



Demographic Research a free, expedited, online journal
of peer-reviewed research and commentary
in the population sciences published by the
Max Planck Institute for Demographic Research
Konrad-Zuse Str. 1, D-18057 Rostock · GERMANY
www.demographic-research.org

DEMOGRAPHIC RESEARCH

VOLUME 12, ARTICLE 13, PAGES 323-380

PUBLISHED 28 JUNE 2005

www.demographic-research.org/Volumes/Vol12/13

DOI: 10.4054/DemRes.2005.12.13

Research Article

**Geographical diversity of cause-of-death
patterns and trends in Russia**

Jacques Vallin

Evgeni Andreev

France Meslé

Vladimir Shkolnikov

© 2005 Max-Planck-Gesellschaft.

Table of Contents

1	Introduction	324
2	Data and methods	335
2.1	Grouping causes of death	335
2.2	Hierarchical cluster analysis	336
2.2.1	Period-specific geographical patterns	338
2.2.2	Time-scale overall geographical patterns	339
3	To each period its own geographical pattern of causes of death	339
3.1	1969-1970	342
3.2	1978-79	344
3.3	1988-89	346
3.4	1993-94	348
4	Constant geographical contrasts	350
5	Four main clusters explanatory of how geographical contrasts contribute to general mortality dynamics	358
5.1	Mortality dynamics by clusters	358
5.2	Contribution of geographical dynamics to all-Russia mortality changes	363
6	Conclusion	365
7	Acknowledgements	367
	References	368
	Annex I	371
	Annex II	374
	Annex III	378
	Annex IV	379

Geographical diversity of cause-of-death patterns and trends in Russia

Jacques Vallin¹

Evgeni Andreev²

France Meslé³

Vladimir Shkolnikov⁴

Abstract

This paper performs a systematic analysis of all currently available Russian data on mortality by region, census year (1970, 1979, 1989, and 1994) and cause of death. It investigates what links may be found between these geographical variations in cause-specific mortality, the negative general trends observed since 1965, and the wide fluctuations of the last two decades. For that, four two-year periods of observation were selected where it was possible to calculate fairly reliable mortality indicators by geographic units using census data for 1970, 1979, 1989, and micro-census data for 1994, and a clustering model was used.

Behind the complexity of the studied universe, three main conclusions appeared. Firstly, in European Russia, there is a stark contrast between south-west and north-east, both in terms of total mortality and of cause-of-death patterns. Secondly, analysis of overall cause-of-death patterns for all periods combined clearly confirms that contrast at the whole country level by the prolongation of the southern part of European Russia through the continuation of the black soil (“chernoziom”) belt along the Kazakhstan border, while the rest of Siberia presents a radically different picture to European Russia. Thirdly, while it is difficult to infer any permanent geographical pattern of mortality from that very fluctuating piece of history, 1988-89 appears to be a base period for at least the entire period from 1969-1994.

¹ Institut national d'études démographiques, E-mail: vallin@ined.fr

² Centre of Demography and Human Ecology, Moscow, E-mail: evg_andreev@ns.cnt.ru

³ Institut national d'études démographiques, E-mail: mesle@ined.fr

⁴ Max Planck Institute for Demographic Research, E-mail: Shkolnikov@demogr.mpg.de

1. Introduction

Russian life expectancy has been stagnating for females and declining for males since the mid-1960s, with a wide fluctuation in the 1980s and 90s related to the 1985 anti-alcohol campaign and the 1992-93 socio-economic crisis (Figure 1). The long-run adverse trends are mainly due to increases in cardiovascular diseases, alcohol-related mortality and violent deaths (Shkolnikov *et al.*, 1996; Meslé *et al.*, 2003). Changes in alcohol consumption are responsible for the abrupt rise in life expectancy observed in 1985-86, and the subsequent decrease of 1990-92 (Meslé *et al.*, 1994; Shkolnikov and Nemtsov, 1997; Leon *et al.* 1997), while the even sharper decrease observed in 1992-94 stems from more varied causes of death, and is involved with social and economic difficulties encountered by individuals and families in the transition to a market economy (Meslé *et al.*, 1998; Shkolnikov *et al.*, 1998; Gavrilova *et al.*, 2001). With adaptation of people to the new economic situation, life expectancy returned to its previous levels, but since 1998, it has resumed its long-term downwards trend (Meslé *et al.*, 2003).

It is, however, a matter of record that mortality varies within Russia from place to place. Patterns of regional mortality variation have been documented in a number of studies (Andreev, 1979; Shkolnikov, 1987; Shkolnikov and Vassin, 1994; Vassin and Costello, 1997; Jozan and Prokhorskas, 1997; Walberg *et al.*, 1998).

The general pattern of mortality increase from south-west to north-east, with lower mortality in the blacksoil regions of southern European Russia and the southern part of West Siberia, and higher mortality in northern European Russia, Ural, Siberia and the far East, was identified by Evgueni Andreev (1979) from 1970 data, and by Vladimir Shkolnikov (1987) from mortality data around the all-Soviet censuses of 1970 and 1979. The latter study found correlations between this geographical pattern and inter-regional differences in general socio-economic development, climate conditions, and alcohol consumption. Studies by Vladimir Shkolnikov and Serguei Vassin (1994), and by Peter Jozan and Remigijus Prokhorskas (1997), described geographical patterns of mortality from principal and “avoidable” causes around the following census of 1989. Serguei Vassin and Christine Costello (1997) used the same data to classify Russian regions by the shapes of their mortality age curves. Finally, Peder Walberg *et al.* (1998) analyzed decreases in male life expectancy from 1989 to 1994 (the last Russian micro-census) across regions of European Russia, and found some associations between these decreases and the acuity of labor market changes.

This paper performs a systematic analysis of all currently available Russian data on mortality by region, census year (1970, 1979, 1989, and 1994) and cause of death. It investigates what links may be found between these geographical variations in cause-specific mortality, the negative general trends observed since 1965, and the wide

fluctuations of the last two decades. For that, four two-year periods of observation were selected where it was possible to calculate fairly reliable mortality indicators by geographic units using census data for 1970, 1979, 1989, and micro-census data for 1994. As Figure 1 shows, 1970 may be taken as the base year (1965 would have been preferable, but it is not a census year); 1979 is the closest census year prior to the anti-alcohol campaign; 1989 is fairly representative for the highest post-campaign life expectancy, and 1994 is the lowest point at the peak of the socio-economic crisis. Unfortunately, no reliable geographic mortality data are available for the most recent years⁵. However, these four points may already go a long way to explaining the relationships between time changes in life expectancy and geographical variations of mortality.

Also of interest would be to use life expectancy by regions as an indicator of geographical variations of mortality. However, life expectancy has complicated non-linear relations with cause-specific mortality depending on their weights and age patterns. In particular, life expectancy is not easily decomposable by cause (Shkolnikov *et al.*, 2001). In this sense, age-standardized death rate (SDR) is a more suitable indicator for a geographical analysis of cause-specific mortality. The choice of indicator will have little impact on the outcome since the all-cause SDR is closely correlated with life expectancy at birth across the Russian regions (Table 1).

Table 1: Correlation between SDR and life expectancy at birth among the 73 Russian administrative units

Sex	All periods	1969-70	1978-79	1988-89	1993-94
Males	-0.91	-0.88	-0.86	-0.75	-0.86
Females	-0.90	-0.89	-0.92	-0.87	-0.92

⁵ The 2002 census results are not yet available.

For each of the 73 administrative units used here ⁶ (see Annex I) SDR ⁷ were computed on the basis of deaths by age and sex-specific mortality in the immediately pre-census and census years. Since censuses were taken at the very beginning of the census year ⁸, it was possible to use the age-specific population numbers reported by the census as denominators.

Figure 2 shows the geographical variations in age-standardized mortality rates for each of the four periods. For each sex, the four maps were drawn according to the same 7 value classes. The size of these classes is of one standard deviation, and the central one is centred on the mean all-Russia value for the four periods. In fact, since the range of values is much wider above than below this mean value, the central class is the third one ⁹.

When considering the maps, it must be borne in mind that the Russian population is very unequally distributed over its territory. A large part of the total population lives in administrative units mostly concentrated in the European part of the country and the south-western part of Siberia. By contrast, the very large units of northern Siberia, especially Tyumen Oblast, Krasnoyarsk Kray, and the Republic of Sakha (Yakutia), are sparsely-populated. For that reason, when interpreting the maps, more weight must be given to colour variations among the smaller western and southern areas than the large Siberian ones. Moreover, the population of territories like Tyumen Oblast or Krasnoyarsk Kray, is mainly concentrated in their very small southernmost tips (identified by dotted lines on the maps). Unfortunately, however, no data were available for these administrative sub-units, and so the entire territory is coloured as for that small part, which may give a misleading impression at first glance.

⁶ The present Russian territory is divided into 2 cities (Moscow and St Petersburg), 49 oblasts, 6 krais, 21 republics, 1 autonomous district (Chukchi AD), and 1 autonomous oblast (Jewish AO) (Goskomstat, 2002, p. 13). Prior to 1991, instead of these 21 republics, there were 15 "autonomous republics" that were direct members of the Russian Federation (including the Chechnya-Ingushia, which, in 1993, was split into two republics: Chechnya and Ingushia), and 4 "autonomous oblasts" that were parts of 4 krais. Chukchi autonomous district was part of Magadan Oblast and the Jewish autonomous oblast was included in Khabarovsk Kray. Data for these 5 autonomous oblasts, 1 autonomous district, and for Chechnya and Ingushia separately, were not available and we can only use here data at the level of whole krais and the whole autonomous republic of Chechnya-Ingushia. Accordingly, 73 geographical units (2 cities, 49 oblasts, 6 krais, and 16 republics, see Annex I) are used here. Furthermore, no data is available for Chechnya for 1993-94, and the entire autonomous republic of Chechnya-Ingushia has been left in blank.

⁷ SDR were computed on the basis of the WHO (1992) standard European population.

⁸ Exact census dates were 15 January 1970, 17 January 1979, and 12 January 1989. In 1994, the micro-census accounted for 14 February; however, this was not used directly as a denominator, but only to correct the post-census population estimate for 1 January 1994.

⁹ To depict those maps in life expectancy terms, mean life expectancies of each class of SDR are given in Annex IV.

Figure1: Trends in life expectancy since 1965 and the four points selected for geographical analysis

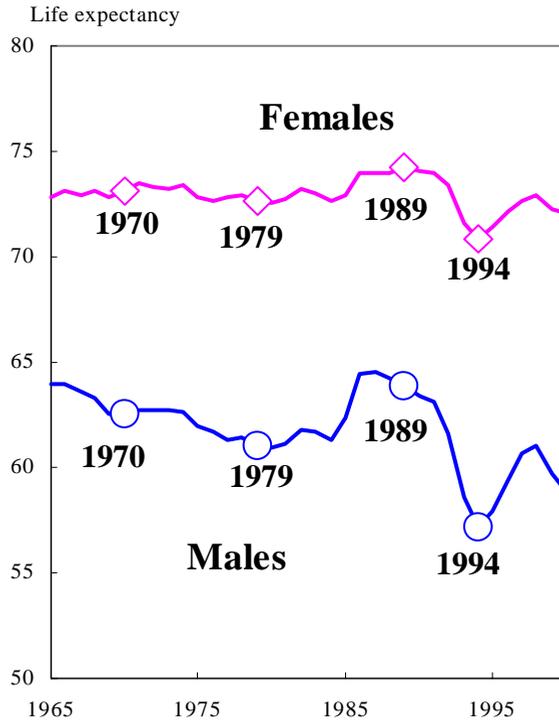


Figure 2a: Variation of age-standardized mortality rate among 73 administrative units in 1969-70, 1978-79, 1988-89 and 1993-94. Males

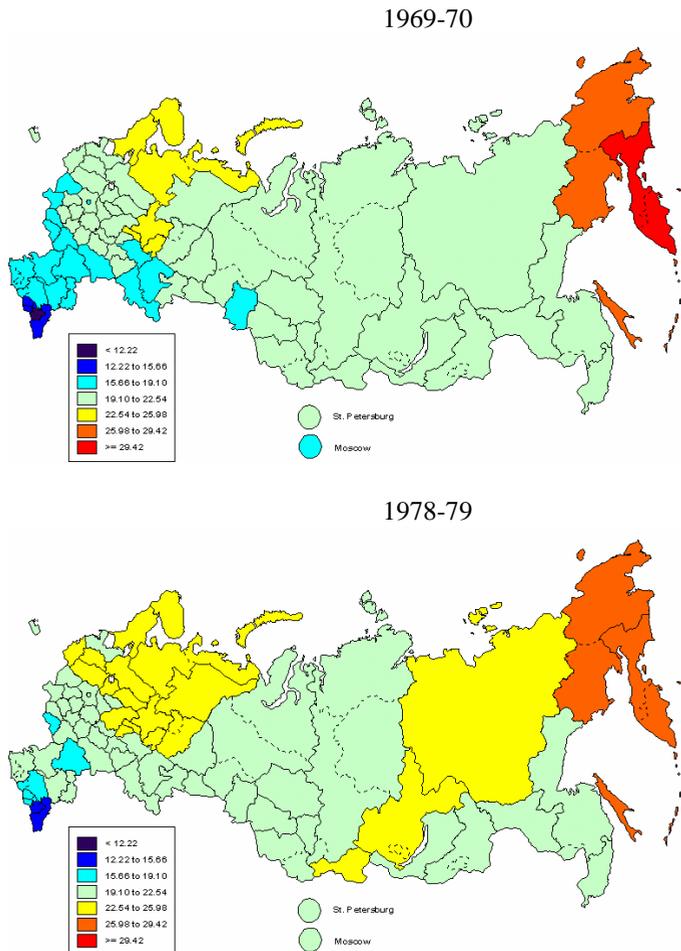


Figure 2a: Males (continued)

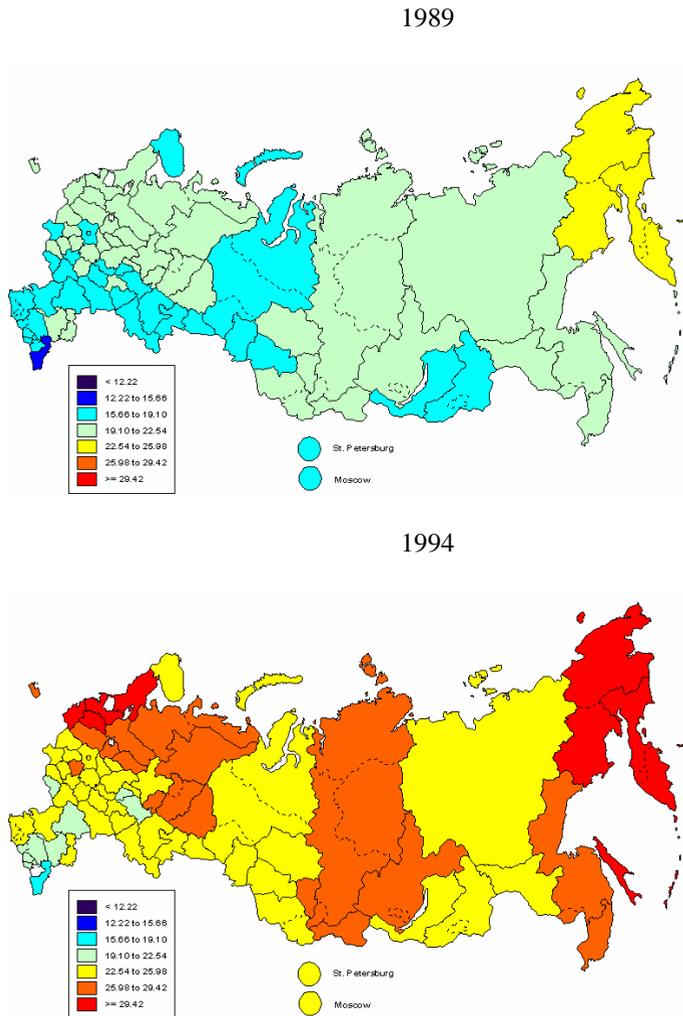


Figure 2b: Variation of age-standardized mortality rate among 73 administrative units in 1969-70, 1978-79, 1988-89 and 1993-94. Females

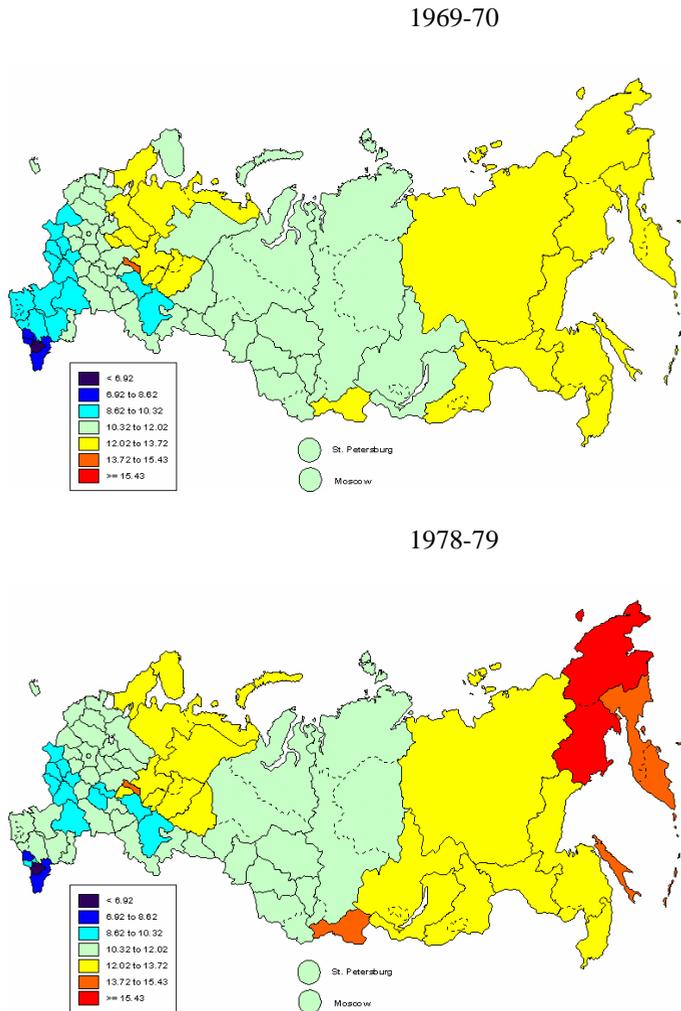
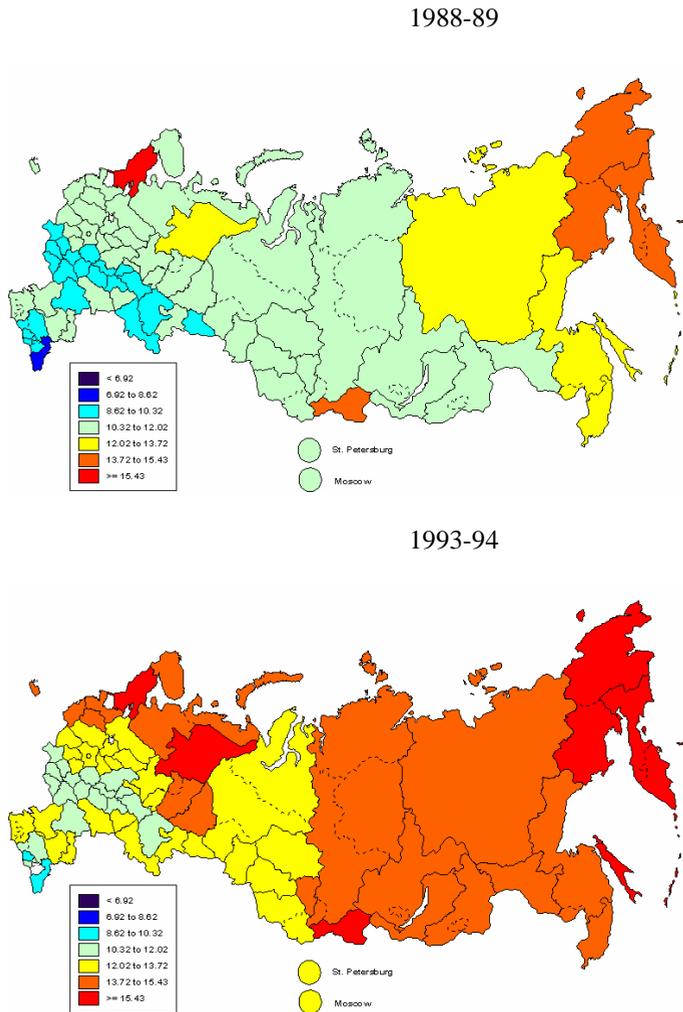


Figure 2b: Females (continued)



On the male side (Figure 2a), in 1970, a large part of Russia was less than half a standard deviation from the mean (1970-1994) all-Russia SDR value. This included most oblasts of central European Russia and almost all Siberian territories, and accounted for 58.6% of the total male population (Table 2). This means that, at a time when the big post-war gains in life expectancy from the decline in infectious diseases had occurred, Russia had grown increasingly homogeneous. However, two main deviations from the mean were observed: a much better situation in southern European Russia plus Moscow (32.1% of the total population), and a slightly worse one in its northern part (5.3%). The far East (Magadan, Kamchatka, Sakhalin) was affected by much higher mortality, but, again, these are sparsely-populated territories (1.1%). Some Caucasian republics (Dagestan, Kabardin-Balkar, North Ossetia, Chechnya-Ingushia), accounting for 2.9% of the total population, seemed to enjoy notably low mortality, but part at least of that advantage may only be apparent, stemming from under-registration of infant and old-age mortality (Andreev and Kvasha, 2002; Meslé *et al.*, 2003).

Table 2: Percentage of total population living in regions included in each of the SDR intervals, by sex.

SDR intervals	Males			
	1970	1979	1989	1994
<12.22	0.8	0	0	0
12.22-15.66	2.1	2.1	1.2	0
15.66 to 19.10	32.1	5.5	54.2	1.3
19.10 to 22.54	58.6	71.8	43.8	9.4
22.54 to 25.98	5.3	19.4	0.8	61.1
25.98 to 29.42	0.8	1.2	0	23.5
>= 29.42	0.3	0	0	3.9
Not available				0.9

SDR intervals	Females			
	1970	1979	1989	1994
<6.92	0.8	0.8	0	0
6.92-8.62	1.9	1.7	1.2	0
8.62 to 10.32	23.6	16.0	24.4	1.8
10.32 to 12.02	59	61.3	68.5	23.7
12.02 to 13.72	14.1	18.5	4.6	49.5
13.72 to 15.43	0.5	1.4	0.8	21.7
>= 15.43	0	0.3	0.5	2.5
Not available				0.9

In 1979, after ten years of rising mortality, the map has shifted towards higher level classes of SDR and the yellow areas (71.8% of population) have expanded, while the two blue ones have shrunk (5.5% and 2.1%). In particular, in European Russia, the northern disadvantage has spread towards a large number of new oblasts, while the southern advantage involves only a very small number of territories.

By contrast, in 1989, the remarkable progress made with the anti-alcohol campaign has produced a reverse shift (Shkolnikov and Nemtsov, 1997). The low mortality area expands extensively from southern European Russia to many oblasts along the border between Siberia and Central Asia (again, bearing in mind that the large blue splash of Tyumen Oblast is only due to the weight of its small southernmost part), and the two blue areas account for 55.4% of the total male population. Moscow is again within that advantaged area, now joined by St Petersburg. By contrast, only two special far-Eastern territories (0.8%) are still affected by slightly higher mortality.

From 1989 to 1994, all territories experience a sharp rise in mortality. Almost the entire map is shaded in high mortality colours, with only a couple of territories, accounting for 10.7% of the total male population, remaining in the three lower mortality classes (Table 2).

Apart from these shifts in value levels between successive maps, it can be observed that the homogenisation at work in the 1960s continued in the 1970s. The unweighted standard deviation for SDR was 2.95 for males in 1969-70, falling to 2.64 in 1978-79 (Table 3), with an accompanying, though less pronounced, decrease for the weighted¹⁰ standard deviation. In other words, the decrease in life expectancy was higher for advantaged than disadvantaged areas. By contrast, the gains due to the anti-alcohol campaign were far more marked for disadvantaged areas, and the unweighted standard deviation fell dramatically to 1.70 in 1988-89, and even more so for the weighted SD. Correspondingly, the socio-economic crisis had a much greater impact on disadvantaged areas, while the unweighted standard deviation of SDR jumped to 3.01 in 1993-94 (2.32 for weighted SD).

¹⁰ As the unweighted standard deviation give the same weight to all regions regardless of population, it is useful to also check trends in standard deviation weighted by population size.

Table 3: Standard deviation of SDR (unweighted and weighted by population of regions) among 73 administrative units at four periods

Sex	1969-70	1978-79	1988-89	1993-94
Males				
Unweighted	2.95	2.64	1.70	3.01
Weighted	2.42	2.36	1.32	2.32
Females				
Unweighted	1.43	1.51	1.65	1.46
Weighted	1.35	1.29	1.28	1.10

At the same time, the asymmetric geographical distribution mapped in Figure 2 minimises in 1988-89, while increasing in the other three periods¹¹.

In all periods, the geography of female mortality is much more homogeneous than that of males (Table 3). Furthermore, standard deviation (weighted or not) varies very little from one period to the other, meaning that neither trends nor fluctuation greatly affect mortality level distribution. It is true that trends and fluctuation are equally less pronounced among females than males. Where male life expectancies declined, female expectancies stagnated, and the wide fluctuations observed among males were much slighter among females. At a lower level of contrast, however, women experienced the same type of geographical variations as men in all periods.

How does the cause-of-death structure of mortality contribute to these geographical changes over time? The first thing is to know if, for a specific period,

¹¹ Asymmetry a can be measured by the ratio of the distance from the median M to the upper limit Q_1 of the first quartile to that from M to the upper limit Q_3 of the third quartile:

$$a = \frac{M - Q_1}{Q_3 - M}$$

Thus, for the SDR sets mapped in Figure 2, a takes the following values :

Sex	1969-70	1978-79	1988-89	1993-94
Males	0.82	0.95	1.07	0.91
Females	0.65	0.85	1.25	0.58

For both sexes, a is relatively closer to 1 in 1988-89, indicating minimum asymmetry that year. Homogeneisation is thus accompanied by a decrease in asymmetry. Furthermore, 1988-89 is the only period where asymmetry is on the lower mortality side, which implies a reversal of asymmetry.

geographical regularities may be found in the cause-of-death structure of mortality. It would be then interesting to determine whether such regularities display any time-bound continuity in spite of or in relation to trends and changes in overall mortality. The first evident requirement is to select a way of measuring the distances between geographical units, in terms of cause-of-death patterns. Since geographical and time contrasts are more pronounced among males, the following analysis will be limited to them.

2. Data and methods

Two questions must be addressed: How to use discontinuous cause-of-death statistics? How to cluster geographical variations in cause-specific mortality?

2.1 Grouping causes of death

When using cause-of-death data at such a level, two problems are encountered. First, observed time-bound changes are partly due to disruptions in statistical time series. Second, observed geographical differences can be influenced by variations in data quality.

Since the late sixties, several classifications of causes of death have been in use in Russia, and deaths published according to old Soviet classifications had to be reclassified accordingly. As a result of a previous work, all cause-of-death data have been reclassified according to the most recent Soviet classification¹² (1981, as modified in 1988 for violent deaths) for all Russia, for each year, from 1956 to 1998 (Meslé *et al.*, 1996 and 2003). For this particular geographical analysis, all-Russia transition coefficients¹³ have been applied to cause-of-death data by region for the two-year period used¹⁴.

¹² During the Soviet era, a specific cause-of-death classification was applied in all the Republics. Even after 1990, that classification remained in use for around a decade in the new independent countries. The Soviet classification itself changed over time, and four successive versions were used in the review period. Over time, however, its main structure evolved to more closely mirror the WHO International Classification of Diseases, although it remains much less detailed (Meslé *et al.*, 1996).

¹³ The method of reclassification used for the all-Russia level relies on transition coefficients from one classification to the next, calculated after a complete analysis of medical definitions and statistical contents of individual items of both classifications (Meslé *et al.*, 1996).

¹⁴ Assuming that changes in classification had the same impact in each region as at all-Russia level, is certainly not always true. However computerised complete cause-of-death time series with which to check the actual differences are not available at the region level for the whole period. The regional analysis, therefore, is based on an assumption that the differences will not be significant for the results.

After testing different possibilities, the 175 items of the Soviet classification have been put into 16 groups meaningful for the regional analysis. That grouping retains as such the causes of death that are responsible for high mortality rates, but also includes more specific groups that could be clearly discriminant in geographic terms (Table 4).

In spite of the apparent specificity of their title, the items “atherosclerotic cardiosclerosis” (with or without hypertension) were merged with “ill-defined deaths” because, in the Soviet system still in use, “atherosclerotic cardiosclerosis” includes a significant proportion of imprecise diagnoses that should have been classified as “ill-defined” (Meslé and Vallin, 2003) (See also Section II).

Using such broad groups of causes partly but not entirely circumvents the data quality issue. In the Soviet system, causes of death were identified and coded at the local level and only the results were aggregated at higher levels. Theoretically, such a system amply accommodates geographical variations in cause determination. However, it may also be argued that the powerful Soviet era bureaucracy left local administration very little room for manoeuvre. Unfortunately, there are no specific studies of geographical variations in cause-of-death data quality that we could ascertain, such that we must use the available data “as found”, with the caveat that part of our observations will be coloured by that issue.

2.2 Hierarchical cluster analysis

We have therefore to work with a 3-dimensional dataset for 73 regions, 4 periods and 16 cause-of-death groups. Most standard methods assume more or less symmetric and interchangeable categories, which is not the case here, since our purpose is to analyse geographical mortality and its dynamics over 4 periods. A straightforward analysis of the 3-dimensional matrix is arguably inappropriate, as interactions among the three dimensions would interfere with interpretation of the results. What is known about cause-specific mortality trends in Russia from the 1970s-1990s suggests that time-scale and regional variations may be greater for certain causes of death, and that these causes may differ between the regional and time-scale dimensions (Meslé *et al.*, 2003; Shkolnikov and Vassin, 1994; Vassin and Costello, 1997). To gain insight into these peculiarities, regions were first clustered by their cause-of-death profiles independently for each time-period, after which an aggregated clustering was performed based on 64-dimensional (16x4) vectors, which made it possible to identify a more general structure characteristic for the whole time-period on average. Hierarchical analysis was applied to all clusters.

That analysis, based on Euclidian distance between cause-specific SDR patterns¹⁵ of the 73 geographical units (Annex I), was applied using within-group linkage methods¹⁶. This procedure attempts to identify relatively homogeneous clusters of geographical units based on selected characteristics, using an algorithm that starts with each unit in a separate cluster and combines clusters until only one is left. Thus, the numbers of clusters change in the process of calculation.

Table 4: Groups of causes used for geographical patterns of mortality by causes of death

Group number	Causes of death	Soviet Classification item numbers
1	Infectious diseases	1-44
2	Stomach cancer	47
3	Other digestive cancers	45, 46, 48-51
4	Cancer of larynx, trachea, bronchus and lung	52, 53
5	Other cancers	54-67
6	Atherosclerotic cardiosclerosis + Ill-defined	92, 93, 158, 159
7	Ischaemic heart diseases	90, 91, 94, 95
8	Other heart diseases	84-89, 96, 97
9	Cerebrovascular diseases and other diseases of the circulatory system	98-102
10	Influenza and pneumonia	104-107
11	Chronic respiratory diseases	103, 108-114,
12	Digestive diseases	115-127,
13	Other diseases	68-83, 128-157
14	Accidental alcohol poisoning	163
15	Suicide, homicide, and injury undetermined whether accidentally or purposely inflicted	173-175
16	Other external causes	160-162, 164-172

¹⁵ In fact, Russian mortality trends are heavily influenced by age-group-specific problems, especially at adult ages. A forthcoming study will focus on changes in the geographical mortality patterns in adult males only. This study seeks only to estimate changes in total mortality.

¹⁶ SPSS program, release 11.5 (SPSS, 2002), was used for the calculations.

2.2.1 Period-specific geographical patterns

For our classification, each territory in period t was presented as a vector of 16 values $W(r,t,i)$:

$$W(r,t,i) = \frac{SDR(r,t,i)}{\sigma_{SDR}(t,i)},$$

where:

SDR is the standardized death rate,

r is the geographical unit ($r = 0$ for Russia),

t represents the period (1969-1970, 1978-1979; 1988-1989, or 1993-1994),

i represents the cause-of-death group (from 1 to 16),

$\sigma_{SDR}(t,i)$ is the standard deviation among geographical units for the cause i during the period t .

To obtain the clusters for a specific period, the classification was made on vectors of the 16 values of W of each unit, and the distance between units is the Euclidian distance weighted by the Russian level of mortality for each cause, weighted by the square root of the all-Russia SDR ($\sqrt{SDR(0,t,i)}$), or:

$$\left(\sum_i \left((W(r_2,t,i) - W(r_1,t,i)) \cdot \sqrt{SDR(0,t,i)})^2 \right)^{\frac{1}{2}} \right)^{\frac{1}{2}}$$

Weighting is used here to reduce the distance's sensitivity to fluctuations in mortality from causes with relatively low, but highly variable, mortality. The SDR -weightings are square-rooted to reduce the cluster analysis's dependency on cause-of-death choices (see Annex III). The distance function with such weightings would be unaffected by, for example, the division of a broad cause into two more specific causes.

The distances thus calculated for each pair of units were then compared to produce clusters. Each unit starts out as a cluster. In step two, the two closest units are aggregated into one cluster, and the distances between clusters are recalculated. In step three, the two closest clusters are again merged, and so on until a reasonable number of clusters is reached. The analysis was continued until all units were aggregated into just two clusters, from which point the previous results were compared to the last one to determine the materiality of each additional cluster. Finally, a division into 10 clusters appeared to be most informative.

2.2.2 Time-scale overall geographical patterns

To analyse the persistence of geographical cause-of-death patterns over time, a global clustering for all four periods together was obtained on the basis of the distances between the vectors of 16 causes multiplied by 4 periods, i. e. $16 \times 4 = 64$ values of W for each geographical unit. Once again, Euclidian distances are weighted by $\sqrt{SDR(0,t,i)}$:

$$\left(\sum_t \sum_i \left((W(r_2,t,i) - W(r_1,t,i)) \cdot \sqrt{SDR(0,t,i)})^2 \right)^{\frac{1}{2}}$$

3. To each period its own geographical pattern of causes of death

Figure 3 shows the results of the cluster analysis applied to each of the four periods, independently. Although such maps are not comparable, and must be interpreted independently of one another, we defined colour sets for each to maximize ease of interpretation. The results of the time-scale global analysis (next section) were used for that, and detailed explanations are given in annex II. The principle is to assign the same colour to the period-specific clusters that are least distant from the four-period cluster as that assigned in the global analysis. For example, in the four following maps, areas of the cluster that was closest to the first cluster of the global analysis were shaded red, and so on. The same colours were therefore used for clusters showing similarities. In some cases, however, the same global cluster was selected as the closest for two period-specific clusters of the same period. In these cases, the same colour was used, but with distinctive hatching. In this way, for example, the individual cluster of Kamchatka Oblast is associated in 1988-89 with global cluster number 6, already associated with St Petersburg and Moscow cities that formed another cluster in that period. So, while each map shows the results for 10 clusters, the colour sets of some maps include a specific 11th colour, while one of the ten basic colours is missing (see Table II-2 in Annex II).

In all periods, each Figure 3 map shows dissimilarities that echo some of those appearing in Figure 2. In particular, the northern and southern parts of Western Russia present contrasting pictures. However, these contrasts are less evident, as the links between cause-of-death patterns and total mortality levels are complex. There are also wide between-period variations in cause-of-death geography, requiring each period to be considered independently. To clarify the differences between clusters, stars have been plotted for each according to the standardized mortality rates (SMR) by cause of

Figure 3: Males. Geographical distribution of 10 clusters according to cause-of-death structure

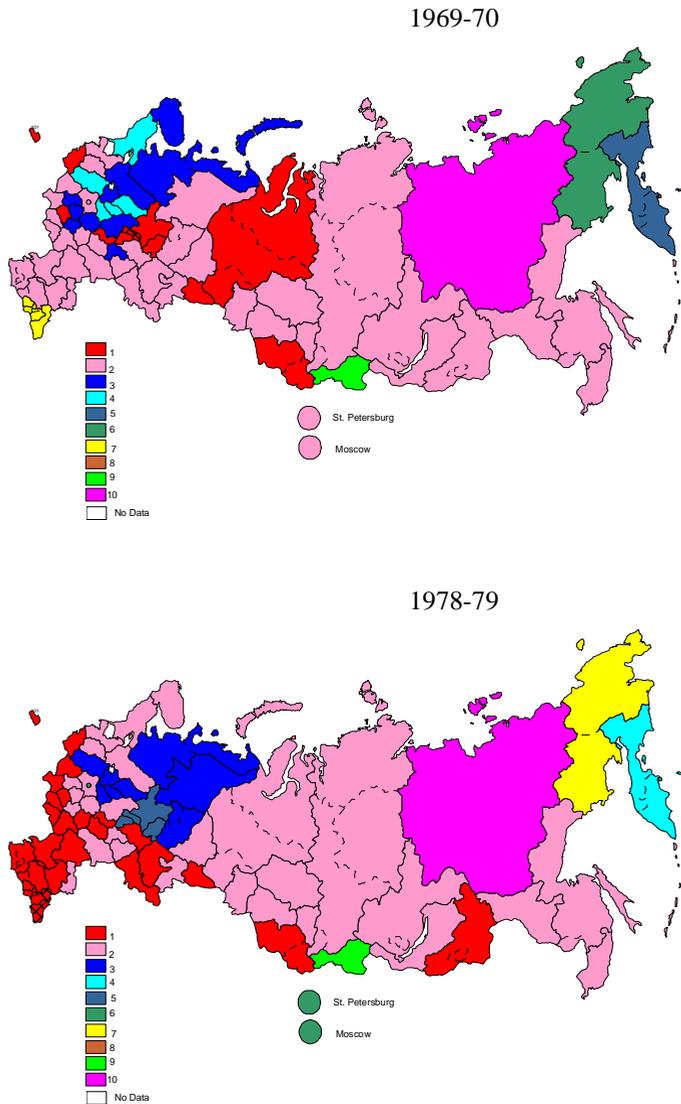
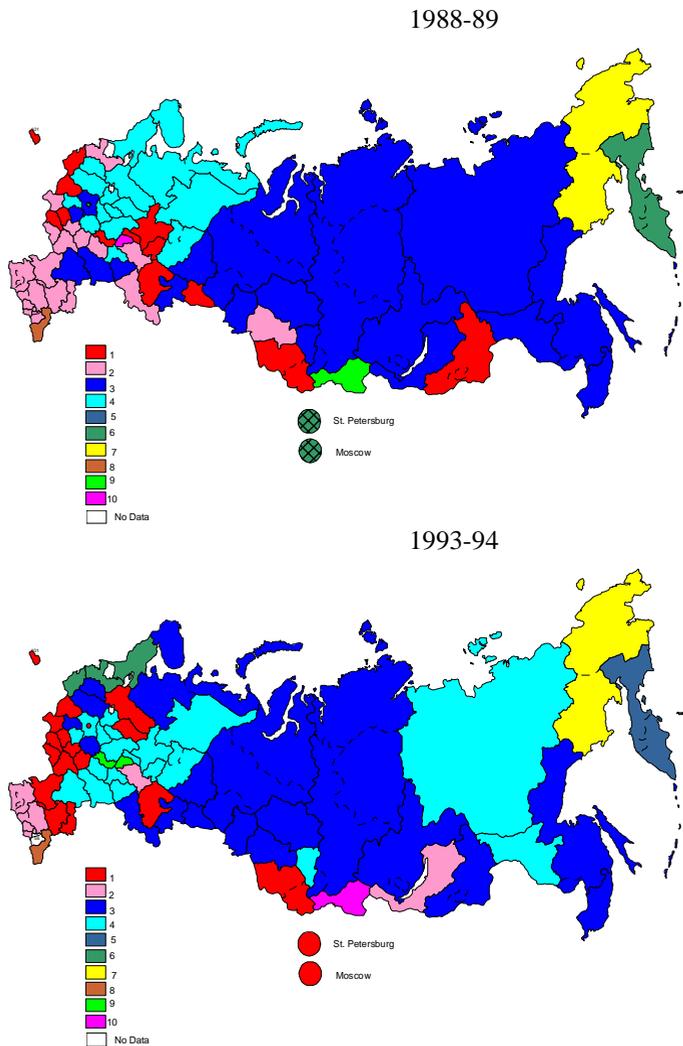


Figure 3: (Continued)



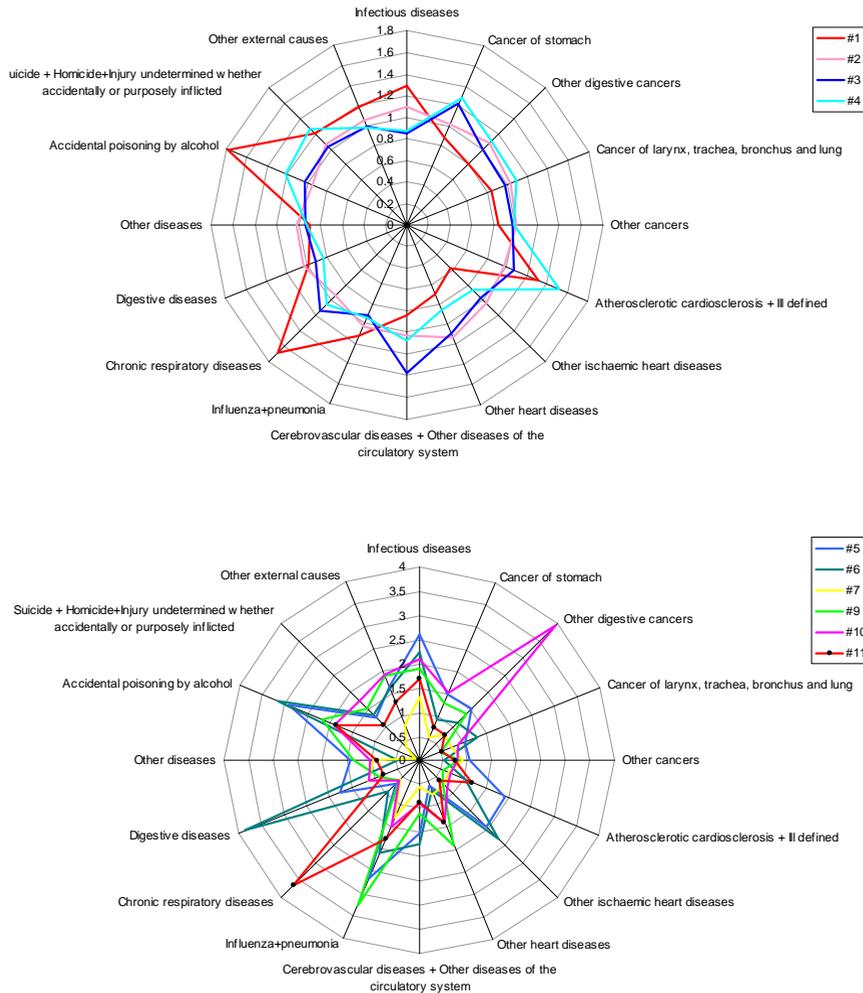
death for the cluster, divided by the Russian SMR. The stars are shown by period in Figures 4a to 4d. For each period, two graphs were plotted. The first aggregates stars of the four major clusters (those closest to the first 4 global clusters). The second aggregates the six remaining clusters, which are mostly quite singular clusters with very specific cause-of-death patterns, but comprising only a few territorial units. Most significance attaches to the shape of the first four stars.

3.1 1969-1970

Notwithstanding that Russian male life expectancy had been declining since 1965, the situation found at the end of the sixties still largely reflects the outcomes of previous decades' strong gains from the country-wide reduction of infectious diseases that produced greater homogeneity. This largely explains why most territorial entities fall in the same cluster (n° 2), coloured light pink on the 1969-70 map. This cluster aggregates 40 of the 73 units, including Moscow and St Petersburg (represented by two circles below the map to delineate them more clearly). The star for this cluster shown at the top of figure 4a appears clearly very close to circle 1 representing the all-Russia cause-of-death pattern. The second most important cluster (cluster 3, in dark blue on the map) aggregates 10 units, all located in the northern European part of Russia. It is also very close to the Russian mean, except for two groups of causes: stomach cancer and cerebrovascular diseases¹⁷. Stomach cancer was also a cause of excess mortality in a third cluster geographically located in the same part of Russia (cluster 4, in light blue), however was not particularly affected by cerebrovascular diseases, but showed strong excess mortality from atherosclerotic cardiosclerosis. The latter group is known to be quite ill-defined, and probably includes many circulatory system diseases, including cerebrovascular disease (Meslé *et al.*, 1998), although that fact may not be significant. By contrast, excess mortality from accidental alcohol poisoning, suicide and homicide is more representative of that cluster. Both of these groups are closely associated with alcohol consumption in Russia (Shkolnikov and Nemtsov, 1997). Alcohol consumption and especially accidental alcohol poisoning, is an even greater cause of excess mortality in the cluster shaded red (cluster 1), which aggregates 10 units, also mainly located in the northern part of European Russia, plus 3 units in western Siberia. That cluster is also strongly characterised by chronic respiratory diseases and to a lesser extent by

¹⁷ That group also includes "other diseases of the circulatory system", but that component is much less significant than cerebrovascular diseases, so the entire group is labelled "cerebrovascular diseases" here for simplicity's sake.

Figure 4a: Cause-of-death patterns of each cluster compared to the all-Russia cluster (cluster SMR/Russian SMR). Ten clusters, males, 1969-70



infectious diseases. There is also a notable excess mortality from atherosclerotic cardiosclerosis, but once again, this category is probably misleading, since its importance is very likely associated with an underestimation of cardiovascular mortality (and especially of “other ischaemic heart diseases”).

Of the other clusters, 5 comprise a single territorial unit. Each appears to display exceptional excess mortality for one specific group of causes. For example, in Yakutia (cluster 10), “other digestive cancers” mortality was 4 times higher than in Russia in that period, while chronic respiratory diseases are more than 3.5 times higher in the Mari Republic (cluster 11) than for all-Russia. Likewise, in Magadan (cluster 6), digestive diseases mortality was 4 times that of Russia, while influenza and pneumonia mortality was 4 times higher in the republic of Tuva (cluster 9). More interestingly, a specific cluster (n° 7) aggregating 4 units that comprise a continuous zone in the Caucasian region (the Republics of Dagestan, Chechnya-Ingouchia, North-Ossetia and Kabardin-Balkaria), is characterised by very low mortality for almost all causes except “infectious diseases” and “influenza and pneumonia”. Its “accidental alcohol poisoning” mortality was almost negligible and its “suicide and homicide” mortality was only one third that of Russia.

3.2 1978-79

The difference between 1969-70 and 1978-79 portrays here the results of the first step of a long-term decline in life expectancy. In 1978-79, cause-of-death patterns appear more heterogeneous among territories. The very large cluster coloured light pink on the 1969-70 map is split into two clusters coloured light pink (cluster 2) and red (cluster 1) on the 1978-79 map. While the light pink cluster still presents very similar cause-of-death pattern, quite close to the Russian mean, the red one is now different, but also close to the Russian mean (Figure 4b). They are distinguished by the fact that their slight differences from the all-Russia pattern are not the same (alcohol-related mortality is higher in the light pink cluster and lower in the red one, while the converse is true for “other ischaemic heart diseases” and cerebrovascular diseases).

By contrast, the new dark blue cluster (n° 3) tends to aggregate territories in the former dark and light blue clusters (northern part of European Russia), but is in fact characterised by a cause-of-death pattern closer to that of the former light blue than dark blue one. Indeed, it is mainly associated with excess mortality in alcohol-related causes of death and atherosclerotic cardiosclerosis, and to a lesser extent, stomach cancer. The fourth most interesting cluster in that period is the one comprised of the two big Russian cities (cluster 6, in dark green on the 1978-79 map). At that time, these

cities showed a specific cause-of-death pattern with higher mortality from all cancers (especially digestive cancer), “other ischaemic heart diseases” and “other diseases”. To a large extent, this pattern typically reflects improved diagnostic accuracy. By contrast, violent deaths, infectious diseases, and chronic respiratory diseases are much less significant than for all-Russia.

Once again, the lower part of Figure 4b displays stars for marginal clusters formed by small numbers of units. The five single-unit clusters (7 Magadan, 4 Kamchatka, 9 Tuva, 10 Yakutia, and 11 Dagestan) show marked peculiarities, almost identical to 1969-70 when the same unit is concerned. By contrast, the remaining cluster, grey-blue in 1978-79, corresponds to a part of the red cluster of 1969-70, aggregating the Republics of Mari, Chuvach, and Udmurtia, and Kirov Oblast (cluster 5). That cluster effectively shows a cause-of-death pattern close to that of the red cluster of 1969-70.

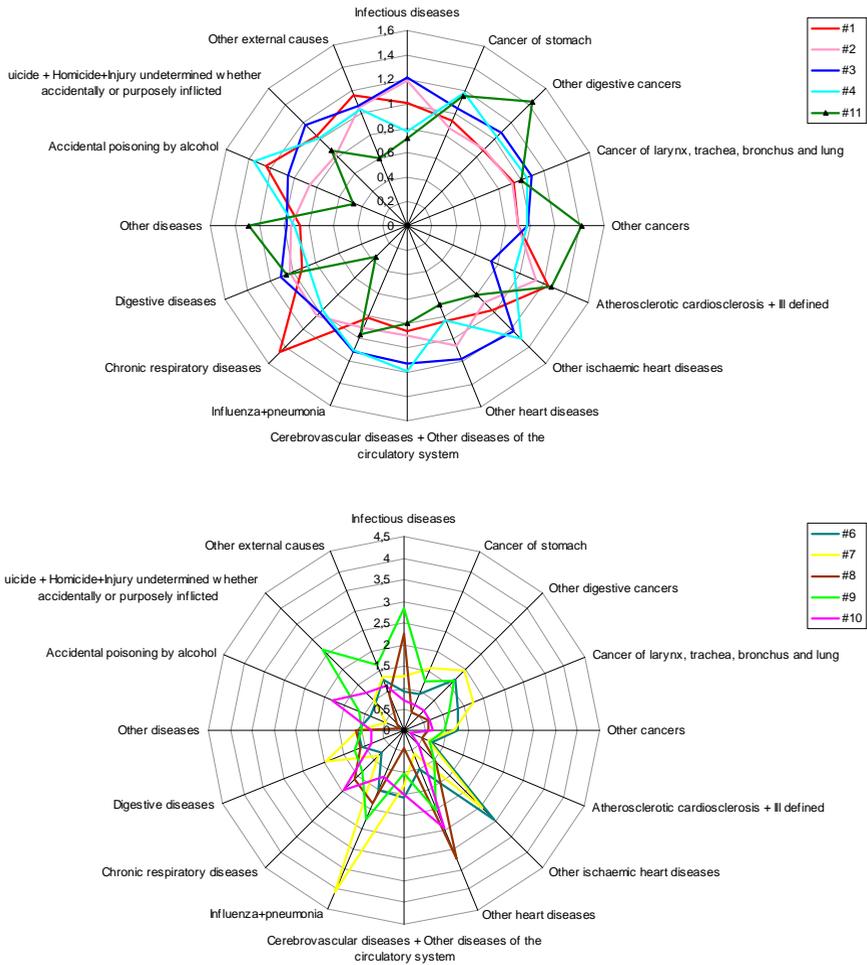
3.3 1988-89

The period 1988-89 characterised the situation produced by the positive results of the Gorbachev anti-alcoholism campaign. At that time, as shown in table 2, the geographical heterogeneity of total mortality was much less than during the other three review periods. Life expectancy gains from reductions in alcohol- and cardiovascular diseases-related mortality had been higher in regions where mortality was high. The 1988-89 map in Figure 3 also shows a much clearer geographical structure of cause of deaths than in any other period. There is a very sharp contrast between the northern and southern parts of European Russia, and it is the only time where there is a clear demarcation between the European part (here extended to the southern tip of western Siberia) and the Asian part of Russia. Homogenisation in term of levels attributable to the reduction in alcohol-related causes of death, which have a spurious effect on cause-of-death patterns, seems to have been associated with the emergence of a geographical pattern more related to the more fundamental cause-of-death structure.

The cause-of-death patterns of the three main clusters, coloured light pink (southern European Russia, cluster 2), light blue (northern European Russia, cluster 4), and dark blue (most of Siberia, cluster 3), are quite close to the all-Russia pattern. They are slightly discriminated, however, by specific causes of death: alcohol-related causes are more significant in northern European Russia and much less in southern European Russia, but the converse is true for infectious diseases. Most cardiovascular categories and infectious diseases are more significant in Siberia.

A fourth cluster (n° 1), coloured red, aggregates a series of geographic units along a West-East axis from the Baltic Sea (Kaliningrad Oblast) through Kirov Oblast,

Figure 4c: Cause-of-death patterns of each cluster compared to the all-Russia cluster (cluster SMR/Russian SMR). Ten clusters, males, 1988-89



Udmurt Republic, Bashkortostan Republic, and Kurgan Oblast to the Altai. Departing appreciably from the Russian pattern, it is marked by higher mortality from chronic respiratory diseases, atherotic cardio-sclerosis and accidental alcohol poisoning.

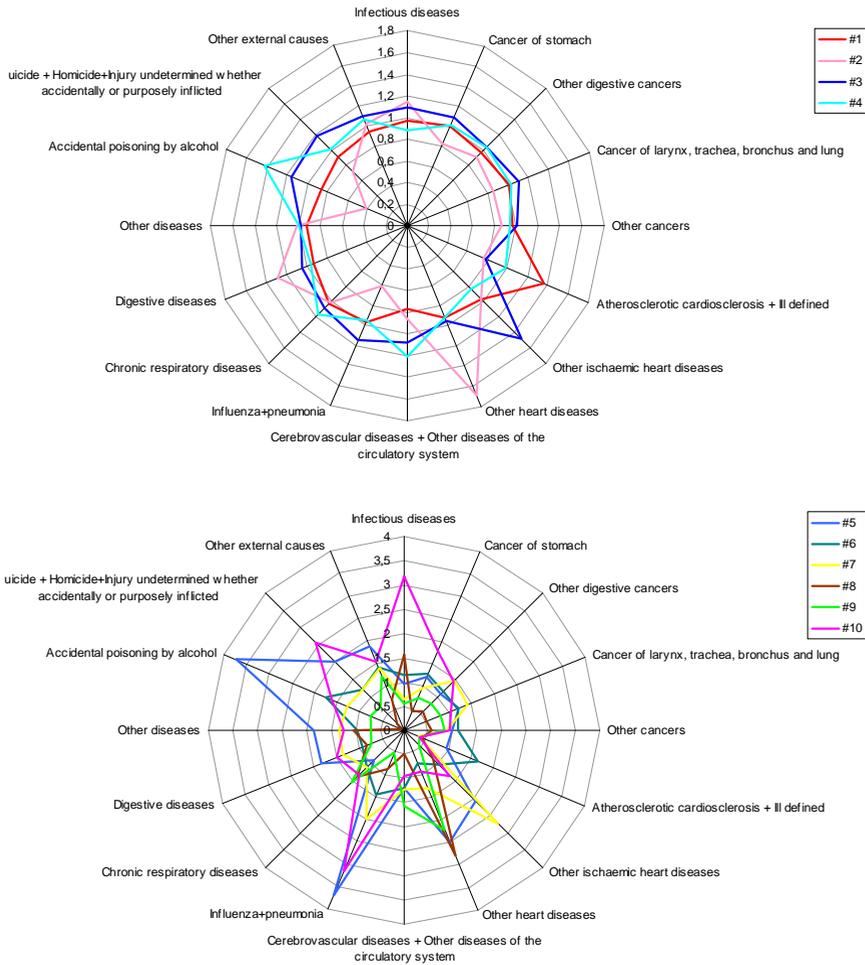
As in the previous period, Moscow and St Petersburg form specific cluster n° 11. Exceptionally, this is included with the four main ones to account for the population sizes of these two entities. Also, that cluster's profile departs much less from the mean than the following ones, included in the second star graph. The 1988-89 Moscow and St Petersburg profile closely matches that of the previous period. However, mortality from "other ischaemic heart diseases" is below the Russian mean, compared to above it in the previous period, and the converse applies to the broad category of atherosclerotic cardiosclerosis and ill-defined causes. In fact, this difference from the previous period reflects less any real change in the disease profile than in coding practice. This is because a new instruction from the Health Ministry in 1989 required any death over the age of 80 where no specific cause was mentioned in the medical file or autopsy report to be classed as "senility", and forbade deaths at younger ages from being reported as acute cardiovascular failure unless the diagnosis was confirmed by an autopsy report (Meslé and Vallin, 2003). The first requirement resulted mainly in transfers from atherosclerotic cardiosclerosis to ill-defined causes, which is without consequence for the broad group used here, but the second led to numerous transfers of specific cardiovascular diseases (here aggregated under "other ischaemic heart diseases") toward ill-defined causes, which is the main explanatory factor in the change of the star for that cluster.

As usual, small clusters generally comprising a single geographic unit, present singular profiles very far from the Russian mean. Significantly, however, the second graph of Figure 4b shows Yakutia no longer identified as a specific cluster, but incorporated in the large cluster 3.

3.4 1993-94

The fourth period is dominated by the effects of the economic and social crisis. Unsurprisingly, the corresponding map (Figure 3) shows clusters differently demarcated and much less compact than the previous three. The northern part of European Russia, very homogeneous and specific in the previous period, is now divided into four different clusters, each also including what may be far-flung other regions. What is shown by the Figure 4d stars has very little in common with what could be said of the previous ones. The most salient feature is that total mortality increased dramatically, with specific impacts in terms of causes of death, giving new geographical forms to the

Figure 4d: Cause-of-death patterns of each cluster compared to the all-Russia cluster (cluster SMR/Russian SMR). Ten clusters, males, 1993-94



clusters and new disease profiles to their stars. One general effect of that change is to make the disease profiles closer to the mean and less different from each other (Figure 4d). The first cluster (in red) that includes for the first time Moscow and St Petersburg is close to the Russian mean for all causes of death except atherosclerotic cardiosclerosis and ill-defined causes. It aggregates fairly well developed regions, including the two metropolitan areas where the mortality rise closely mirrored the national mean, but where also the new coding rules mentioned above were still applied.

The second cluster (in pink), showing the most favourable profile, with often sub-mean mortalities, relatively low impact of accidental alcohol poisoning and greater incidence of other heart diseases, has been reduced to a much smaller number of geographic units, almost entirely limited to Caucasian regions. For that reason, it is also marked by a relative importance of digestive and infectious diseases. Clusters 3 (dark blue) and 4 (light blue) display considerable geographical discontinuity, both including regions in the far East as well as middle Siberia and north-western European Russia. Mortality levels are almost invariably equal to or above the mean for all causes. The main difference between the two lies in “other ischaemic heart diseases”, which are the cause of much higher mortality in the third than fourth cluster.

Smaller clusters are more singular than ever. Kamchatka (cluster 5), Magadan (cluster 7), Tuva (cluster 10), and Dagestan (cluster 8), in particular, are once more isolated in specific clusters with generally the same type of deviations from the mean. For Kamchatka, however, the impact of accidental alcohol poisoning is very prominent.

4. Constant geographical contrasts

The wide mortality fluctuations and general time trends notwithstanding, are there any major geographical features that can be reliably retained when aggregating data for all four periods? To characterize the global geographical pattern, we used the distance measurement between the vectors of 16 causes multiplied by four periods for each geographical unit, as explained earlier in section I-B-2. As previously, we first defined a set of 73 clusters (as many clusters as geographical units), incrementally merging the two closest clusters, leaving only two clusters at the end, splitting all-Russia into two geographical entities.

This final result, shown in the first chart of Figure 5, reveals only an approximate mix of geographical differences, but is already highly illustrative of the geographical structure of Russian mortality. One interesting point is that the units aggregated in these two geographical entities display a high degree of geographical continuity. This includes mainly areas of the western and southern parts of European Russia, extended by a belt of south-western Siberia at the Kazakhstan border. The inclusion of Chita

oblast and the Republic of Sakha (former Yakoutia) arguably slightly weakens that continuity, but the importance of these very large sparsely-populated areas should probably not be over-stated. These two units aside, the second cluster clearly aggregates the rest of Asian Russia with all the north-eastern part of European Russia. A second point is that these two final clusters embrace two very significant and contrasting parts of geographical and historical Russia: the rich agricultural part running from the Northern Caucasus to the Lithuanian border, the Southern Volga basin and the Siberian prolongation of the chernoziom (black soil) belt, and the less hospitable regions of northern European Russia and Siberia. All-period mortality has been consistently lower in the former than the latter part. And Figure 6 shows that these two big clusters present two very different disease profiles.

Mortality is lower for almost all causes in cluster 1 than in cluster 2. The only exception is for chronic respiratory diseases that produce a slightly higher mortality in cluster 1 than all-Russia. However, the difference is much less than for the high excess mortality revealed by cluster 2 for two large groups of causes: violent deaths (including alcohol-related) and cardiovascular mortality. The standardized mortality rate by accidental alcohol poisoning is 40% above the Russian mean in cluster 2, and slightly below in cluster 1. Excess mortality in cluster 2 is also highly significant for the two other groups of violent death, especially suicide and homicide. Likewise, that cluster presents high excess mortality for the three broad groups of cardio-vascular diseases (ischaemic heart diseases, other heart diseases and cerebrovascular diseases). Another major cause of excess mortality for that group is related to acute respiratory diseases (influenza and pneumonia). Even digestive and respiratory cancer mortalities are also above the Russian mean.

However, there is a useful purpose to looking beyond this approximate split into the two largest clusters, and refining the analysis by returning to the previous steps of our clustering procedure. The three remaining charts of Figure 5 show the maps with 4, 6, and 8 clusters, and, finally, Figure 7, the map with 10 clusters. Arguably, the latter is the most appropriate for final comments. This is because the 4- and 6-cluster maps only isolate very specific geographical units, without disrupting the big areas shown by the 2-cluster map. It is only with the 8-cluster map that the first large area is split into two large parts, while the 10-cluster map (Figure 7) does likewise for the second large area. And these two large partitions are, again, of great geographical interest.

Of the 10 clusters, six comprise a single (clusters 10 Yakoutia, 6 Magadan, 5 Kamchatka, 9 Tuva, and 7 Dagestan) or at most two (cluster 8: Republic of Mordovia and Chuvach Republic) geographic units. These units are the same as those frequently isolated by the single period analyses.

Figure 5: Global clustering for the four periods together, according to 2 to 8 clusters. Males

Males, 2 clusters



Males, 4 clusters



Figure 5: (Continued)

Males, 6 clusters



Males, 8 clusters

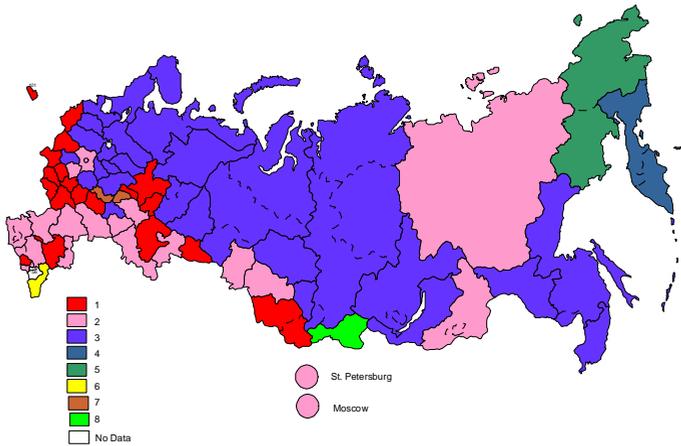


Figure 6: Cause-of-death patterns of the two final clusters compared to the Russian cluster (cluster SMR/Russian SMR). Males, four periods altogether

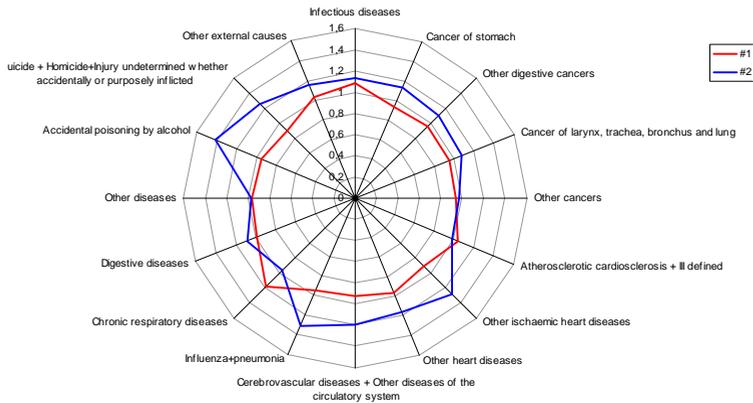
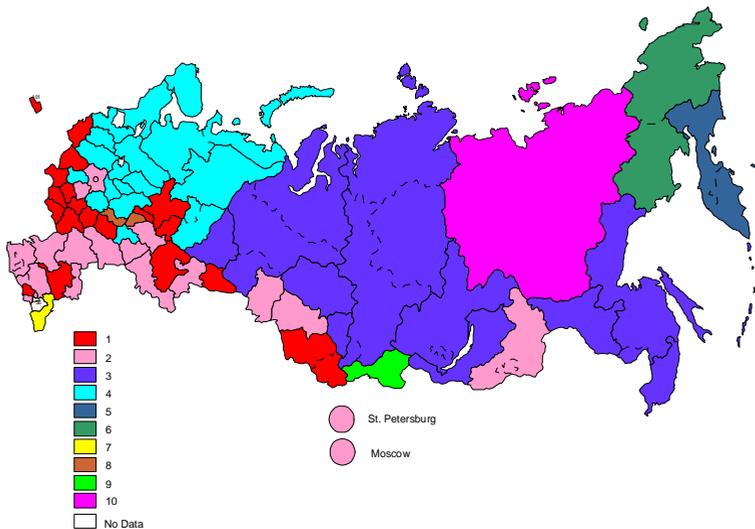


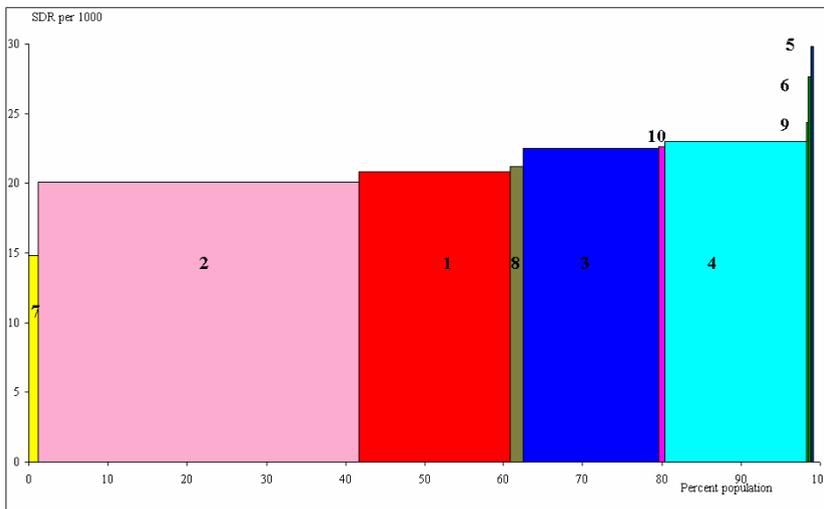
Figure 7: Global clustering for the four periods combined, according to 10 clusters. Males



The other four clusters, by contrast, comprise numerous units, which, again, present strong continuities. Two show a very clear split between the European (cluster 4) and Asian (cluster 3) parts of the second cluster of the 2-cluster map, adhering closely to the Ural frontier. The disjunction is less clear for the subdivision in the first one, but roughly-speaking, it counterposes the southern (cluster 2) and central (cluster 1) parts of European Russia, even if Moscow and St Petersburg are associated with the southern part.

These 4 main clusters therefore concentrate a very large part of the total Russian population. Figure 8 distributes the male Russian population (mean population for the entire interval between the first and last period) among the 10 clusters, ranked by SDR level (four-period mean). While the lowest and highest SDR levels are observed in some extreme small clusters, the four main clusters vary closer to the mean. They differ more by disease profile than overall mortality level. Figure 8 illustrates very clearly their dominant demographic weight.

Figure 8: Distribution of male Russian population by the 10 clusters, ranked by level of SDR

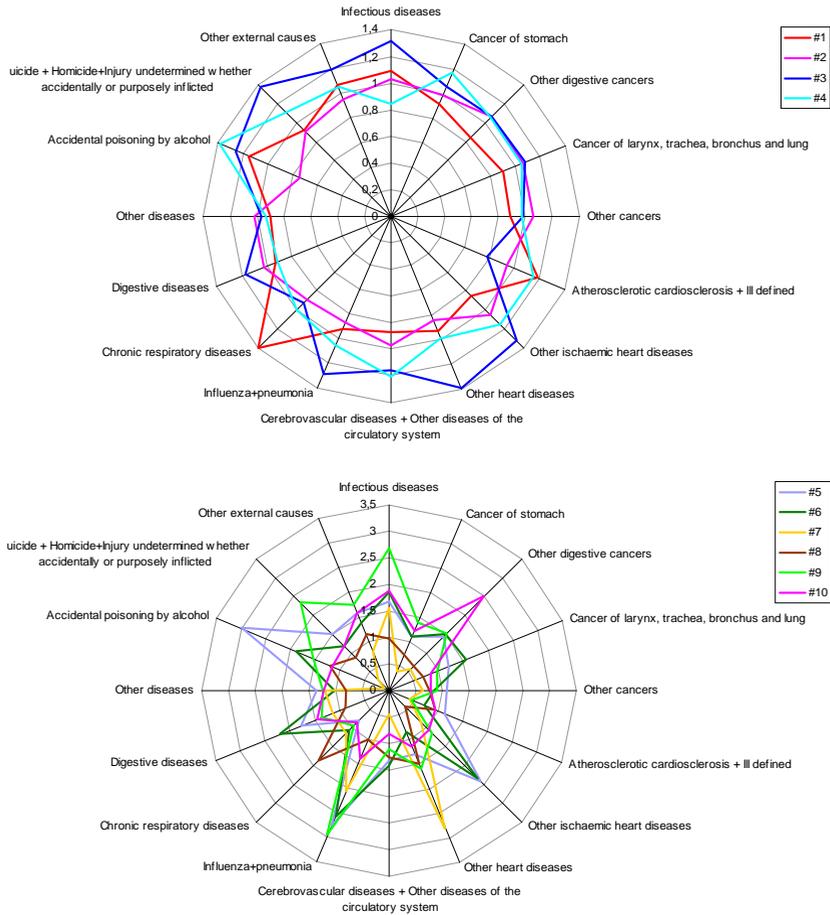


The disease profiles, illustrated by the stars in Figure 9, show that the excess mortality from chronic respiratory diseases observed in the final cluster 1 relates only to the central part of European Russia. The northern Caucasus and the units along the Kazakhstan border are not concerned at all. Conversely, the lower mortality from accidental alcohol poisoning characterizes only southern, not central European Russia, where it is slightly above the Russian mean. For these two disease groups, therefore, the South is much better placed than central European Russia. The reverse is true for cancers, which cause higher mortality in the South than in central European Russia, which probably explains Moscow and St Petersburg's association to that cluster.

The difference between the northern part of European Russia and most of Siberia is related to both overall mortality levels and disease profile. For most causes of death, mortality is higher in the eastern than western Ural. But there are also marked differences in terms of specific causes of death. First, its lower general mortality notwithstanding, the European side presents higher mortalities for accidental alcohol poisoning and stomach cancer. Second, mortality is particularly high in the Asian part for violent deaths (suicide, homicide and other accidental deaths), infectious diseases and acute respiratory diseases. And finally, mortality from heart diseases (ischaemic and others) is also much higher in the Asian than the European part.

The second part of Figure 9 confirms what can be conjectured from period analyses: clusters comprised of one or two geographic units only are highly singular, with sharply contrasting disease profiles. Often, just one or two causes of death will give a very clearly-defined star. What is less easy is to distinguish what is due to true differences from what stems from insufficiently robust data. Continuity here stems more from the individual units concerned being almost invariably the same than from being same-cause related. To comment further would require a forensic examination of each particular context, which is beyond the scope of this paper.

Figure 9. Cause-of-death patterns of each global cluster compared to the all-Russia cluster (cluster SMR/Russian SMR). Ten clusters, males



5. Four main clusters explanatory of how geographical contrasts contribute to general mortality dynamics

To determine the extent to which geographical contrasts contribute to general Russian mortality dynamics, we focused on the four main clusters identified in the previous section. For simplicity, these clusters are named as:

cluster 1 (red): Chernoziom,
cluster 2 (pink): South,
cluster 3 (dark blue): Siberia,
cluster 4 (light blue): European North.

5.1 Mortality dynamics by clusters

For each of these four clusters, we calculated the SDR for the 16 groups of causes at each period. After merging various groups with strong similarities, Table 5 gives the results for 12 new groups.

At the level of all-cause mortality, the dynamics of the four clusters are quite similar (Figure 10). Differences between clusters are mostly a matter of level: the Chernoziom cluster shows a trajectory that is almost identical to that for all-Russia. The trajectories of the other three clusters are almost the same, but at a lower level for the South cluster and a higher level for Siberia and the European North¹⁸. Significantly, however, with the 1988-89 mortality decline due to the anti-alcohol campaign, the differences between clusters narrowed. In particular, the European North and South approached the all-Russia level, at a time when the European North and Siberia were almost at the same level. The inter-cluster distance also increased somewhat from 1989 to 1994, as the mortality increase was slightly steeper for clusters 3 and 4 than for clusters 1 and 2.

¹⁸ Only one cluster is below the all-Russia level, while three are equal to or above it. This is due to the fact that South, above the mean, is more heavily populated than the other three clusters.

Table 5: Male SDR (per 1000) by cause of death for the 4 main clusters and 4 periods

Cluster	Groups of causes												
	Total	1+10	2	3-5	6	7	8	9	11	12+13	14	15	16
1969-70													
1	19.28	1.11	1.01	1.61	4.91	0.65	0.54	3.35	2.64	1.01	0.34	0.73	1.37
2	18.56	0.96	1.14	2.24	3.97	1.14	0.56	3.80	1.54	1.14	0.21	0.66	1.21
3	21.11	1.30	1.17	2.14	4.41	0.99	0.87	4.23	1.61	1.18	0.36	1.11	1.74
4	21.71	0.95	1.33	2.02	5.56	0.98	0.71	4.67	1.86	1.02	0.33	0.94	1.35
Russia	19.33	1.01	1.12	1.99	4.47	0.98	0.63	3.82	1.83	1.11	0.28	0.78	1.32
1978-79													
1	20.97	0.73	0.76	1.83	5.29	1.38	0.55	3.88	2.48	1.04	0.47	0.93	1.64
2	19.84	0.65	0.83	2.35	4.19	1.99	0.42	4.18	1.58	1.14	0.26	0.83	1.43
3	22.39	0.95	0.91	2.35	3.55	2.31	0.79	5.10	1.56	1.15	0.52	1.37	1.83
4	23.18	0.62	0.99	2.25	5.78	1.91	0.48	5.18	1.69	1.06	0.53	1.08	1.61
Russia	20.75	0.70	0.86	2.20	4.60	1.79	0.53	4.30	1.77	1.14	0.39	0.96	1.52
1988-89													
1	18.95	0.38	0.63	2.45	4.40	1.98	0.50	3.75	1.67	1.03	0.18	0.73	1.25
2	18.42	0.37	0.65	2.83	3.44	2.18	0.42	4.32	1.12	1.14	0.12	0.74	1.09
3	19.88	0.52	0.71	2.83	2.36	2.88	0.73	5.14	1.15	1.17	0.14	1.04	1.22
4	20.13	0.36	0.81	2.82	3.47	2.77	0.40	5.18	1.15	1.01	0.19	0.81	1.16
Russia	18.79	0.39	0.68	2.75	3.53	2.16	0.50	4.41	1.21	1.12	0.14	0.79	1.12
1993-94													
1	24.20	0.57	0.58	2.69	5.62	2.71	0.80	4.38	1.77	1.34	0.55	1.46	1.72
2	23.50	0.65	0.59	2.94	4.34	3.00	0.77	4.90	1.24	1.53	0.41	1.56	1.56
3	26.59	0.84	0.65	2.93	2.95	4.29	1.19	5.94	1.36	1.60	0.66	2.31	1.86
4	26.98	0.65	0.70	3.00	4.87	3.51	0.94	6.19	1.33	1.38	0.82	1.83	1.76
Russia	24.45	0.66	0.61	2.87	4.47	3.04	0.91	5.17	1.34	1.49	0.54	1.72	1.64

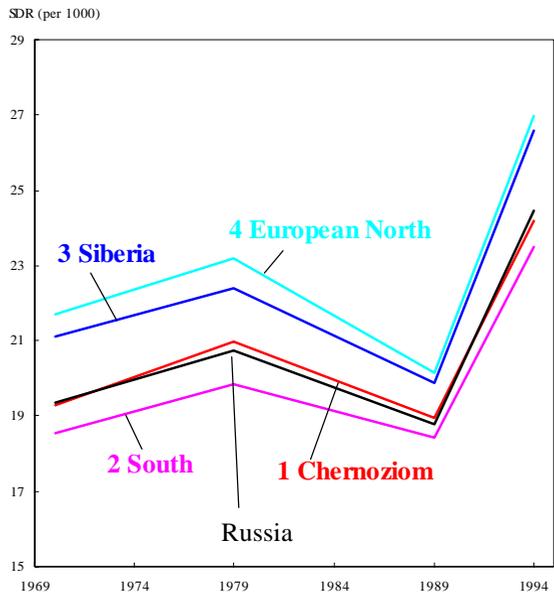
Groups of causes:

- 1+10: Infectious diseases and influenza and pneumonia
 2: Stomach cancer
 3+4+5: Other cancers
 6: Atherosclerotic cardiovascular disease + Ill-defined
 7: Ischaemic heart diseases
 8: Other heart diseases
 9: Cerebrovascular diseases and other diseases of the circulatory system
 11: Chronic respiratory diseases
 12+13 Other diseases
 14: Accidental alcohol poisoning
 15: Suicide, homicide, and injury unspecified whether accidentally or purposely inflicted
 16: Other external causes

Clusters

- 1: (Red) Chernoziom
 2: (Pink) South
 3: (Dark blue) Siberia
 4: (Light blue) European North

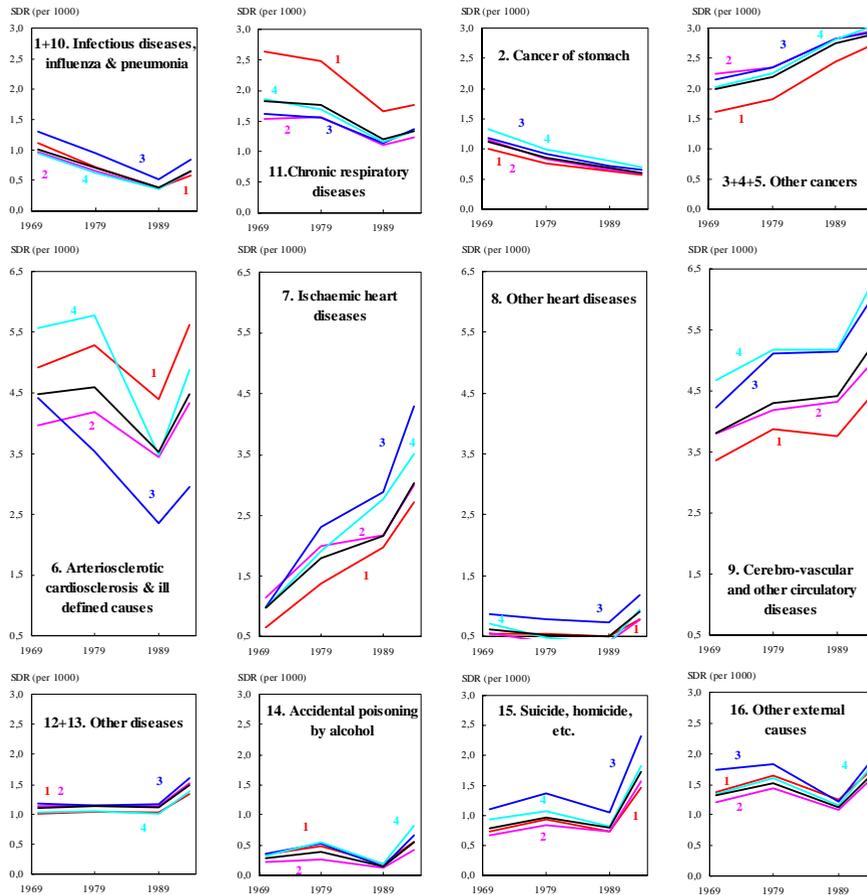
Figure 10: Change in male all-cause SDR in the 4 main clusters, compared to Russia



Almost the same phenomenon is observable here as when Russia is compared to other European republics of the former USSR, like Ukraine or the Baltic countries (Meslé and Shkolnikov, 2000; Meslé and Vallin, 2003; Meslé, 2004): so powerful and generalized were the impacts of first Gorbachev's anti-alcohol campaign, and then the socio-economic crisis due to the abrupt transition to a market economy, as to dominate all the European Republics of the former USSR and all the main Russian regions in comparable fashion. Only slight differences can be identified around these very dramatic changes in general mortality dynamics.

However, greater differences appear when causes of death are taken into account (Figure 11). In fact, causes of death add three types of differentiation that do not appear at the level of all-cause mortality. First, some causes do not at all follow the big changes that are typical of the general dynamic. Second, some clusters are more heavily influenced than others by specific groups of causes. And, finally, some clusters show very specific mortality schedules for some groups of causes.

Figure 11: Changes in SDR by 12 groups of causes in the 4 main clusters, compared to Russia



1= Chernoziom
 2=South
 3 = Siberia
 4 = European North

In the first category, cancers appear highly specific, in that changes are wholly independent of the general scheme, either as steady, across-the-board decreases (stomach cancers) or increases (other cancers) over the review period. The trends in infectious diseases (including influenza and pneumonia), “other heart diseases” and all “other diseases” present no major inter-cluster differentiation, and are quite independent from the anti-alcohol campaign but markedly affected by the 1993-94 crisis.

Several notable examples can be seen of clusters demonstrating particular all-period effects of specific groups of causes. The Chernoziom cluster, for instance, which is the closest to the all-Russia level of all-cause mortality (Figure 10), shows much higher chronic respiratory diseases mortality than the all-Russia level in all four periods, and that mortality is continuously slightly below the mean in all the other three clusters. By contrast, the same Chernoziom cluster is consistently well below Russia for other cancers, while very slight differences are to be found between the other clusters and all-Russia. A third geographical peculiarity of that type affects the Siberia cluster, where the suicide and homicide SDR is constantly well above the rest of the country including the European North cluster, the total mortality of which was higher on Figure 10.

Finally, different mortality schedules are shown for specific groups of causes. For example, while South and Chernoziom clusters track the general schedule for “Atherosclerotic cardiosclerosis & ill-defined causes” quite closely, the European North and Siberia change quite differently. In the European North, mortality decline during the anti-alcohol campaign is much more spectacular than elsewhere. What is remarkable for Siberia is that mortality from that group of causes decreased exceptionally between 1969-70 and 1979-80 while it was increasing everywhere else. Some of these abnormalities may be explained by reverse observations made about ischaemic heart diseases, however. In Siberia, mortality from that group of causes rose far more than elsewhere in the first time interval and continued to increase rapidly in the second, while the rate of increase slowed in the South and Chernoziom. Meanwhile, in the European North that mortality has been rising steadily in both intervals. However, the ischaemic heart disease-related SDR was significantly lower at the time than that for the previous group of causes, especially in the European North. It is very probable that in that part of Russia, a large proportion of mortality from the large ill-defined category including “atherosclerotic cardiosclerosis” is alcohol-related, and that the anti-alcohol campaign had a greater impact in that cluster than elsewhere. The Siberia cluster's situation is probably more related to a general improvement in cause of death registration, which resulted in a general reduction in the proportion of ill-defined causes over the whole period.

The Siberia cluster also presents an atypical schedule, where the SDR from “other external causes” was much higher than elsewhere in 1969-70, but much closer to the

mean in 1988-89 and in 1994-95. It may be that this part of Russia was more affected by these causes during the rapid industrialization years, returning to the mean thereafter.

The main point here is to clarify the impact of the two big events that affected mortality trends in the past 30 years: the anti-alcohol campaign and the socio-economic crisis. Three main features can be singled out. First is the highly differential response of causes of death. For some (cancers), the trends appeared quite independent of both events, while others (infectious diseases, other diseases, other heart diseases) are affected by the socio-economic crisis only. Yet others (violent deaths, most circulatory diseases) are dramatically affected by both events.

Those causes of death whose trends are less affected by these events are also those that display less geographical variation. This is particularly so for stomach cancer. By contrast, geographical variations are wider for circulatory diseases and violent deaths, and for the latter in particular, geographical distance increases as mortality rises and decreases as mortality declines. More specifically, the 1993-94 socio-economic crisis exacerbated geographical inequalities in terms of suicide, homicide and other external causes.

However, differentiated trends, for most causes of death notwithstanding, the hierarchy between the four clusters remained substantially unchanged, producing a clearly constant inter-cluster hierarchy at the level of total mortality, as shown in Figure 10.

5.2 Contribution of geographical dynamics to all-Russia mortality changes

Is there a discernible relation between the geographical clustering of mortality by cause of death, and the observed changes in Russian mortality? We sought to answer that final question by decomposing changes in the Russian SDR observed over the three time intervals studied into the specific contributions of each cluster. Table 6 presents the results for the 10 clusters in terms of total mortality. From 1969-70 to 1978-79 Russian¹⁹ male SDR increased by 142 per 100000. Then, from 1978-79 to 1988-89, it decreased by 199, and finally, from 1988-89 to 1993-94, it increased again by 594. *Prima facie*, a very large share of these changes was always attributable to the 4 first clusters alone. This is mainly due to the fact that these 4 clusters concentrate the overwhelming majority of the total mean population (95.6%). It can also be seen that cluster 2, which accounts for 41% of the total population, consistently dominates the all-Russia changes: it accounts for more than 30% of the total change at each period, and even close to 40% in the first and third periods. But it is also clear that the three

¹⁹ In fact, only 72 regions are considered here as data for Chechnya are still not available.

other main clusters play different roles from period to period, despite having population weights of the same magnitude. For example, cluster 1 accounts for 25% of the Russian mortality increase in the first period, cluster 3 only 15% and cluster 4 under 20%. By contrast, cluster 4 accounts for 26% of the mortality decrease in the second period, while clusters 1 and 3 account for less than 20% each. So the first period deterioration was more attributable to lower than higher mortality regions, while, conversely, higher mortality regions contributed more to the mortality decline during the anti-alcohol campaign. Finally, no such difference appeared during the third period of the huge mortality increase due to the socio-economic crisis, which affected all Russian territory more or less equally.

Table 6: Contribution of each of the 10 clusters to the changes in Russian SDR for three periods

Clusters	Absolute changes in SDR (per 100000)			Proportion of the total change (%)			Percent of population
	1978-79 - 1969-70	1988-89 - 1978-79	1993-94 - 1988-89	1978-79 - 1969-70	1988-89 - 1978-79	1993-94 - 1988-89	
Russia	142.4	-196.3	565.9				
72 regions	141.5	-198.7	594.9	100.0	100.0	100.0	100.0
1	35.4	-38.9	103.9	25.0	19.6	17.5	19.4
2	52.9	-62.3	227.5	37.4	31.4	38.2	40.8
3	21.2	-37.5	106.5	15.0	18.9	17.9	17.3
4	27.2	-51.8	118.3	19.2	26.1	19.9	18.1
5	-0.1	-1.1	1.7	-0.1	0.6	0.3	0.3
6	0.1	-1.2	2.1	0.1	0.6	0.3	0.4
7	1.7	-0.5	2.4	1.2	0.2	0.4	1.2
8	2.9	-7.2	6.7	2.0	3.6	1.1	1.6
9	0.6	-0.7	1.2	0.4	0.3	0.2	0.2
10	0.3	-1.8	2.7	0.2	0.9	0.5	0.7
Changes in the population distribution	-0.5	4.3	22.0	-0.4	-2.2	3.7	

Table 7 examines in detail the contributions by causes of death of each of the 4 main clusters. Major contributions are shown in bold type (black if positive, red if negative). During the first period, cardiovascular mortality is responsible for the biggest share of the contribution to the rise in national mortality in all clusters. However, that contribution differs by cluster. Cluster 2 is obviously the greatest contributor because of its population size. But, cardiovascular mortality is also the reason why cluster 1's contribution to the overall mortality increase is greater than that of clusters 3 and 4. Intriguingly, cluster 3 displays a major counterbalance between “atherosclerotic cardiosclerosis and ill-defined causes” and both ischaemic heart diseases and cerebrovascular diseases, reflecting the big improvement in cause of death registration occurring in that cluster, which contains most of Asian Russia, at the time.

During the second period, “atherosclerotic cardiosclerosis and ill-defined causes” were responsible for the major contributions by clusters to falling Russian mortality, reflecting the fact that the anti-alcohol campaign coincided with a general improvement in cause-of-death registration. This improvement was more significant in the Northern part of European Russia (cluster 4), which thus contributes more than clusters 1 and 3 to the total mortality decline notwithstanding the comparable population size of these clusters.

In the third period, almost every group of causes in all four clusters contributes to the general increase in total Russian mortality. Barring stomach cancer, which continues to play a positive role, all causes and clusters deteriorate to an equal extent with the economic and social crisis.

6. Conclusion

The complex geography of mortality and causes of death, and their time-dependent variations in a large and heterogeneous country like Russia, make it very difficult to advance simple, firm conclusions for any single period, and even more so for a global overview of the four periods analysed here. Arguably, however, at least three interesting conclusions can be drawn from the present attempt.

Firstly, while there may be no clear and complete coincidence between geographical variations in total mortality and cause-of-death levels for the territory of Russia as a whole, there is in European Russia a clear partition between south-west and north-east, both in terms of total mortality and cause-of-death patterns.

Secondly, when analysing global cause-of-death patterns for all periods combined, that contrast is clearly confirmed at whole-country level by the prolongation of the southern part of European Russia by the continuation of the chernoziom belt along the

Table 7: Contribution of causes of death of each of the 4 main clusters to the changes in Russian SDR for three periods (percent)

	1978-79 - 1969-70				1988-89 - 1978-79				1993-94 - 1988-89			
	1	2	3	4	1	2	3	4	1	2	3	4
Total	25.0	37.4	15.0	19.2	19.6	31.4	18.9	26.1	17.5	38.2	17.9	19.9
of which, by cause of death												
Infectious diseases	-3.4	-5.3	-2.7	-2.6	2.0	3.4	1.7	1.3	0.2	1.0	0.4	0.2
Stomach cancer	-3.6	-8.9	-3.1	-4.4	1.2	3.9	1.5	1.6	-0.2	-0.5	-0.2	-0.3
Other digestive cancers	-0.1	-0.2	0.0	0.1	-2.1	-3.2	-1.3	-1.5	0.2	0.3	0.1	0.2
Cancer of larynx, trachea, bronchus and lung	2.4	2.9	1.9	2.3	-2.8	-5.0	-1.9	-2.3	0.4	0.1	0.0	0.2
Other cancers	0.9	0.7	0.5	0.7	-1.1	-2.4	-0.5	-1.0	0.2	0.4	0.1	0.2
Atherosclerotic cardiosclerosis + Ill-defined	5.6	6.3	-10.2	2.9	8.6	16.4	8.9	19.7	4.1	6.7	1.6	4.1
Other ischaemic heart diseases	10.7	24.6	15.5	12.2	-5.9	-4.2	-4.3	-7.3	2.4	6.2	3.8	2.1
Other heart diseases	0.1	-4.1	-0.9	-3.1	0.5	-0.1	0.5	0.7	1.0	2.6	1.2	1.6
Cerebrovascular diseases + Other diseases of the circulatory system	7.9	11.2	10.2	6.6	1.2	-3.1	-0.3	0.0	2.1	4.4	2.1	2.9
Influenza+pneumonia	-2.3	-3.7	-1.3	-1.7	1.4	2.7	1.5	1.0	0.4	1.0	0.5	0.6
Chronic respiratory diseases	-2.5	1.0	-0.6	-2.2	7.8	10.1	3.1	4.6	0.4	0.9	0.6	0.5
Digestive diseases	0.8	0.3	0.3	0.0	0.2	0.3	0.0	0.2	0.5	1.3	0.6	0.5
Other diseases	-0.4	-0.2	-0.7	0.6	-0.1	-0.2	-0.1	0.2	0.6	1.6	0.6	0.6
Accidental alcohol poisoning	2.0	1.4	1.9	2.7	2.8	3.0	2.9	2.9	1.2	2.2	1.4	1.8
Suicide + Homicide+Injury unspecified	2.9	4.9	3.0	1.8	1.9	2.1	2.5	2.3	2.4	6.2	3.4	2.9
Other external causes	4.0	6.5	1.1	3.4	3.8	7.7	4.6	3.8	1.5	3.6	1.7	1.7

Kazakhstan border. At the same time, however, a clear split is also made by a second frontier, that of Ural, between Europe and Asia, for the north-eastern part of Russia. With the exception of the chernoziom belt, therefore, Siberia appears as a very different world from European Russia.

Thirdly, looking at relations between the global cause-of-death patterns (Figure 7) and that for each of the four periods (Figure 3), the period 1988-1989 shows the closest similarities with the global view, while the three other periods are more peculiar. It is true that the period 1988-89, which is that of the lowest variance in total mortality, is also characterized by the highest life expectancy, due to the transient decline in alcoholism, while cardiovascular diseases had already risen to prominence in the Russian disease profile. The other three periods have their individual characteristics. The first remains a transition period in which infectious diseases still play an important role in the Russian disease profile and above all its geographical variations. In the second period, mortality is higher and much more tied to cardiovascular diseases and alcoholism, but geographical differences in total mortality are less important. The geographical contours become closer to that of the period 1988-89, especially with the clear split made by Ural. By contrast, the profound economic and social crisis of 1993-94 makes the geography of cause-of-death patterns of this final period again highly discrete. So, 1988-89 appears to be a base period, and thus fairly typical of the whole period. However, that global period marked by such wide fluctuations may itself be a singular one. Once the 2002 census results become available, it will be interesting to see whether the geographical pattern of 1988-89 remains a robust benchmark or whether the more recent trends in Russian life expectancy will result in a new geographical cause-of-death pattern.

7. Acknowledgements

This paper is the product of collaborative work done under INTAS project n° 1722 “Expectation of life and causes of death in different republics of the ex-USSR: long term trends and recent changes”.

References

- Andreev, Evgueni M., 1979. – Prodolzhitel'nost' zhizni v SSSR: differentsialniy analiz. [Length of life in the USSR: a differential analysis], in : Evgueni Andreev and Anatoli Vishnevski (ed.), *Prodolzhitel'nost' zhizni: analiz i modelirovaniye*. [Length of life: analysis and modelling]. – Moskow, Statistika, 136 p.
- Andreev, Evgeni and Kvasha Ekaterina, 2002. – Osobennosti pokazateley mladencheskoy smertnosti v Rossii. [Peculiarity of infant mortality indicators in Russia], *Problemy sotsial'noy gigiyeny, zdravookhraneniya i istorii meditsiny*. [Problems of social hygiene, public health services and history of medicine], n° 4, p. 15-20.
- Gavrilova Natalia S., Evdokushkina Galina N., Semyonova Victoria G., and Gavrillov Leonid A., 2001. – *Economic crises, stress and mortality in Russia*. – Chicago, Center on Aging, University of Chicago, 28 p. (Paper presented at the 2001 PAA meeting, Washington).
- Goskomstat, 2002. – *The demographic yearbook of Russia. 2002: Statistical Handbook*. – Moscow, Goskomstat, 397 p.
- Jozan Peter and Prokhorskas Remigijus (ed.), 1997. – *Atlas of leading and avoidable causes of death in countries of Central Eastern Europe*. – Budapest, Hungarian CSO Publishing House, Hungarian Central Statistical Office, World Health Organisation, 324 p.
- Leon David A., Chenet Laurent, Shkolnikov Vladimir M., Zakharov Sergei, Shapiro Judith, Rakhmanova Galina, Vassin Sergei, and McKee Martin, 1997. – Huge variation in Russian mortality rates 1984-94: artefact, alcohol, or what?, *The Lancet*, vol. 350, August 9, p. 383-388.
- Meslé France, 2004. – Mortality in Eastern Europe and the former Soviet Union : long-term trends and recent upturns, *Demographic Research*, p. 46-70. (Special Collection 2. Determinants of Diverging Trends in Mortality).
- Meslé France and Shkolnikov Vladimir, 2000. – Russie: une crise sanitaire sans précédents, *Espace, Populations, Sociétés*, n° 2, p. 253-272.
- Meslé France, Shkolnikov Vladimir, Hertrich Véronique, and Vallin Jacques, 1996. – *Tendances récentes de la mortalité par cause en Russie, 1965-1994*. – Paris, INED, 140 p. + 2 floppy disks. (Données statistiques n°2).
- Meslé France, Shkolnikov Vladimir, and Vallin Jacques, 1994. – Brusque montée des morts violentes en Russie, *Population*, vol. 49, n° 3, p. 780-790.

- Meslé France and Vallin Jacques, 2003. – *Mortalité et causes de décès en Ukraine au XXe siècle*. – Paris, INED, xvi + 396 p. (Les cahiers de l'INED, cahier n° 152, with contributions by Vladimir Shkolnikov, Serhii Pyrozhkov and Serguei Adamets).
- Meslé France, Vallin Jacques, Hertrich Véronique, Andreev Evgueni, and Shkolnikov Vladimir, 2003. – Causes of death in Russia: assessing trends since the 1950s, in : Irena E. Kotowska and Janina Józwiak (ed.), *Population of Central and Eastern Europe. Challenges and opportunities*, p. 389-414. – Warsaw, Statistical Publishing Establishment, 724 p.
- Meslé France, Vallin Jacques, and Shkolnikov Vladimir, 1998. – Reversal of mortality decline: the case of contemporary Russia, *World Health Statistics Quarterly/Rapport trimestriel de statistiques sanitaires*, vol. 51, n° 2-3-4, p. 191-206. (“Historical epidemiology: mortality decline, and old and new transitions in health”, special issue edited by Odile Frank).
- Shkolnikov Vladimir M., 1987. – Geograficheskiye faktori prodoljitelnosti jizni. [Geographical factors of length of life]. *Izvestiya AN SSSR, Seriya Geograficheskaya*, vol. 3, n° 12s, p. 35-44. [Academy of Sciences of the SSSR].
- Shkolnikov Vladimir, Cornia Giovanni, Leon David, and Meslé France, 1998. – Causes of the Russian mortality crisis: evidence and interpretations, *World Development*, vol. 26, n° 11, p. 1995-2011.
- Shkolnikov Vladimir, Meslé France, and Vallin Jacques, 1996. – Health crisis in Russia. I. Recent trends in life expectancy and causes of death from 1970 to 1993, II. Changes in causes of death: a comparison with France and England and Wales (1970 to 1993), *Population, An English Selection*, vol. 8, p. 123-190.
- Shkolnikov Vladimir and Nemtsov Alexander, 1997. – The anti-alcohol campaign and variations in Russian mortality, in: Jose Luis Bobadilla, Christine A. Costello, and Faith Mitchell (eds), *Premature death in the New Independent States*, p. 239-261. – Washington D.C, National Academy Press, 404 p.
- Shkolnikov, Vladimir, Valkonen Tapani, and Begun Alexander, 2001. – Measuring inter-groups inequality in length of life, *Genus*. Vol. 57, n° 3-4, p. 33-62.
- Shkolnikov Vladimir M. and Vassin Sergei A., 1994. – Spatial differences in life expectancy in European Russia in the 1980s, in : Wolfgang Lutz, Sergei Scherbov, and Andrei Volkov (éd.), *Demographic trends and patterns in the Soviet Union before 1991*, p. 379-402. – London and New York, Routledge, 496 p.

SPSS, 2002. – SPSS 11.5 Syntax Reference Guide. – Chicago, SPSS Inc.

Vassin Sergei A. and Costello Christine A., 1997. – Spatial, age, and cause-of-death patterns of mortality in Russia, 1988-1989 , in : José Luis Bodabilla, Christine A. Costello, and Faith Mitchell (éd.), *Premature death in the new independent states* , p. 66-119. – Washington (DC), National Academy Press, 404 p.

Walberg Peder, McKee Martin, Shkolnikov Vladimir M, Chenet Laurent, and Leon David, 1998. – Economic change, crime and Russian mortality crisis: a regional analysis, *British Medical Journal*, vol. 317, August, p. 312-318.

WHO, 1992. – *World Health Statistic Annual, 1992*. – Geneva, World Health Organisation, 400 p.

Annex I. Administrative units of the Russian Federation

Table I-1: List of the 73 administrative units in alphabetical order

Altai Kray & Republic of Altai	71	Lipetsk Oblast	38	Republic of Tatarstan	42
Amur Oblast	96	Magadan Oblast & Chukchi autonomous district	99	Republic of Tuva	81
Arkhangelsk Oblast	5	Moscow	20	Rostov Oblast	59
Astrakhan Oblast	43	Moscow Oblast	21	Ryazan Oblast	23
Belgorod Oblast	35	Murmansk Oblast	8	Sakhalin Oblast	100
Bryansk Oblast	15	Nizhny Novgorod Oblast	33	Samara Oblast	46
Chechen & Ingush Republics	56	Novgorod Oblast	12	Saratov Oblast	47
Chelyabinsk Oblast	68	Novosibirsk Oblast	73	Smolensk Oblast	24
Chita Oblast	88	Omsk Oblast	74	St. Petersburg	10
Chuvash Republic	31	Orenburg Oblast	64	Stavropol Kray & Karachaev-	58
Irkutsk Oblast	86	Oryol Oblast	22	Circassian republic	
Ivanovo Oblast	17	Penza Oblast	45	Sverdlovsk Oblast	67
Kabardin-Balkar Republic	53	Perm Oblast	65	Tambov Oblast	39
Kaliningrad Oblast	101	Primorsky Kray	94	Tomsk Oblast	75
Kaluga Oblast	18	Pskov Oblast	13	Tula Oblast	26
Kamchatka Oblast	97	Republic of Bashkortostan	61	Tver Oblast	25
Kemerovo Oblast	72	Republic of Buryatia	80	Tyumen Oblast	76
Khabarovsk Kray & Jewish autonomous oblast	95	Republic of Dagestan	51	Udmurt Republic	62
Kirov Oblast	32	R. of Kalmykia - Khalmg Tangch	41	Ulyanovsk Oblast	48
Kostroma Oblast	19	Republic of Karelia	3	Vladimir Oblast	16
Krasnodar kray & R. of Adygeya	57	Republic of Komi	4	Volgograd Oblast	44
Krasnoyarsk Kray & Republic of Khakasia	83	Republic of Mari El	29	Vologda Oblast	7
Kurgan Oblast	63	Republic of Mordovia	30	Voronezh Oblast	36
Kursk Oblast	37	Republic of North Ossetia	55	Yaroslavl Oblast	27
Leningrad Oblast	11	Republic of Sakha (Yakutia)	91		

Table I-2: List of the 73 administrative units in numerical order

3	Republic of Karelia	32	Kirov Oblast	64	Orenburg Oblast
4	Republic of Komi	33	Nizhny Novgorod Oblast	65	Perm Oblast
5	Arkhangelsk Oblast	35	Belgorod Oblast	67	Sverdlovsk Oblast
7	Vologda Oblast	36	Voronezh Oblast	68	Chelyabinsk Oblast
8	Murmansk Oblast	37	Kursk Oblast	71	Altai Kray & Republic of Altai
10	St. Petersburg	38	Lipetsk Oblast	72	Kemerovo Oblast
11	Leningrad Oblast	39	Tambov Oblast	73	Novosibirsk Oblast
12	Novgorod Oblast	41	Republic of Kalmykia - Kholm Tangch	74	Omsk Oblast
13	Pskov Oblast	42	Republic of Tatarstan	75	Tomsk Oblast
15	Bryansk Oblast	43	Astrakhan Oblast	76	Tyumen Oblast
16	Vladimir Oblast	44	Volgograd Oblast	80	Republic of Buryatia
17	Ivanovo Oblast	45	Penza Oblast	81	Republic of Tuva
18	Kaluga Oblast	46	Samara Oblast	83	Krasnoyarsk Kray & Republic of Khakasia
19	Kostroma Oblast	47	Saratov Oblast	86	Irkutsk Oblast
20	Moscow	48	Ulyanovsk Oblast	88	Chita Oblast
21	Moscow Oblast	51	Republic of Dagestan	91	Republic of Sakha (Yakutia)
22	Oryol Oblast	53	Kabardin-Balkar Republic	94	Primorsky Kray
23	Ryazan Oblast	55	Republic of North Ossetia	95	Khabarovsk Kray & Jewish autonomous oblast
24	Smolensk Oblast	56	Chechen & Ingush Republics	96	Amur Oblast
25	Tver Oblast	57	Krasnodar kray & Republic of Adygeya	97	Kamchatka Oblast
26	Tula Oblast	58	Stavropol Kray & Karachaev- Circassian rep.	99	Magadan Oblast & Chukchi autonomous district
27	Yaroslavl Oblast	59	Rostov Oblast	100	Sakhalin Oblast
29	Republic of Mari El	61	Republic of Bashkortostan	101	Kaliningrad Oblast
30	Republic of Mordovia	62	Udmurt Republic		
31	Chuvash Republic	63	Kurgan Oblast		

Figure annex: Map of administrative units used for the geographical analysis.
Numbers are listed above.



Solid lines are for units used, dotted lines are for autonomous territories (not distinguished for the analysis)

Annex II. The choice of colour sets for period-specific maps

To make the maps and figures as understandable as possible, we endeavoured to use colors whose signification would remain near-constant even if the borders of clusters that they cover varied. The color of each global cluster (for the four periods simultaneously, Section III) was selected on the basis of the dendrogram produced by the clustering process, and was chosen from the last step of the process. Two contrasting colors (blue and red) were attributed to the two large clusters obtained at the end (Figure 5 above). Then, at every reverse step towards a greater numbers of clusters, one cluster is subdivided into two. We kept the former color for the larger, and chose a new but allied color for the smaller. Finally, we selected 10 colours for the global clusters.

When repeating the analysis separately for each period, the geographic contents of clusters clearly vary. But most share many characteristics with clusters of the global analysis, and we aimed to use the same colors to reflect such similarities. To do that, we calculated the distance between period-specific clusters and global clusters in terms of cause-of-death levels and patterns for each period. We first computed the Euclidian distance between each region at one period to global clusters:

$$\left(\sum_i \left((W(r, t, i) - W(R_k, i)) \cdot \sqrt{SDR(0, t, i)} \right)^2 \right)^{\frac{1}{2}},$$

where R_k is the set of the n_k regions included in cluster k and

$$W(R_k, i) = \frac{1}{4n_k} \sum_t \sum_{r \in R_k} W(r, t, i),$$

and, finally, the distance between each period-specific cluster and the global clusters was obtained as the average of the Euclidian distances of regions belonging to the period-specific cluster.

On the basis of these distances, we assigned the color of the nearest global cluster to each period-specific cluster (Table II-1). Note that a period-specific cluster and the nearest global cluster are not necessarily the closest geographically. We also used the same colors to plot the stars representing the cause-of-death patterns (Figures 4, 6, and 9).

In 3 of 4 periods, the same global cluster was the nearest for two period-specific clusters. We addressed this problem by adding a specific pattern to the selected color to maintain the differentiation between the two period-specific clusters. For example, in 1978-1979 the global cluster 1 (red) is the nearest for period-specific clusters 1 and 8, but at this period cluster 8 includes only the Republic of Dagestan, and we attributed the

red color to cluster 1, shadowed by a same-colour pattern for cluster 8, which will be numbered 11, in that case. Conversely, the color of global cluster 8 (orange) is not used for any period-specific cluster, since that global cluster never appeared as the nearest for any period-specific cluster.

Table II-2 gives the number of units for each cluster (either global or period specific), including a column showing the color used for each on the maps. Note that lines 1 and 6 also show the alternative pattern used to identify the additional cluster 11.

Table II-1: Distance between period-specific clusters and global clusters

Period-specific	Global cluster number									
cluster number	1	2	3	4	5	6	7	8	9	10
1970										
1	0.0005	0.0080	0.0197	0.0130	0.0737	0.0850	0.0591	0.0151	0.0757	0.0383
2	0.0123	0.0004	0.0078	0.0089	0.0553	0.0563	0.0652	0.0297	0.0633	0.0378
3	0.0136	0.0074	0.0031	0.0059	0.0371	0.0465	0.0848	0.0237	0.0385	0.0283
4	0.0129	0.0120	0.0144	0.0012	0.0464	0.0606	0.1114	0.0222	0.0737	0.0473
5	0.1117	0.1187	0.1097	0.0811	0.0538	0.1049	0.2808	0.1350	0.1454	0.1180
6	0.1155	0.1039	0.0696	0.0797	0.0483	0.0313	0.2278	0.1034	0.0983	0.1141
7	0.0555	0.0463	0.0633	0.0859	0.1382	0.1367	0.0101	0.0806	0.0996	0.0683
8	0.0092	0.0310	0.0446	0.0288	0.0966	0.1150	0.0858	0.0158	0.1056	0.0611
9	0.0735	0.0615	0.0399	0.0661	0.0758	0.0827	0.0948	0.0682	0.0165	0.0438
10	0.0780	0.0657	0.0501	0.0707	0.0818	0.0820	0.1319	0.0956	0.0254	0.0139
1979										
1	0.0002	0.0093	0.0199	0.0122	0.0715	0.0841	0.0645	0.0142	0.0777	0.0406
2	0.0099	0.0001	0.0085	0.0088	0.0571	0.0594	0.0620	0.0277	0.0653	0.0368
3	0.0248	0.0109	0.0003	0.0105	0.0357	0.0338	0.0885	0.0290	0.0331	0.0365
4	0.0128	0.0111	0.0138	0.0008	0.0424	0.0576	0.1119	0.0239	0.0732	0.0461
5	0.0718	0.0554	0.0384	0.0541	0.0154	0.0213	0.1406	0.0966	0.0688	0.0597
6	0.0699	0.0549	0.0331	0.0406	0.0277	0.0128	0.1657	0.0667	0.0561	0.0720
7	0.0866	0.0810	0.0984	0.1248	0.1909	0.1842	0.0048	0.1046	0.1366	0.1066
8	0.0148	0.0413	0.0525	0.0333	0.0969	0.1230	0.1040	0.0187	0.1140	0.0704
9	0.1322	0.1037	0.0608	0.0994	0.0714	0.0589	0.1902	0.1216	0.0131	0.0736
10	0.0448	0.0470	0.0437	0.0506	0.0634	0.0797	0.1067	0.0732	0.0443	0.0029

Table II-1: (Continued)

Period-specific	Global cluster number									
cluster number	1	2	3	4	5	6	7	8	9	10
1989										
1	0.0003	0.0102	0.0218	0.0133	0.0689	0.0843	0.0661	0.0196	0.0803	0.0388
2	0.0070	0.0002	0.0088	0.0068	0.0556	0.0616	0.0653	0.0244	0.0651	0.0359
3	0.0236	0.0092	0.0006	0.0110	0.0402	0.0338	0.0795	0.0278	0.0382	0.0382
4	0.0144	0.0056	0.0054	0.0016	0.0413	0.0449	0.0933	0.0235	0.0596	0.0423
5	0.1090	0.0731	0.0505	0.0622	0.0304	0.0230	0.2009	0.1129	0.0938	0.0990
6	0.1164	0.0806	0.0599	0.0773	0.0391	0.0154	0.2022	0.1406	0.0772	0.0813
7	0.0639	0.0710	0.0898	0.1087	0.1787	0.1760	0.0037	0.0840	0.1403	0.1005
8	0.0429	0.0413	0.0316	0.0367	0.1021	0.0903	0.0946	0.0091	0.0794	0.0848
9	0.0633	0.0584	0.0330	0.0568	0.0629	0.0646	0.1323	0.0700	0.0097	0.0319
10	0.0270	0.0254	0.0269	0.0320	0.0501	0.0587	0.0842	0.0527	0.0497	0.0040
1994										
1	0.0006	0.0093	0.0241	0.0132	0.0738	0.0875	0.0674	0.0218	0.0885	0.0437
2	0.0070	0.0002	0.0103	0.0084	0.0592	0.0649	0.0620	0.0260	0.0669	0.0359
3	0.0271	0.0122	0.0012	0.0139	0.0348	0.0295	0.0824	0.0331	0.0348	0.0377
4	0.0142	0.0075	0.0066	0.0006	0.0367	0.0439	0.1008	0.0232	0.0602	0.0414
5	0.1202	0.1057	0.0705	0.0918	0.0268	0.0530	0.1992	0.1371	0.0568	0.0880
6	0.1234	0.0879	0.0618	0.0912	0.0465	0.0264	0.1685	0.1406	0.0825	0.0945
7	0.0757	0.0818	0.1048	0.1202	0.2135	0.2077	0.0077	0.0871	0.1687	0.1241
8	0.0730	0.0639	0.0503	0.0605	0.1302	0.1077	0.1161	0.0272	0.0973	0.1126
9	0.1020	0.0866	0.0560	0.0908	0.0788	0.0725	0.1521	0.1164	0.0113	0.0545
10	0.0383	0.0359	0.0432	0.0480	0.0713	0.0841	0.0787	0.0650	0.0622	0.0076

Table II-2: Number of administrative units in the global clusters and in corresponding nearest cluster(s) for specific periods

Global cluster		Number of units in the nearest cluster(s)				
Number	Color	Number of units	1969-70	1978-79	1988-89	1993-94
1	 / 	19	10/1	25/1	13	18
2		18	39	29	18	6
3		11	10	8	19	20
4		17	5	1	16	19
5		1	1	4	-	1
6	 / 	1	1	2	1/2	3
7		1	4	1	1	1
8		2	-	-	1	1
9		1	1	1	1	2
10		1	1	1	1	1
No data		1				1
Total		73	73	73	73	73

Annex III. Why use \sqrt{SDR} as a weight?

Suppose there are two causes with very similar variations across regions. A desirable property of a distance measure would be that such measure does not change if we aggregate the two causes of death into one.

Let $SDR(r, t, 2) = \lambda \cdot SDR(r, t, 1) + C$. Let 1&2 denote the sum of causes 1 and 2. Then $SDR(r, t, 1 \& 2) = (1 + \lambda) \cdot SDR(r, t, 1) + C$ and for any pair of regions r_1 and r_2 :

$$(W(r_2, t, 1) - W(r_1, t, 1)) = (W(r_2, t, 2) - W(r_1, t, 2)) = (W(r_2, t, 1 \& 2) - W(r_1, t, 1 \& 2))$$

It is clear that the sum of simple Euclidian distances $D[W(r_2, t, 1), W(r_1, t, 1)]$ and $D[W(r_2, t, 2), W(r_1, t, 2)]$ will be higher than the distance for the summary cause 1&2 $D[W(r_2, t, 1 \& 2), W(r_1, t, 1 \& 2)]$. However with weights, equal to $\sqrt{SDR(0, t, i)}$, the two distances are equal:

$$(W(r_2, t, 1) - W(r_1, t, 1))^2 SDR(0, t, 1) + (W(r_2, t, 2) - W(r_1, t, 2))^2 SDR(0, t, 2) = (W(r_2, t, 1 \& 2) - W(r_1, t, 1 \& 2))^2 SDR(0, t, 1 \& 2)$$

Annex IV. Life expectancy at birth for regions in each SDR interval
(see footnote 10)

Males	1969-70	1978-79	1988-89	1993-94
SDR intervals				
<12.22	69.80			
12.22-15.66	66.82	66.22	67.42	
15.66 to 19.10	64.43	64.43	64.96	65.25
19.10 to 22.54	61.98	61.55	63.50	61.27
22.54 to 25.98	60.56	59.25	62.94	58.87
25.98 to 29.42	58.60	57.90		56.45
>= 29.42	56.44			48.70

Females	1969-70	1978-79	1988-89	1993-94
SDR intervals				
<6.92	82.62	81.16		
6.92-8.62	77.08	75.43	76.68	
8.62 to 10.32	74.68	74.86	75.29	74.74
10.32 to 12.02	73.27	73.23	74.12	73.12
12.02 to 13.72	70.99	71.15	72.55	71.66
13.72 to 15.43	70.26	68.18	70.20	69.59
>= 15.43		68.21	68.90	58.16

