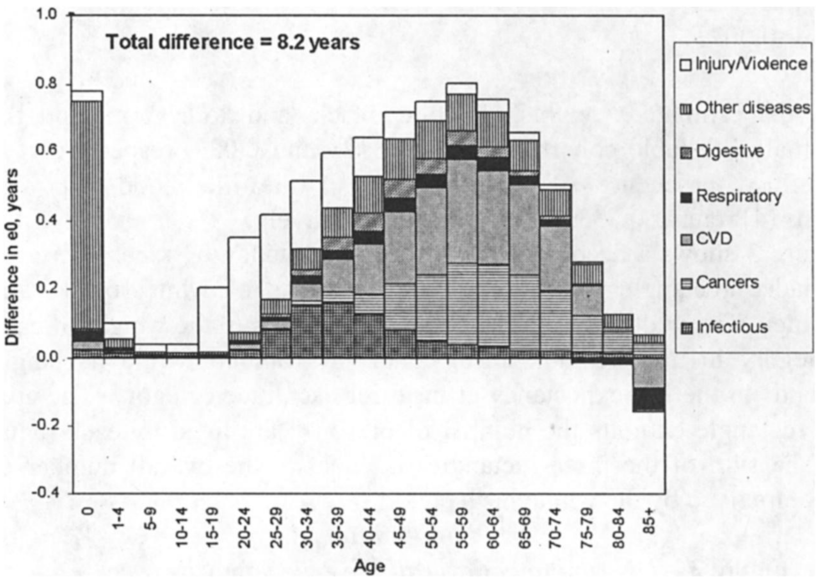
[5] The quality of the data for the residual small group is questionable, but it would not significantly affect our further results.

Table 1 – *Decomposition of the difference in life expectancy at birth between white and black men by age and cause of death*

Age	Infectious diseases	Cancers	CVD	Respiratory diseases	Digestive diseases	Other diseases	Accidents and violence	**All causes combined**
0	0.024	0.001	0.022	0.031	0.009	0.664	0.030	**0.780**
1-4	0.012	-0.001	0.008	0.008	0.001	0.028	0.045	**0.101**
5-9	0.005	0.000	0.000	0.003	0.001	0.007	0.025	**0.040**
10-14	0.001	-0.001	0.002	0.004	0.001	0.006	0.028	**0.041**
15-19	0.002	-0.002	0.005	0.003	0.001	0.007	0.186	**0.203**
20-24	0.022	0.004	0.015	0.007	0.002	0.023	0.279	**0.352**
25-29	0.080	0.004	0.023	0.013	0.008	0.041	0.250	**0.419**
30-34	0.154	0.009	0.050	0.022	0.018	0.062	0.202	**0.519**
35-39	0.162	0.029	0.094	0.028	0.038	0.087	0.160	**0.598**
40-44	0.127	0.058	0.148	0.037	0.051	0.107	0.116	**0.644**
45-49	0.083	0.131	0.212	0.041	0.056	0.115	0.082	**0.720**
50-54	0.053	0.187	0.259	0.039	0.047	0.108	0.056	**0.749**
55-59	0.042	0.235	0.302	0.040	0.043	0.109	0.033	**0.804**
60-64	0.031	0.242	0.275	0.039	0.030	0.104	0.036	**0.756**
65-69	0.022	0.218	0.268	0.026	0.014	0.089	0.020	**0.657**
70-74	0.022	0.176	0.189	0.016	0.012	0.078	0.012	**0.505**
75-79	0.016	0.109	0.086	-0.012	0.003	0.058	0.006	**0.266**
80-84	0.014	0.064	0.011	-0.013	0.001	0.040	-0.003	**0.113**
85+	0.010	0.036	-0.088	-0.057	-0.006	0.021	-0.005	**-0.089**
Total	**0.883**	**1.500**	**1.880**	**0.275**	**0.328**	**1.752**	**1.558**	8.176

Note: USA, 1990 (in years).

Figure 2 – *Age- and cause-specific contributions to the overall difference in the life expectancy at birth between white and black men in the USA in 1990*



$$
\mathbf{A} = \begin{bmatrix} 2 & 0 & 0 & 1 & 72.7483 \\ 0 & 2 & 0 & 1 & 64.5717 \\ 0 & 0 & 2 & 1 & 78.1041 \\ 1 & 1 & 1 & 0 & 0 \\ 72.7483 & 64.5717 & 78.1041 & 0 & 0 \end{bmatrix}
$$

$$
\mathbf{b} = \begin{bmatrix} 2*0.8413 \\ 2*0.1190 \\ 2*0.0387 \\ 1 \\ 71.8752 \end{bmatrix}
$$

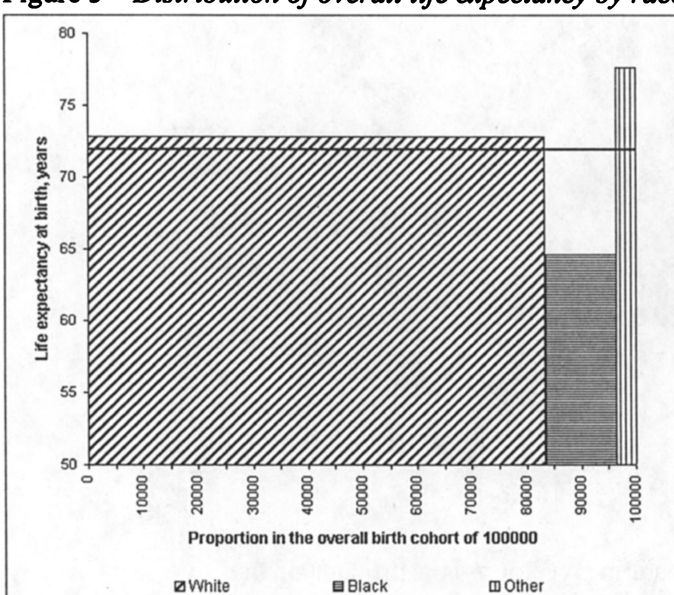
The resulting vector **z** is a product of the inverse matrix $\mathbf{A}^{-1}$ and the vector **b**:

$$\mathbf{z} = \begin{bmatrix} 0.8406 \\ 0.1276 \\ 0.0318 \\ -0.1652 \\ 0.0023 \end{bmatrix}$$

So, the estimated weights of white, black, and "other" men in the overall male life table cohort are 0.841, 0.128 and 0.032, respectively. A more refined procedure of maximization of the likelihood [8] with constrains [4] returns 0.833, 0.131, 0.036, respectively.

Figure 3 shows a decomposition of life expectancy by race. There are three shaded rectangles representing the fractions of white, black, and "other" men. The width of each rectangle corresponds to the weight of each race category in the life table birth cohort of 100,000, while its height corresponds to the life expectancy at birth for each race category. The area of each rectangle exhibits the number of person-years lived by each racial group. The sum of the three rectangles is equal to the overall number of person-years lived by the whole birth cohort.

Figure 3 – ***Distribution of overall life expectancy by race***



***Note*: USA, males, 1990.**

Figure 3 suggests that in 1990 high mortality among black men did not affect the level of the length of life of the total US population very much. Nevertheless, its negative impact was probably significant enough to contribute to a relatively unfavorable position of the USA in comparison to other Western countries according to the level of life expectancy at birth.

Table 2 – *Changes in life expectancy at birth associated with different proportional decreases in age-specific mortality rates*[6]

Percentage of proportional change in mortality rates		Life expectancy at birth			Difference
White (w)	Black (b)	White (w)	Black (b)	Total	(w)-(b)
0	0	72.8	64.6	71.9	8.2
-5	-5	73.5	65.4	72.6	8.1
-10	-10	74.2	66.4	73.3	7.8
-15	-15	75.0	67.4	74.1	7.6
-20	-20	75.8	68.5	75.0	7.3
-25	-25	76.7	69.6	75.9	7.1
-25	-15	76.7	67.4	75.7	9.3
-25	-35	76.7	73.3	76.2	3.4

Note: USA, males.

Table 2 shows expected changes in life expectancy at birth for white, black, and all American men induced by different degrees of proportional reduction in age-specific mortality rates. Proportional reductions are the same for all age groups; the mortality pattern of "other" men is fixed as is the proportion of the races in the overall population. If proportional decrease in mortality rates is the same for white and black men, the difference in life expectancy tends to decline. It is so because life expectancy, as a measure, is more sensitive to the same proportional changes in higher mortality rates among black men than to equivalent changes in lower mortality rates among white men.

Two bottom lines in Table 2 suggest that the overall life expectancy can mask a very low level of length of life even in big population minorities. In one case the inter-race difference in the life expectancy at birth is about 9 years, in another case it is about 3 years. The difference in the respective

6 The baseline level of mortality rates and life expectancy at birth corresponds to the year 1990.

overall life expectancies is only 0.5 year: 75.7 years versus 76.2 years.

Formulae [9] and [10] are used to compute summary indices of the inter-race inequality in life expectancy in the USA:

$$PALL = \frac{(78.1-72.7)*0.84+(78.1-64.6)*0.13}{71.9} = 0.087$$

$$PALL_{abs} = 0.087*71.9 = 6.2$$

Unusually high life expectancy in a small group of "other" men results in a high *PALL*. 6.2 years of length of life could be achieved by the overall population if white and black Americans experience mortality conditions characteristic of the "other" Americans. This account for 8.7% of the overall number of years lived by the overall birth cohort.

$$IDLL = \frac{|71.9-72.7|*0.84+|71.9-64.6|*0.13+|71.9-78.1|*0.03}{71.9} = 0.020$$

$$IDLL_{abs} = 0.020*71.9 = 1.4$$

It means that 1.4 years of length of life (or 2% of the overall life expectancy) should be redistributed to achieve the situation of complete equity between racial groups at the level of life expectancy of the overall male population of the USA.

3.2 Differences between socio-occupational classes in Finland

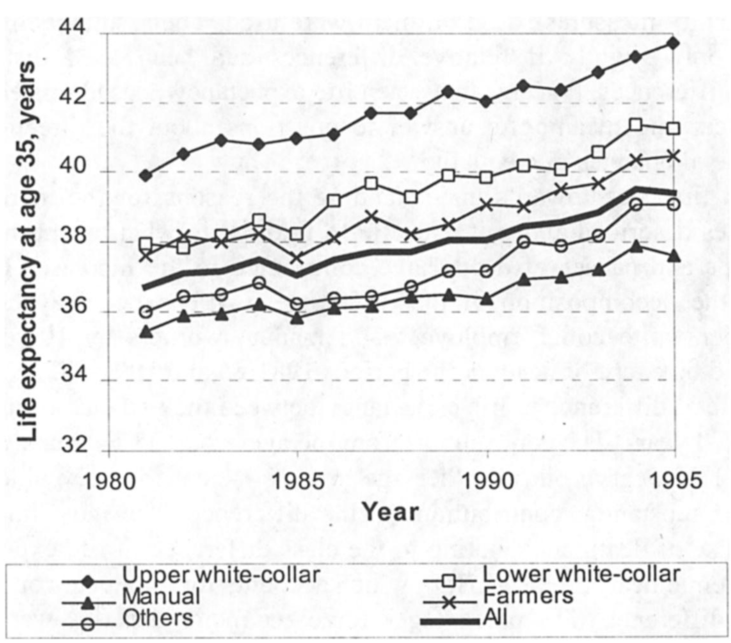
Statistics Finland permits one to link census records to death certificates issued after census by means of personal identification numbers. The data sets, used in this section, were constructed by linking the 1980 census records to deaths occurred in 1981-85, the 1985 census records to deaths occurred in 1986-90 and the 1990 census records to deaths occurred in 1991-95. More than 99.5% of all death certificates could be linked to the deceased persons' records of the previous census (Valkonen *et al.*, 1993, Martikainen *et al.*, 1995). The present consideration is restricted to men aged 35 years and over during the years from 1981 to 1995. Women and younger age groups have been excluded from the analysis due to the problems of classifying them by occupational class.

Economically active people were classified according to their occupation reported in the census on which the relevant subset of data on deaths was based. To avoid the bias associated with so-called "healthy worker effect" it was necessary to classify economically non-active people as well (Valkonen *et al.*, 1997). Therefore, pensioners, unemployed persons and others for whom information on current occupation was not available were classified according to occupational information in earlier censuses. One per cent of the people could not be classified and were included in the category of "others". The following classification was used: upper white collar employees (managers and higher administrative and clerical employees), lower white collar employees, manual workers, farmers, and others (self-employed, students or people with current or former occupation unknown).

Person-years at risk and numbers of deaths by cause for the period 1981-95 were tabulated by year, sex, age (5-year age groups) and social class. These data are used for the analysis of the socio-occupational differentials in life expectancy.

Life expectancy of the Finnish men at age 35 was 36.7 years in in 1981. There were marked differences between social classes: the life expectancy for upper white collar employees was 39.9 years but only 35.4 years for manual workers. Other classes fell between these two extremes (Figure 4).

Figure 4 – *Male life expectancy at age 35 by socio-occupational class in Finland in 1981-1995*



Male life expectancy at age 35 increased rapidly over the 1980s and 1990s and was 2.7 years higher in 1995 than in 1981. Life expectancy increased in all social classes but not equally. For example, the increase was 3.8 years among the upper white-collar employees, but only 2.2 years among manual workers. As a result, the gap between the two extreme classes was 1.6 years larger in 1995 than in 1981.

Figure 5 summarizes the changes in class inequalities using the index of dissimilarity in length of life (*IDLL*). In spite of relatively large annual fluctuations in inequalities, it is clear that the degree of inequality tends to increase. This increase was relatively slow in the first years of the 1980s and became much more rapid in the second half of the 1980s. The *IDLL* for men increased from 1.36 years in 1981-85 to 1.71 years in 1991-95 (26.3%).

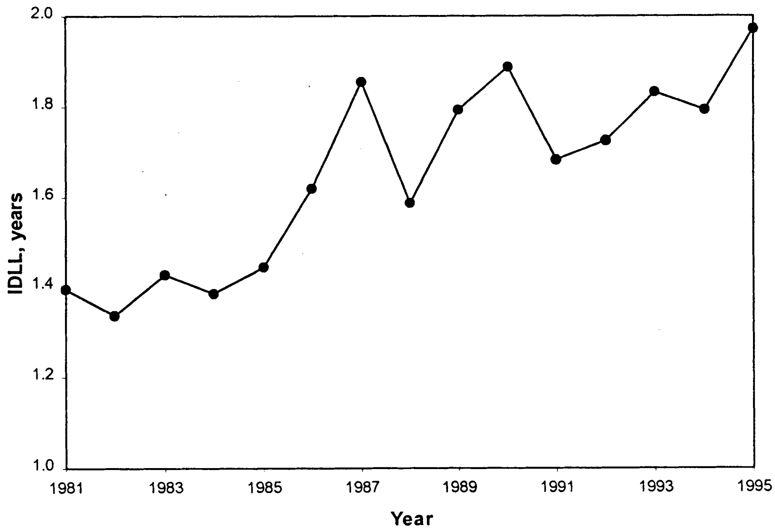
When Mackenbach *et al.* (1997) used Finnish data on class differences in death rates to illustrate the use of different measures of inequality they observed that different measures gave different results. The choice between absolute and relative measures of inequality turned out to be particularly important. When the level of mortality decreases – as has been the case in most countries – relative differences in age-standardized mortality may increase at the same time when absolute differences diminish. In this situation it is difficult to say whether inequalities have decreased or increased.

In the analysis presented above, absolute differences in life expectancy and inequality measures based on them were used. The results would have changed only slightly if relative differences had been used instead of absolute differences. It seems that when life expectancy is used it is easier to give a clear and meaningful answer to questions about the direction and magnitude of changes in inequality.

As a first step towards understanding the reasons for the increase in inequalities described above it is useful to find out to what extent different age groups and causes of death have contributed to the increase. Table 3 presents the decomposition of the difference in life expectancy between male upper white collar employees and manual workers by 10-year age groups and by cause of death in the periods 1981-85 and 1991-95.

The total difference in life expectancy between the two classes in 1981-85 was 4.79 years. The contribution from the age group 55-64 years was the greatest (1.27 years), but all other age groups, except the very oldest one also made substantial contributions to the difference. The most important single cause of death contributing to the class difference in life expectancy was ischaemic heart disease (IHD), which accounts for 1.65 years or 34% of the total difference. The percentages for other major diseases were: lung cancer -14.5%, respiratory diseases -10.2%, cerebrovascular disorders - 5.5% and other cardiovascular diseases -6.7%.

Figure 5 – *Index of dissimilarity* ($IDLL_{abs}$) *in male length of life in 1981-1995*



In the youngest age-group (35-44) almost 60% of the total contribution were deaths from alcohol-associated causes (including alcoholic liver cirrhosis, other alcohol-related diseases, and accidental poisoning by alcohol), but the absolute and relative contributions of these causes were diminishing with age. About 20% of the overall class difference were due to these causes.

In 1991-95 the total difference between upper white-collar employees and manual workers was 5.68 years, 0.90 years more than ten years earlier. The general pattern of the contributions from age groups and causes of death was the same as in 1981-85: the contributions from the age groups 55-64 and 65-74 were the greatest, and ischaemic heart disease was the most important single cause of death. The differences between age- and cause-specific contributions in 1981-85 and in 1991-95 reveal the causes of the widening of the gap in more detail: 0.63 years or 70% of the total increase in the gap was due to alcohol-associated causes, suicide, and other accidents and violence. This increase took place almost entirely in the ages below 65. The contribution from the age-group 35-44 increased by more than 40%.

Table 3 – *Contributions of ages and causes of death to the overall difference in life expectancy at age 35 between white collar and manual workers in Finland in 1981-85 and 1991-95*

Age	35-44	45-54	55-64	65-74	75-84	85+	**All ages**
1981-85							
Lung cancer	0.008	0.089	0.264	0.246	0.083	0.008	**0.698**
Other cancers	-0.001	0.083	0.034	0.031	0.029	-0.002	**0.173**
Ischaemic heart diseases	0.161	0.385	0.482	0.490	0.107	0.007	**1.632**
Cerebrovascular disorders	0.034	0.045	0.078	0.045	0.050	0.004	**0.256**
Other cardiovascular diseases	0.033	0.035	0.086	0.091	0.084	0.013	**0.343**
Respiratory diseases	0.018	0.042	0.109	0.161	0.146	0.029	**0.504**
Alcohol-associated causes	0.095	0.091	0.058	0.010	0.002	0.000	**0.254**
Other diseases	0.035	0.044	0.028	0.056	0.039	-0.005	**0.197**
Suicide	0.138	0.082	0.052	0.022	0.005	0.001	**0.299**
Other accidents and violence	0.170	0.129	0.082	0.033	0.012	0.004	**0.429**
All causes	**0.690**	**1.025**	**1.271**	**1.184**	**0.557**	**0.058**	**4.786**
1991-95							
Lung cancer	0.010	0.070	0.180	0.256	0.088	0.016	**0.619**
Other cancers	0.018	0.019	0.094	0.061	0.049	0.000	**0.241**
Ischaemic heart diseases	0.106	0.243	0.521	0.526	0.230	0.034	**4.660**
Cerebrovascular disorders	0.047	0.051	0.073	0.075	0.064	-0.013	**0.297**
Other cardiovascular diseases	0.032	0.070	0.080	0.071	0.034	0.026	**0.314**
Respiratory diseases	0.031	0.050	0.094	0.178	0.144	0.048	**0.543**
Alcohol-associated causes	0.173	0.198	0.112	0.014	0.003	0.000	**0.500**
Other diseases	0.062	0.087	0.084	0.081	0.060	0.025	**0.399**
Suicide	0.233	0.117	0.060	0.021	0.002	0.003	**0.435**
Other accidents and violence	0.274	0.217	0.102	0.050	0.031	0.005	**0.678**
All causes	**0.985**	**1.121**	**1.400**	**1.332**	**0.705**	**0.143**	**5.686**

The group of "other diseases" contributed more than 20% to the increase of the mortality gap, so the role of other diseases was not insignificant.

Ischaemic heart disease had no effect on the change in the overall class difference. This seems to have been due to a decrease in the class difference in IHD mortality in the two youngest age groups, which compensated for the increase in the older age groups.

3.3 Differences between occupational groups in Moscow and St. Petersburg (the LRC cohort)

The Lipid Research Clinics program was part of a US-USSR collaboration in cardiovascular epidemiology in the 1970s. The two LRC-clinics were established in Moscow and St Petersburg (formerly Leningrad, renamed in 1991) to perform surveys of lipids, lipoproteins and other factors of cardiovascular diseases (US-USSR Steering Committee, 1977). A cohort of men born in 1916-1935 was randomly selected from voting lists of 1974 from one residential district in each of the cities. 7,815 men (from 10,034 selected) were examined and interviewed according to the North American LRC Prevalence Study Protocol in 1975-77 (3,908 in Moscow and 3,907 in St Petersburg). At the initial stage 57% of men were aged 40 to 49, 38% were aged 50 to 59 and the rest were aged 60 years and over. The initial average age of the cohort was about 48 years.

The information about socio-demographic characteristics of individuals was collected only once at the beginning of the follow-up. The mortality in the cohort was traced beginning with the initial survey. In 1994 the data collection process was stopped due to luck of funding from the Government of the Russian Federation. The activity resumed only in 1998 within the framework of the Global Health Equity Initiative (Shkolnikov *et al.*, 2000). The vital status of the cohort subjects was updated until spring of 1997. At this time 3,462 deaths (or 44.3%) were recorded and 243 subjects (or 3.1%) were lost from observation.

Although the LRC cohort represents relatively large numbers of deaths and person-years at risk, the data from the two metropolitan cities are not representative of the overall Russian population[7].

[7] In the 1970s and the 1980s the life expectancy in Moscow and St Petersburg was significantly higher than the Russian average due mostly to lower mortality from accidents and violence. However, in the early 1990s steep increases in violent deaths in the two metropolitan cities have resulted in narrowing the gap with the rest of the Russian population.

As the cohort under consideration is not homogenous by age, the survivorship function cannot be obtained simply from the temporal decrease in the size of the cohort. In such a situation, the method of "person years" can be applied for the computation of death numbers ${}_1D_x$ and exposure time ${}_1L_x$ (Kahn *et al.*, 1989).

$$ {}_1D_x = \sum_{t\in[0,T]} \sum_i d_{i,t}(x,x+1); {}_1L_x = \sum_{t\in[0,T]} \sum_i \tau_{i,t}(x,x+1) - \sum_{t\in[0,T]} \sum_\omega \tau_{\omega,t}(x,x+1), $$

where T is the duration of the follow up period; $d_{i,t}(x,x+1)=1$ if the individual i dies between the exact ages x and x+1 in calendar year t, otherwise $d_{i,t}(x,x+1)=0$; $\tau_{i,t}(x,x+1)$ is the time lived by the individual i between the exact ages x and x+1 in calendar year t; $\tau_{\omega,t}(x,x+1)$ is the time lived between the exact ages x and x+1 by the individual ω, who is withdrawn (due to the end of the follow-up period) or lost from observation between the exact ages x and x+1 in the calendar year t.

The age-specific mortality rate for age x is equal to ${}_1D_x / {}_1L_x$ is based on observations on all individuals who used to be at age x sometime during the period of observation. In other words, the survival function is determined for the whole period of follow-up disregarding chronological effects on age-specific mortality. Other life table functions can be estimated from the age-specific mortality rates in conventional way.

Occupational status of individuals in 1975-77 was categorized into four broad groups: the elite (state officials of higher ranks, functionaries of the communist party, trade unions and Soviet organizations as well as individuals with liberal professions). This group accounted for 8.6% of the cohort. The white-collar employees (engineers, teachers, physicians, clerical workers, *etc.*) accounted for 42.8% of the cohort. The blue-collar employees (manual workers of all qualifications) accounted for 43.1% of the cohort. Finally, the residual group (students, pensioners, jobless, disabled, *etc.*) accounted for another 5.5%. The white-collar employees and especially the elite demonstrate a considerably better survival rate than the blue-collar employees (Figure 6). The small group of "others" stands out due to its very low survival level, probably because of a considerable number of people within this group with initially poor health status.

Temporary life expectancies between exact ages 40 and 80 are equal to respective areas under respective survival curves in Figure 6. The estimates of life expectancies with their 95% confidence limits for the elite, the white collar, the blue-collar employees and the "others" are: 33.1 (31.9, 34.3), 31.0 (30.4, 31.6), 27.2 (26.4, 28.0), and 23.4 (20.6, 26.2) years, respectively. The difference between the two biggest groups of white and blue-collar employees in years lived within the range of ages 40 to 79 is 3.8 years.

Figure 6 – *Survival rate by occupational group in the Russian LRC cohort*

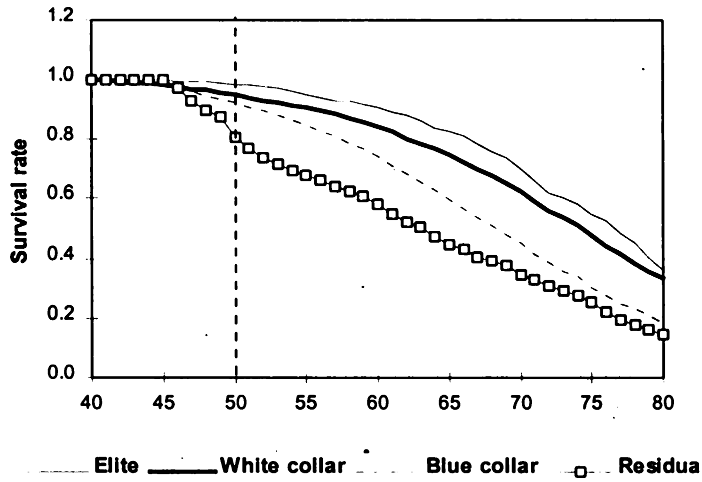
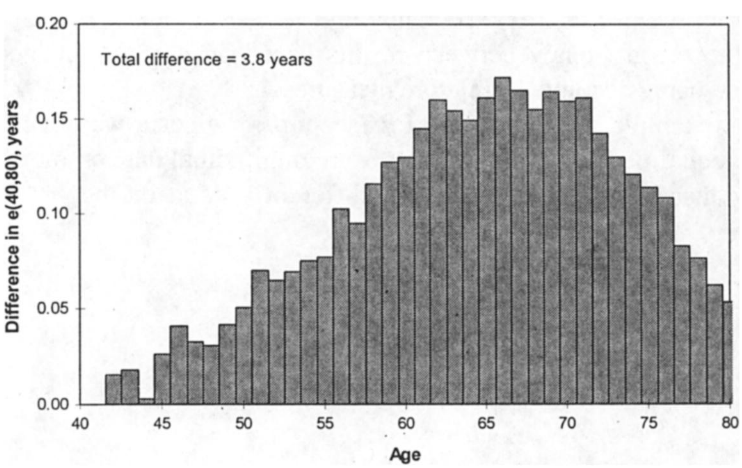


Figure 7 suggests that the maximum contributions to this gap are induced by differences in mortality rates between the ages of 60 and 70.

Figure 7 – *Age-specific contributions to the difference in life expectancy within the range of ages 40-79 between white and blue-collar employees in the Russian LRC cohort*



4. CONCLUSION

Life expectancy directly reflects prematureness of death, revealing a prevalence of younger or older ages at death in different social groups. In principle, the present study shows that it is possible to examine socioeconomic inequalities in mortality schedules in terms of life expectancy. An approximate linear decomposition of the life expectancy of the overall population by population group permits to build aggregate indices of inequality in length of life similar to aggregate indices of inequality in mortality rates.

The empirical examples illustrate techniques of operating with life expectancies. The examples of the racial differences in the USA in the 1980s and socio-occupational differences in Finland in the 1980s and 1990s show interesting cases of significant and growing inequalities in length of life. They also highlight important properties of the life expectancy as a measure of inter-group inequality. In particular, it appears that the average life expectancy for a nation can be quite stable in spite of big changes in life expectancy of relatively big population minorities (*i.e.*, of African Americans in the case of the USA). At the same time, the gap between better-off and worse-off tends to become smaller if both groups experience the same proportional decrease in age-specific mortality rates. This is so because life expectancy as a measure is more sensitive to the same proportional change in higher mortality rates than in lower ones.

In many western countries mortality is declining in all social groups, but this decline is usually steeper for better-off than for worse-off groups. In this case, absolute differences in mortality rates between rich and poor decrease, while the relative differences (ratios) in mortality rates increase. The example of changes in class differences in life expectancy in Finland shows that the use of life expectancy allows one to avoid, at least in some cases, the inconsistency between results based on changes in inequalities measured using absolute or relative measures.

The example of the Russian LRC cohort provides a way of measuring social inequalities in life expectancy from longitudinal data on mortality in a cohort consisting of people being at different ages at the beginning of the follow-up.

Acknowledgements

This study was completed within the framework of the Global Health Equity Initiative of the Rockefeller Foundation.

References

ANDREEV E.M. (1982), "Metod komponent v analize prodoljitelnosty zjizni. [The method of components in the analysis of length of life]", *Vestnik Statistiki*, 9, 42-47.

ANTONOVSKY A. (1967), "Social class, life expectancy and overall mortality", *Milbank Memorial Fund Quarterly,* 43, 31-73.

ARRIAGA E. (1984), "Measuring and explaining the change in life expectancies", *Demography*, 21(1), 83-96.

ARRIAGA E. (1989), "Changing trends in mortality decline during the last decades" in RUZICKA L. *et al.* (eds.), *Differential Mortality: Methodological Issues and Biosocial Factors*, Oxford, Clarendon Press, International Studies in Demography, 105-129.

CHIANG C.L. (1984), *The life table and its applications*, Malabar, Florida.

KAHN H.A., SEMPOS C.T. (1989), *Statistical methods in epidemiology*, Oxford University Press, New York-Oxford.

KOCHANEK K.D., MAURER J.D., ROSENBERG H.M. (1994), "Why did black life expectancy decline from 1984 through 1989 in the United States?", *American Journal of Public Health*, 84, 938-944.

KUNST A.E., MACKENBACH J.P. (1994), *Measuring socio-economic inequalities in health*, WHO Regional Office for Europe, Copenhagen.

LEE E.T. (1992), *Statistical methods for survival data analysis*, John Willey & Sons Inc.

MACKENBACH J.P., KUNST A.E. (1997), "Measuring the magnitude of socio-economic inequalities in health: an overview of available measures illustrated with two examples from Europe", *Social Science and Medicine*, 44(6), 757-771.

MARTIKAINEN P., VALKONEN T. (1995), *Lama ja ennenaikainen kuolleisuus*, Helsinki, Statistics Finland, Population 1995:11.

NCHS-WEB (1994), *National Center for Health Statistics*, http://www.cdc.gov/nchswww/datawh.

POLLARD J.H. (1982), "The expectation of life and its relationship to mortality", *Journal of the Institute of Actuaries*, 109(442), Part II, 225-240.

PRESSAT R. (1985), "Contribution des écarts de mortalité par âge à la différance des vies moyennes", *Population*, 4-5, 766-770.

PRESTON S.H., HEUVELINE P., GUILLOT M. (2000), *Demography. Measuring and Modelling Population Process*, Blackwell Publishers, Oxford, Malden, Massachusetts.

SHKOLNIKOV V., ANDREEV E., MALEVA T. (2000), *Neravenstvo i smertnost v Rossii [Inequality and mortality in Russia]*, Signal, Moscow.

THE MATHWORKS INC. (1999), *Optimization Toolbox. For use with MATLAB.*

US-USSR STEERING COMMITTEE FOR PROBLEM AREA 1 (1977), "The pathogenesis of atherosclerosis: collaborative US-USSR study on the prevalence of dyslipoproteinemia and ischeamic heart disease in American and Soviet populations", *American Journal of Cardiology*, 40, 260-268.

VALKONEN T., MARTELIN T., RIMPELÄ A., SAVELA S. (1993), *Socio-economic Mortality Differences in Finland 1981-90*, Helsinki, Statistics Finland, Population 1993:1.

VALKONEN T., MARTIKAINEN P. (1997), "Problems of comparability resulting from the exclusion of economically inactive persons in studies on social class differences in mortality" in HANCIOGLY A. *et al.* (eds.), *Hacettepe Institute of Population Studies*, Population No. IPS-HU.97-01, Ankara.

WILLEKENS F., SHAH I., SHAH J.M., RAMACHANDRAN P. (1983), "Multi-state analysis of marital status life tables: theory and application", *Population Studies*, 37(3), 129-143.

Appendix 1

According to Andreev (1982):

$$e_0^2 - e_0^1 = \sum_x {}_n\varepsilon_x^{2,1} = \sum_x \left[l_x^2(e_x^2 - e_x^1) - l_{x+n}^2(e_{x+n}^2 - e_{x+n}^1)\right],$$
$$e_0^1 - e_0^2 = \sum_x {}_n\varepsilon_x^{1,2} = \sum_x \left[l_x^1(e_x^1 - e_x^2) - l_{x+n}^1(e_{x+n}^1 - e_{x+n}^2)\right], \qquad [A]$$

Generally speaking, ${}_n\varepsilon_x^{2,1} \neq -{}_n\varepsilon_x^{1,2}$. The age-components of a difference between the two life expectancies depend on permutation of populations.

To avoid this problem Andreev (1982) and Pressat (1985) suggested using

$${}_n\varepsilon_x = \frac{1}{2}({}_n\varepsilon_x^{2,1} - {}_n\varepsilon_x^{1,2}) \qquad [B]$$

for a "symmetrical" decomposition of $e_0^2 - e_0^1$. It is easy to see that [A] and [B] lead to expression [2] from section 1.

The formula by Arriaga replicated in the recent textbook by Preston *et al.* (2000) is written in the following form:

$$e_0^2 - e_0^1 = \sum_x {}_n\Delta_x^{2,1} = \sum_x \left[l_x^1 \left(\frac{{}_nL_x^2}{l_x^2} - \frac{{}_nL_x^1}{l_x^1}\right) - T_{x+n}^1 \left(\frac{l_x^1}{l_x^2} - \frac{l_{x+n}^1}{l_{x+n}^2}\right) \right] \qquad [C]$$

Taking into account $T_x = l_x e_x$ and ${}_nL_x = l_x e_x - l_{x+n} e_{x+n}$ one can express ${}_n\Delta_x^{2,1}$ in a simpler form:

$${}_n\Delta_x^{2,1} = l_x^1(e_x^2 - e_x^1) - l_{x+n}^1(e_{x+n}^2 - e_{x+n}^1) \qquad [D]$$

Comparison of [A] and [D] shows that ${}_n\Delta_x^{2,1} = -{}_n\varepsilon_x^{1,2}$. The latter means that there is no difference between the components by Arriaga and those by Andreev-Pressat if one uses symmetrical components [B].

Appendix 2

Following Andreev (1982) we consider a component of the difference between two life expectancies at birth produced by the difference in mortality between populations 1 and 2 in a small age interval x, $x+\Delta x$. For simplicity we use here only the first term of the average from the right part of [2].

$$\varepsilon_{x,x+\Delta x} = l_x^2(e_x^2 - e_x^1) - l_{x+\Delta x}^2(e_{x+\Delta x}^2 - e_{x+\Delta x}^1) \qquad [A]$$

From the definitions of life table functions l_x, μ_x, e_x it can be shown that

$$l_{x+\Delta x} = l_x(1-\mu_x \Delta x); e_{x+\Delta x} = e_x - (1-\mu_x e_x)\Delta x,$$

where μ_x is the force of mortality at age x.

Using the latter expressions [A] can be transformed in

$$\varepsilon_{x,x+\Delta x} = l_x^2 e_x^1 \Delta x(\mu_x^1 - \mu_x^2).$$

The component corresponding to the age interval x, $x+t$ can be obtained as an integration by age between ages x and $x+n$:

$$\varepsilon_{x,x+t} = \int_x^{x+n} l_t^2 e_t^1(\mu_t^1 - \mu_t^2)dt \approx ({}_nM_x^1 - {}_nM_x^2)\int_x^{x+n} l_t^2 e_t^1 dt. \qquad [B]$$

Obviously, each central death rate ${}_nM_x$ can be decomposed by cause of death as ${}_nM_x = \sum_j {}_nM_{x,j}$. Thus, the contribution of cause of death j at age group $x,x+n$ is:

$${}_n\varepsilon_{x,j} = ({}_nM_{x,j}^1 - {}_nM_{x,j}^2)\cdot \int_x^{x+n} l_t^2 e_t^1 dt. \qquad [C]$$

If for all x ${}_nM_x^1 \neq {}_nM_x^2$ a comparison of [B] and [C] gives formula [3a] in section 1 for age-cause-specific components of the difference between two life expectancies. If at some age x death rates for two populations are equal ${}_nM_x^1 = {}_nM_x^2$ and cause-specific death rates differ between two populations then [C] can be used.

To make the expression [C] insensitive to permutations of two populations under study it can be transformed into a symmetrical form by averaging

$$ {}_n\varepsilon_{x,j} = ({}_nM_{x,j}^1 - {}_nM_{x,j}^2) \cdot \frac{1}{2} \cdot \left(\int_x^{x+n} l_t^2 e_t^1 dt + \int_x^{x+n} l_t^1 e_t^2 dt \right). \qquad [D] $$

Appendix 3

Let us, for a fixed age x, denote the unknown values $\theta_x^i = l_x^i / l_x$ as θ_i, life expectancies e_x^i as e_i, and population weights P_{x+}^i / P_{x+} as p_i. The following problem of minimization with constraints has to be be solved:

$$ \sum_{i=1}^{N} (\theta_i - p_i)^2 \to \min, $$
$$ \begin{cases} \sum_{i=1}^{N} \theta_i = 1, \\ \sum_{i=1}^{N} e_i \theta_i = e, \end{cases} \qquad [A] $$

The method of Lagrange multipliers allows to transform a problem of minimization with constraints [A] into a problem of minimization without constraints:

$$ L = \sum_{i=1}^{N} (\theta_i - p_i)^2 + \lambda_1 \left(\sum_{i=1}^{N} \theta_i - 1\right) + \lambda_2 \left(\sum_{i=1}^{N} e_i \theta_i - e\right) \to \min \qquad [B] $$

At the minimum point, all the derivatives of L with respect to θ_i and to λ_1, λ_2 should be equal to zero. It results in:

$$\begin{cases} 2\theta_1 + \lambda_1 + e_1\lambda_2 = 2p_1 \\ 2\theta_2 + \lambda_1 + e_2\lambda_2 = 2p_2 \\ 2\theta_3 + \lambda_1 + e_3\lambda_2 = 2p_3 \\ \dots. \\ 2\theta_N + \lambda_1 + e_N\lambda_2 = 2p_N \\ \sum_{i=1}^{N} \theta_i = 1, \\ \sum_{i=1}^{N} e_i\theta_i = e \end{cases} \qquad [C]$$

The system [C] of (N+2) linear equations with (N+2) unknown parameters $\theta_1, \theta_2, \dots, \theta_N, \lambda_1, \lambda_2$ can be rewritten in a matrix form as **A·z=b**, where

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 0 & .. & .. & .. & 0 & 1 & e_1 \\ 0 & 2 & 0 & .. & .. & .. & 0 & 1 & e_2 \\ 0 & 0 & 2 & .. & .. & .. & 0 & 1 & e_3 \\ .. & .. & .. & .. & .. & .. & .. & .. & .. \\ .. & .. & .. & .. & .. & 2 & 0 & .. & .. \\ 0 & 0 & 0 & .. & .. & 0 & 2 & 1 & e_N \\ 1 & 1 & 1 & .. & .. & .. & 1 & 0 & 0 \\ e_1 & e_2 & e_3 & .. & .. & .. & e_N & 0 & 0 \end{bmatrix}, \mathbf{z} = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ .. \\ .. \\ \theta_N \\ \lambda_1 \\ \lambda_2 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 2p_1 \\ 2p_2 \\ 2p_3 \\ .. \\ .. \\ 2p_N \\ 1 \\ e \end{bmatrix},$$

The solution of the system [C] is $\mathbf{z} = \mathbf{A}^{-1}\,\mathbf{b}$.