ABOUT MORTALITY DATA FOR RUSSIA

by Vladimir M. Shkolnikov and Dmitri A. Jdanov

Last Updated: December 20, 2017

GENERAL

In Russia the most detailed data on deaths and population size by single-year age groups have never been published. Between the 1950s and 1973, published statistical annuals on population included only aggregate indices such as the total (all ages combined) numbers of deaths, births, and population, or, sometimes, totals which had been disaggregated into very broad age categories (e.g., children, people of working age, those beyond retirement age).

Between 1974 and 1986, any publication of mortality data was forbidden in the Soviet Union, for ideological reasons. In 1986-87, in the completely new era of Mikhail Gorbachev's 'Glasnost', the Central Statistical Office resumed publication of mortality statistics in demographic yearbooks. These data were much more detailed than those of the 1960s since they included deaths and population by sex and ages 0, 1-4, 5-9, ... 85+. This practice continues now without significant changes.

Since 1946, the Soviet (Russian) statistical system has been producing annual tables (unpublished) of deaths and population by sex and single-year age group.

These statistical tables have been used for various statistical purposes (population projections, life tables etc.), but they have not been published. Demographers had (and have) access to these data by request from the Central Statistical Office of the Russian Federation (currently named "Goskomstat").

Source of Data

Most of the data included in the Human Mortality Database (HMD) were received from E. Andreev and S. Zakharov. The population estimates for the period 1946-59 were made by E. Andreev (Andreev et al., 1998). Birth counts were published in the Demographic Yearbook of Russia (Goskomstat, 2000). The Input Database (InputDB) death and population numbers cover each year since 1946.

TERRITORIAL COVERAGE

The following territorial changes have taken place in Russia during the period covered by the data:

- The Crimea region was excluded from Russia and included in Ukraine in 1956.
- Kaliningrad, Tuva, and Karel provinces were included in Russia after the end of the Great Patriotic war in 1945.

S. Zakharov (see research note, Appendix II) has made adjustments for these territorial changes. All data in the Input Database are for Russia with the Crimea region excluded and with Kaliningrad, Tuva, and Karel provinces included.

In addition, there were some changes in the coverage of vital statistics although they were not official territorial changes (i.e., the national boundaries of Russia did not change). For these changes in population coverage, the methods described in Appendix D of the *Methods Protocol* have been used to make the appropriate adjustments in the formulas for calculation of population estimates and death rates:

Dates	Population Coverage	Area Code†
1959-1992	Vital statistics and official population estimates cover the entire population of Russia.	1
1993-1994	Vital statistics exclude the Chechen-Ingush Republic. Although the official population estimates include this region, these estimates have been adjusted for the purposes of the HMD to exclude the Chechen-Ingush Republic.	11
1995-2002	Vital statistics exclude the Chechen Republic. Again, the official population estimates include this region, but for the purposes of the calculations for this database, these estimates have been adjusted to exclude the Chechen Republic.	12
2003-2015	Vital statistics and official population estimates cover the entire population of Russia.	1

† The area code is used in the raw data files (Input Database) to denote the geographic area covered by the data.

The official Russian population statistics produced by the Federal State Statistics Service have been including the Crimea region and Sevastopol since January 1st, 2015. The current HMD data series ends in 2014 even though vital statistics are available up to 2015 included. To avoid using territorial adjustment factors for the calculation of death rates in 2014, population as of January 1st, 2015 is presented without Crimea and Sevastopol, i.e. with the same territorial coverage as for death counts in 2014

DEATH COUNT DATA

Coverage and Completeness

Registration of deaths in Russia by the centralised state civil registration system was organised in the 1920s. It included the Central Statistical Office, named TCUNHU SSSR/Central Office for the Statistics of the National Economy (later renamed TCSU SSSR/Central Statistical Office of the USSR and then again renamed Goskomstat/State Statistical Committee of the Russian Federation), regional ('oblast') statistical offices, and district statistical bureaux (ZAGS). This system is highly centralised. District statistical bureaus register vital events and send copies of individual records to the regional statistical offices. Regional statistical offices (with the help of standard computer programs since 1988) assemble individual death records and transform them into standard summary tables going to the Central Statistical Office in Moscow. The coverage of the whole population by this system became complete by about the mid-1930s.

The annual statistical table named 'forma 4' (computer table S41 since 1988) includes deaths by sex and single-year age group for Russia as a whole and for each of Russia's oblast-level regions.

For the period between the end of the Second World War and the mid-1960s, there has been a concern in the demographic literature about the quality of infant and old-age mortality values in Russia and other parts of the former Soviet Union (Anderson and Silver, 1986, Shkolnikov, Meslé & Vallin, 1996). For Central Asia and Transcaucasia, the problem was very significant even at the end of the Soviet era in the late 1980s. However, for Russia, the problem was radically reduced by the end of the 1960s. There are some signals from statisticians and researchers

suggesting that in the most recent years, the problem of death under-registration at very old ages might be growing. To the authors' knowledge, however, there is still no scientific evidence on the issue.

A substantial proportion of births with a higher risk of infant death are excluded from live births according to a restricted definition of live birth which results in incomparability with countries using the WHO definition of live birth (see section "Birth Count Data" for more details).

Death statistics for Russia do not include deaths in Chechen-Ingush Republic in 1993-1994 and deaths in the Chechen Republic in 1995-2002.

Specific Details

There is a concern about age misreporting on death certificates mostly at ages 70, 80, and 90 before 1970 (see Appendix II for more details). Significant age heaping and problems with erratic age reporting may exist even now at very old ages. Unfortunately, no systematic individual-level validation of the quality of death records for deaths at old ages has been done.

S. Zakharov has detected a specific problem of death heaping at age 99 (see Appendix II for more details). The authors can only speculate about the nature of the problem. It might be due to some computational procedures by the Goskomstat, which (perhaps) allow exchanges between age 99 and the next age group 100+. To avoid this problem in the estimates made for the HMD, deaths have been aggregated in the open age interval 99+ for the period 1959-1989.

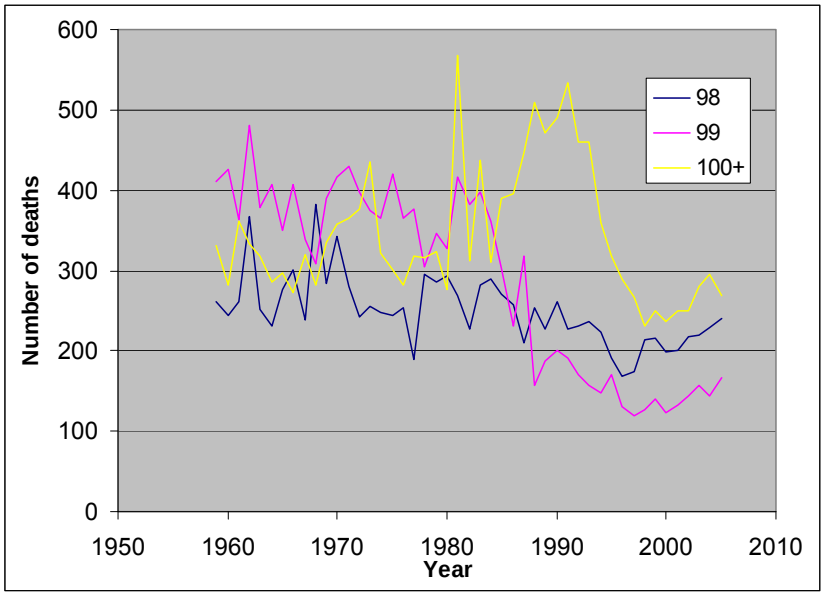


Figure 1: Number of deaths at ages 98, 99 and 100+, Russia, males, 1959-2005.

There is some concern about age overstatement in mortality statistics (see the section relative to Data Quality Issues for details)

POPULATION COUNT DATA

Coverage and Completeness

Population estimates are based on the all-Soviet censuses of 1959, 1970, 1979 and 1989. Censuses were conducted on January 15th in 1959 and 1970, and on January 17th in 1979 and 1989. From the census counts, the Goskomstat produced the official population estimates as of January 1st.

Specific Details

Before the first post-war census of 1959, very little was known about the Russian population. In 1998, Andreev et al. (1998) completed a reconstruction of population by age for the period 1946-58. These estimates are included in the raw data files, but they were not used for the construction of the HMD mortality estimates. In the 1960s, the Central Statistical Office did not produce official estimates of population by age. The all-Soviet census of 1959, the micro-census of 1964, and the census of 1970, along with (incomplete) current data on births, deaths, and migrations provided grounds for an estimation of population. Since the census of 1970, the Central Statistical Office has produced official estimates of population by age on an annual basis from census counts, deaths, births and migrations. After the censuses of 1979 and 1989, the population figures within respective inter-census periods have been re-calculated on the basis of the new census data.

For population counts at advanced ages, just as for death counts at these ages, there is a problem of age heaping and age overstatement (see the Data quality section and Appendix II). In addition, there is a particular problem of over-registration of centenarians (Garson, 1991).

As noted in the section 'Territorial Coverage', deaths in the Chechen Republic were excluded from official statistics of Russia during the period 1993-2002, although official population estimates are available for Russia as a whole (including the Chechen Republic). To avoid inconsistency between the numerator and denominator, population estimates have been used for 1993-2002 which exclude the Chechen-Ingush Republic (1993-1994) and the Chechen Republic (1995-2002). Thus, for this period the estimates do not refer to the whole of Russia. Although these changes in coverage do not represent official territorial changes, they are treated as such in order to make the appropriate adjustments to the formulas.

It should be noted that the adjustment factor for the inclusion/exclusion of the Chechen Republic varies substantially across age. The Chechen Republic comprises less than 1% of the total population of Russia, but at very old ages (100+) the comparable percentage is more than 10%. This indicates a serious age overstatement problem in the Chechen Republic and (perhaps) some other Muslim regions of the North Caucasus (Coale and Kisker, 1986).

BIRTH COUNT DATA

Coverage and Completeness

The registration of live births in Russia (as in many post-Soviet and former Communist countries) differs from conventional Western practices and the WHO recommendations. According to the Soviet definition of a live birth (launched before the Second World War), a live birth is to be officially registered by the statistical system if the gestation period is 28 weeks or longer, the

body mass at birth is 1000g or higher, the body length is 35cm or longer, and the newborn breathes. Such a restrictive rule leads to underestimation of births and population at age 0 and also to underestimation of neonatal mortality by about 50% and infant mortality by about 25% (Anderson and Silver, 1986, Blum and Monnier, 1989, Velkoff and Miller, 1995, Andreev, 1995).

DATA QUALITY ISSUES

Data prior to 1970 should be used with extra caution due to problems of data quality.

1. The problem of death age heaping is still one of the main problems of Russian data, especially for the period before 1970. The authors have not attempted to solve this problem, and leave it open to researchers who would particularly focus on it. Therefore, care should be taken in using the Russian data on number of deaths and mortality rates by one-year age groups (see Appendix II for more details).
2. There are serious inaccuracies in the official estimation of the population of elderly men and women reported by the Goskomstat, especially for persons over age 80. As a result, elderly population changes and, correspondingly, mortality rates, become highly irregular and inaccurate. The older the people, the more problematic the official data. These errors could be of methodological character and, most probably, have to do with use of the wrong model of approximation of very old people's migration. These problems are partly solved in the HMD estimates by using the extinct/almost extinct cohort method.

The sex ratio at ages 90+ also suggests that there has been data quality problems with the population estimates at very old ages during recent years. The issue was first detected in 2005, when male life expectancy at age 90 years (calculated with the HMD methodology using official population data) exceeded female life expectancy by 0.1 years (with values of 3.38 and 3.28 years, respectively). The discrepancy has increased since then so that in 2014 life expectancy at age 90 years for males is higher than that for females by over 0.5 year (with values of 4.39 and 3.87 years, respectively), a result that is quite implausible. Data from the most recent census do not seem to have improved to the extent that the problem is resolved.

For most population estimates at ages 80+, the HMD methodology uses the extinct cohort method, which may provide more accurate estimates than the official population estimates. Nevertheless, for more recent years when older cohorts are not yet 'extinct', the authors' methods for estimating population rely on the official estimates. Thus, data quality problems with these estimates influence the quality of HMD estimates for the previous 10-15 years (see the *Methods Protocol* for more details).

The authors' analyses of data suggest that the male population remains inflated even after excluding the Chechen Republic from the all-Russia population. Moreover, the problem cannot be attributed to increasing numbers of deaths at unknown age in recent years. Thus, it appears to be a systematic error, which may also (to some degree) affect the data for females. To avoid this problem, data from the official population estimates were aggregated to make the open age interval 80+. This approach has been used by Statistical Offices in similar cases. For example, since 2001 Statistisches Bundesamt (Germany) has provided population estimates with open age interval 90+ instead of 95+, and Statistics Lithuania has published data with open age interval 85+ instead of 100+ since 2003. Here the survival ratio method is used to split the population totals for ages 80+ instead of 90+ (see the *Methods Protocol* for details). Correspondingly, the extinct/almost extinct cohort methods were applied for all ages above 80. Using an open age interval 80+ produces more plausible results (see Figure 2).

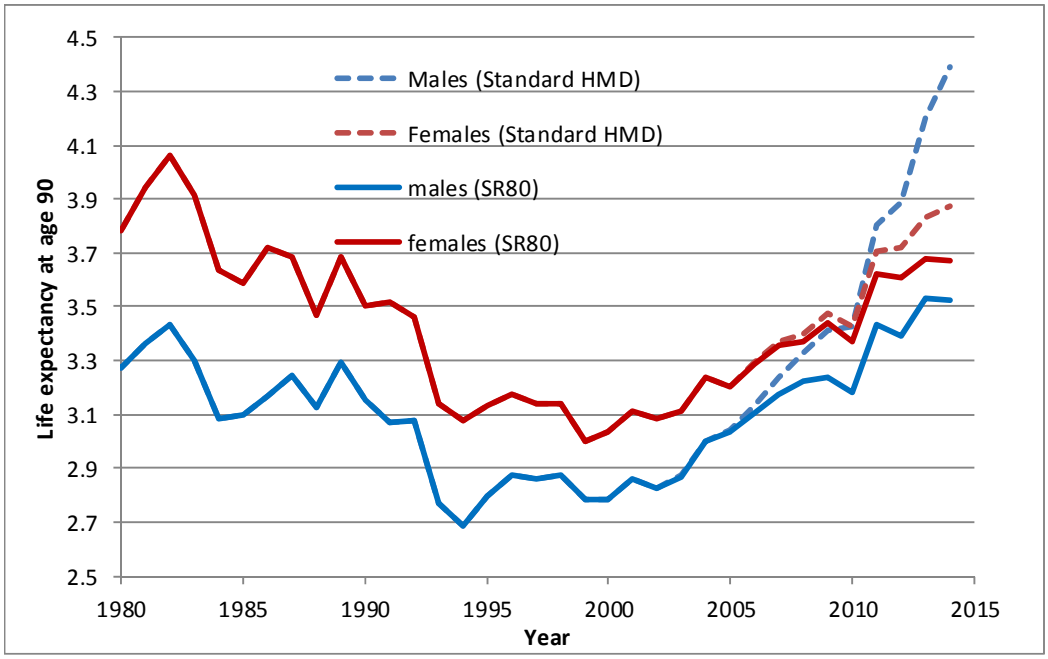
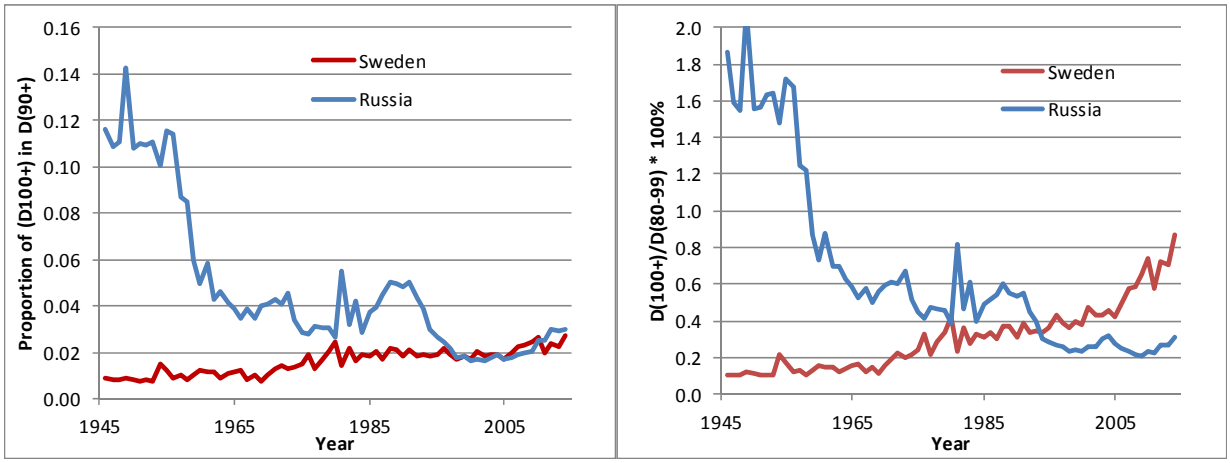


Figure 2: Life expectancy at age 90 estimated based on various open age intervals in 1980-2014. [Note: 'Standard HMD' refers to calculations using the survival ratio method with an open age interval of 90+ years (SR90+) and 80+ years (SR80+), respectively].

Age overstatement is suspected not only in the population estimates but also in the mortality data. Following Kannisto (1994, 1999), the ratio of deaths at ages 100+ to deaths at ages 90+ was used as an indicator of overstatement. Ratios which are considerably above those for Sweden, regarded as a standard for data quality, are considered as evidence of age overstatement. In a country with high quality data (and perfect age recording), this ratio rises as life expectancy at higher ages increases. As one may see on Figure 3a (left side), the ratio for Russian males is absolutely implausible up to the mid-1990s in comparison to the ratios for Swedish males. Moreover, taking into account the significantly higher levels of adult mortality in Russia during this period, the ratios are expected to be lower than in Sweden so that even those for the years 2000s, which are very similar to the ratios in Sweden, are still suspicious. Nevertheless, the progressive decline in the ratio of D(100+)/D(90+) since the mid-1990s may be interpreted as a sign of improvement in the quality of the deaths-by-age statistics. Unfortunately, this improvement appears to be explained not so much by an improvement in the data than by a drop in the number of deaths at ages 100+ years (relative to the number of deaths at ages 90+ years). Such a pattern could arise if the statistical office recoded unbelievably high ages at death into the "age unknown" category, thus removing them from the interval 100+ years but, unfortunately, we were unable to confirm this hypothesis even though we do have confirmation that death certificates have been found for this period with ages higher than 115 years. The ratio of deaths at ages 100+ years to deaths at ages 80-99 years (Figure 3b, right side) confirms our suspicions about the age overstatement in the early period but it shows significantly lower trend than Sweden in 1990s-2000s as expected if the quality of the age reporting was good.



Figures 3a and 3b. Ration of death counts in the open age interval 100+ years to the death count in the open age interval 90+ years (left side) and to the number of deaths at ages 80-99 years (right side) for Russian and Swedish males.

The combination of age overstatement in the population estimates (even partly corrected for by using an open age interval of 80+ years) and in the death-by-age statistics, together with the [suspected] underestimation of the number of deaths in the open age interval creates a complex challenge for mortality estimation. The survival ratio and extinct cohort methods used in the HMD protocol to construct population estimates at ages 80+ years from deaths above that age (see the HMD Methods Protocol for details) result in an implausible pattern in the death rates at ages 100 and above (Figure 4), with rates that are incorrectly lower than at younger ages (95-99). More details is provided in Appendix III.

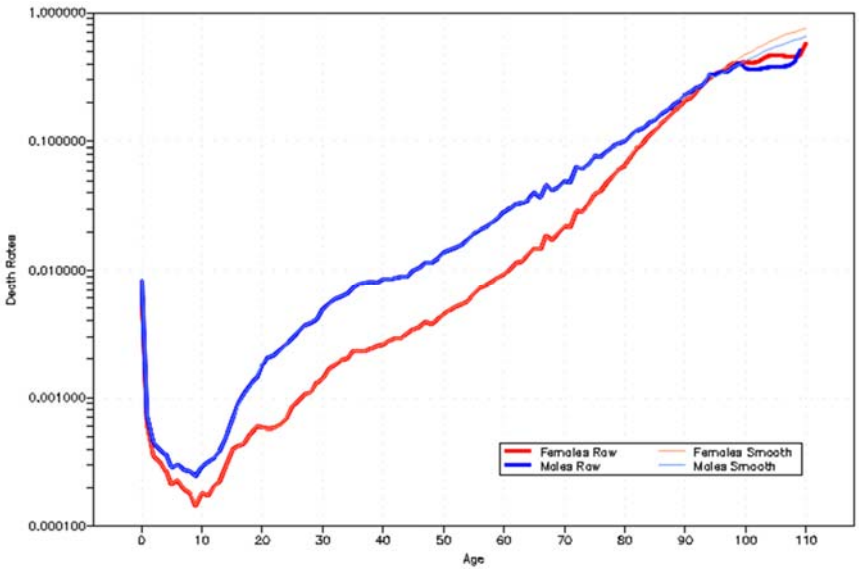


Figure 4. Death rates by age, Russia, 2014

At age 90+ the HMD uses smoothed death rates for the calculation of life tables and this problem does not affect any life table indicator. We strongly recommend using only the corrected HMD data and not the raw estimates for old age mortality.

INPUT DATABASE AND MORTALITY SURFACES IN THE HMD

Although all Russian data included into the Input Database may be of interest to users, the authors have decided to make further estimates of mortality surfaces for the period since 1959only. Nevertheless, data for 1959-1969 should be used with caution due to the problems of data quality discussed above (see also Appendix II).

REVISION HISTORY

The current update (up to 2014) includes not only new data for 2011-2014 but also revised population estimates for the last inter-censal period (2003-2010). Death counts for 2000-2008 were replaced by more detailed data classified by age and year of birth (Lexis triangles).

Changes with the December 2017 revision:

Life tables: All life tables have been recalculated using a modified methods protocol. The revised protocol (Version 6) includes two changes: 1) a more precise way to calculate a0, the mean age at death for children dying during the first year of life and 2) the use of birth-by-month data (where and when available) to more accurately estimate population exposures. These changes have been implemented simultaneously for ALL HMD series/countries. For more details about these changes, see the revised Methods Protocol (at http://v6.mortality.org/Public/Docs/MethodsProtocol.pdf), particularly section 7.1 on Period life tables and section 6 and Appendix E, on death rates. The life tables calculated under the prior methods (Version 5) remain available at v5.mortality.org but will not be further updated in the future.

ACKNOWLEDGEMENTS

We are grateful to Fiona Willis for her help with language editing.

REFERENCES

Anderson B.A., Silver B.D. (1986). "Infant mortality in the Soviet Union: regional differences and measurement issues." *Population and Development Review*, Vol. 12, No 4, pp. 705-738.

Andreev E. (1995). "Mladencheskaya smertnost v Rossii" (Infant mortality in Russia). *Vestnik Statistiki*, Vol. 5, pp. 66-71.

Blum A., Monnier A. (1989). "Recent mortality trends in the USSR: new evidence." *Population Studies*, Vol. 43, pp. 211-241.

Coale, A.J. and E.E.Kisker. 1986. Mortality crossovers: reality or bad data? *Population Studies*, 40: 389-401.

Garson L.K. (1991). "The centenarian question: old-age mortality in the Soviet Union, 1897 to 1970." *Population Studies*, Vol. 45, pp. 265-278.

Kannisto, V. (1994). *Development of oldest-old mortality, 1950-1990. Evidence from 28 developed countries.* In: B. Jeune and J. Vaupel (Eds). *Monographs on Population Aging*. *Vol. 1* (1-108 p.). Odense: Odense University Press.

Kannisto, V. (1999). Assessing the Information on Age at Death of Old Persons in National Vital Statistics. In: B. Jeune and J. Vaupel (Eds.). *Validation of Exceptional Longevity. Monographs on Population Aging. Vol. 6.* (pp. 240-249). Odense: Odense University Press.

Shkolnikov V.M, Meslé F., Vallin J. (1996). "Health crisis in Russia. I.Recent trends in life expectancy and causes of death from 1970 to 1993." *Population: An English Selection*, Vol. 8, pp. 123-154.

Velkoff V., Miller J.E. (1995). "Trends and differentials in infant mortality in the Soviet Union, 1970-90: How much is due to misreporting?" *Population Studies*, Vol. 49, pp. 241-258.

DATA SOURCE REFERENCES

Goskomstat. (2000). *Demograficheskiy ejegodnik Rossii* (The Demographic Yearbook of Russia). Moscow: Goskomstat.

Andreev E., Darskii L., Kharkova T. (1998). *Demograficheskaya istoria Rossii: 1927-1959* (Demographic History of Russia: 1927-1959). Moscow: Informatics.

APPENDIX I: DESCRIPTION OF DATA USED FOR LEXIS DATABASE

DEATHS

Period	Type of Data	Age Grouping	Comments	RefCode†
1959-1999	Annual number of deaths by sex and age	0, 1, 2, 3, 4, 5, …, 99, 100+	Unpublished data; for the purpose of HMD calculations, deaths are aggregated for ages 99+ during the period 1959-1989 (see 'Specific Details' under 'Death Count Data')	1
2000-2014	Annual number of deaths by sex, age, and year of birth	0, 1, 2, 3, 4, 5, …, 99, 100+	Annual officially registered deaths by single year of age and birth cohort, compiled from the death certificates. Unpublished data	42

POPULATION

Period	Type of Data	Age Grouping	Comments	RefCode†
1959, 1970	Population estimates as of January 1st	0, 1, 2, …, 100+	Census data	3, 4
1971-1978	Population estimates as of January 1st	0, 1, 2, …, 85+		6
1979	Population estimates as of January 1st	0, 1, 2, …, 100+	Census data, estimated population as of January 1st	3
1980-2015	Population estimates as of January 1st	0, 1, 2, …, 100+	For the purpose of HMD calculations, population estimates are aggregated for ages 80+ (see 'Data Quality Issues')	8, 21, 43

BIRTHS

Period	Type of Data	Comments	RefCode†
1959-2014	Annual counts of births by sex		9, 10, 20, 29, 32, 37, 40, 41

† The reference code is used in the raw data files (Input Database) to link data with sources

APPENDIX II: DESCRIPTION OF DATA USED IN THE INPUT DATABASE

Research note by Sergei Zakharov[1] April 2001

DEATHS

Type of data:

Annual officially registered number of deaths by age, compiled from the death certificates.

Age grouping:

Single age groups (0, 1, …, 99, 100+)

Period covered:

1946 – 1999

Comments:

1. No adjustment has been made for age heaping or for underestimation.
2. Slight adjustment has been made for territory compatibility:
    - Crimea region was subtracted by direct method (1946-1956);
    - Data for a few regions—namely Kaliningrad, Tuva, and Karel provinces— are not available or incomplete for some years before 1956. Based on fragmentary data available, it was decided to include some extra deaths. About 6,500 deaths for 1946 and 1953-1955, and about 900 for 1948 and 1949 were added and distributed proportionately by age.

POPULATION

Type of data:

Permanently resident population on January 1st, since 1946. Different sources have been used for different calendar periods and different age groups.

Age grouping:

Single age groups (0, 1, …, 99, 100+)

Period covered:

1946 – 1999

---

[1] Centre for Demography and Human Ecology (Moscow)

Comments

1. Official estimates as of January 1st are based on the 1959, 1970, 1979, 1989 Census Data. For the 1959 Census, only an "Actually present population" ("Nalichnoe naselenie") distributed by age is available, because only this category of population was processed according to the program of data processing adopted for this Census.
2. Official (Russian Federation State Committee on Statistics – Goskomstat) estimates for 1979–1999 are based on 1979 and 1989 Census data, annually registered numbers of births, age-classified numbers of deaths, and estimated numbers of net migrants by age. For migration estimates, a special procedure was applied. This procedure was developed by TCSU (USSR Central Statistical Board) in 1978, and then modified several times. Currently, Goskomstat estimates for migration are based on registration information of the Ministry of Internal Affairs (registration of in- and out- migrants) and the State Committee on Migration (registration of forced migrants, refugees etc.). The procedure allows for some adjustments for underestimation and non-formalised expert estimates as well.
3. Official (USSR Central Statistical Board – TCSU) estimates for 1970-1978 are supplemented by E. Andreev's[2] interpolations for people aged 85 and over. Before 1979, the TCSU used only an "Actually present population" ("Nalichnoe naselenie") category rather than a "Permanently resident population" ("Postoyannoe naselenie") category for one-year age distribution of people aged 85+. E. Andreev's estimates are included here since they are the only estimates known to the authors, and are more or less reliable. However, they are necessarily highly smoothed and have to be used carefully.
4. Estimates for the 1959-1970 intercensal period were made by E. Andreev. In the 1990s, E. Andreev made estimates for this period using conventional methods. No official distributions of population by one-year age groups exist for 1959-1970.
5. Backward extrapolation for the period 1946-1959 based on the 1959 Census data was undertaken by E. Andreev in the 1990s (smoothed data). There are no official distributions of population by one-year age groups for this period.
6. For 1959, two population distributions are available: a smoothed one as a base for the 1946-1959 backward extrapolation mentioned above, and a second one estimated as of January 1st based on the official Census data.

SOME DATA PROBLEMS REVEALED

1. Age-heaping effect

The most serious problem with the Russian mortality data is pronounced age heaping on deaths at ages ending in 0 and 5. For all years prior to 1970, it is clear that age heaping in number of deaths results in abnormally high mortality rates at every age ending in 0 and 5, and, consequently lower rates between these ages. At older ages, the age-heaping effect is more pronounced. It is also evident that by applying a smoothed distribution of population by age to calculate mortality rates (e.g., for the 1946-1959 period), the age-heaping effect is just reinforced.

---

[2] Andreev, Evgenij, Ph.D., Leader of Expert Group on Methodology and Population Estimates in the Russian Federation State Committee on Statistics. He is also a Head of the Laboratory in the Centre for Demography and Human Ecology, Russian Academy of Sciences Institute for Economic Forecasting.

The figures below demonstrate the problem for the male population. The same is true for females (not presented here).

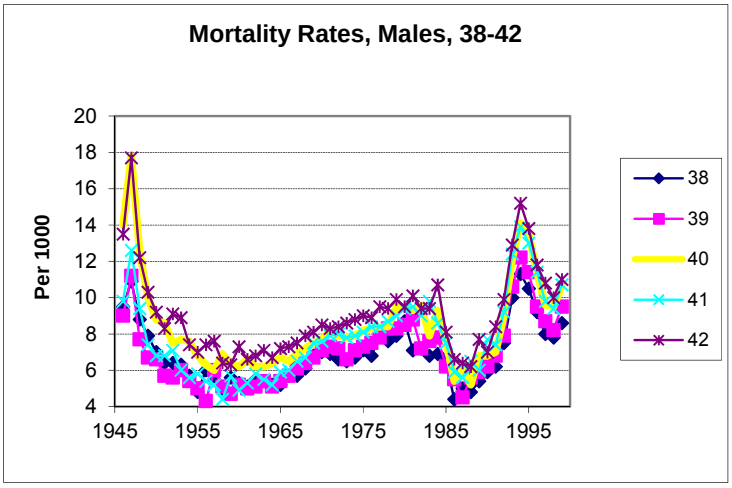


Figure 1

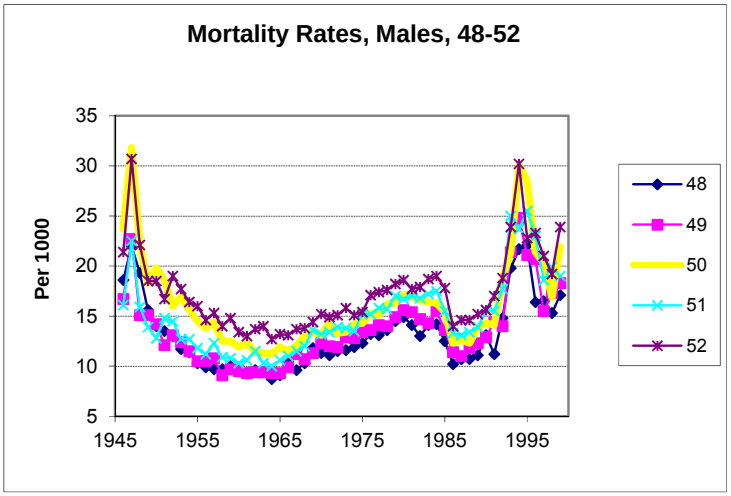


Figure 2

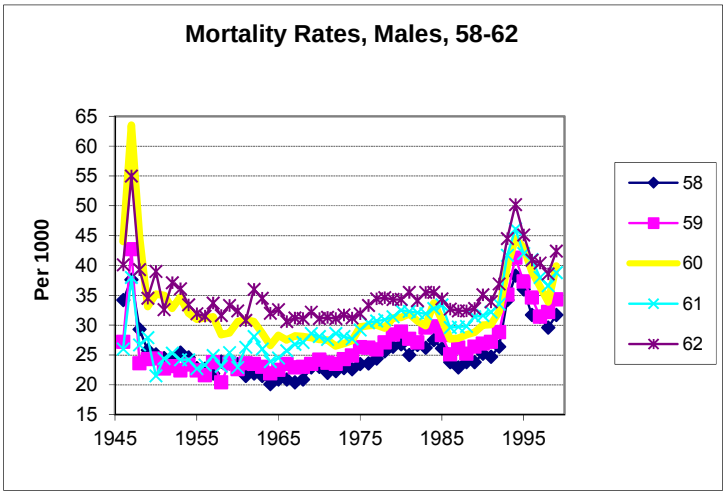


Figure 3

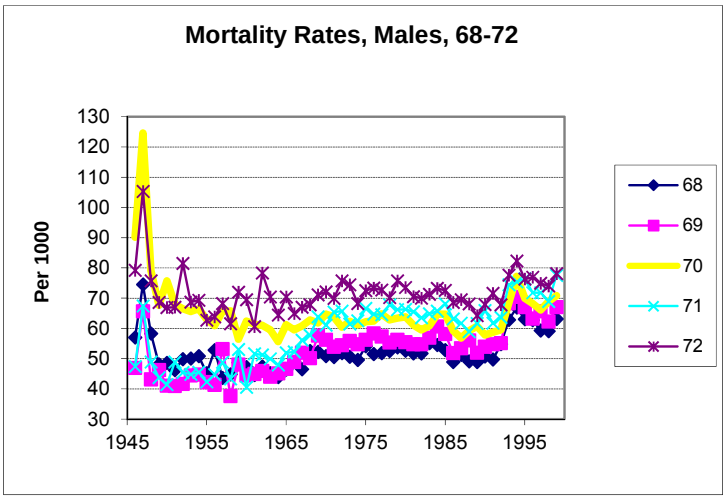


Figure 4

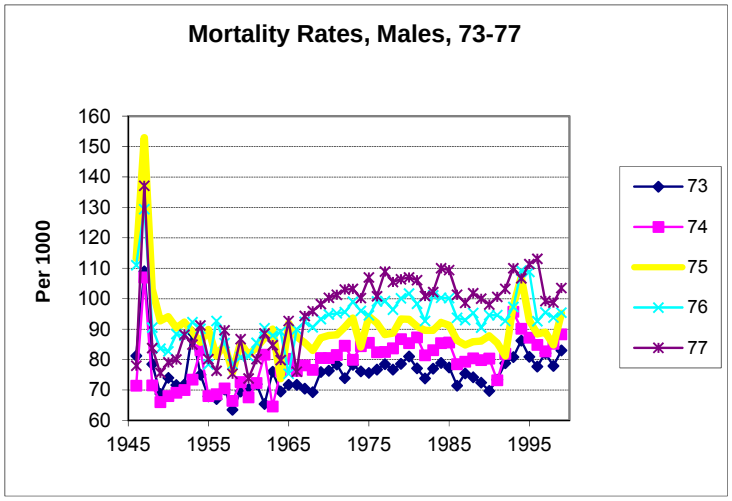


Figure 5

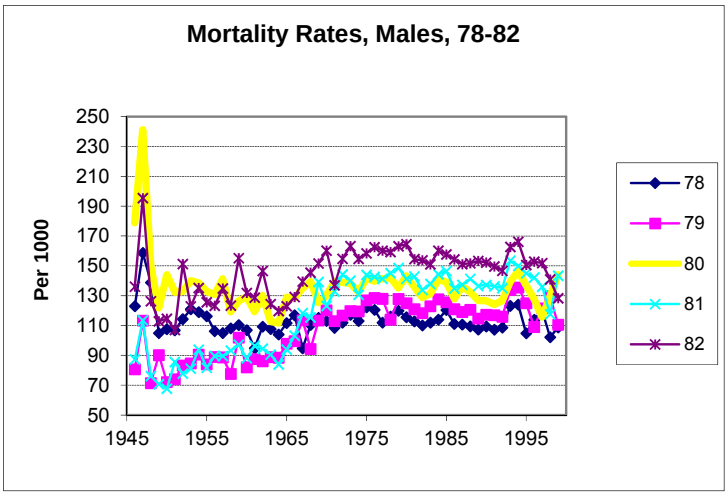


Figure 6

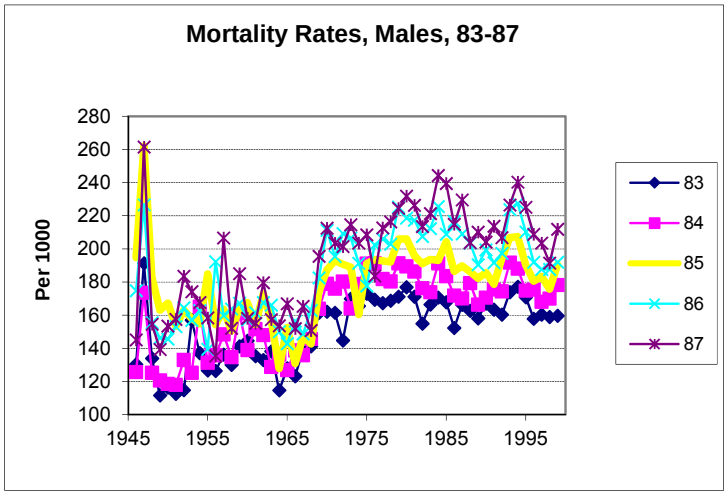


Figure 7

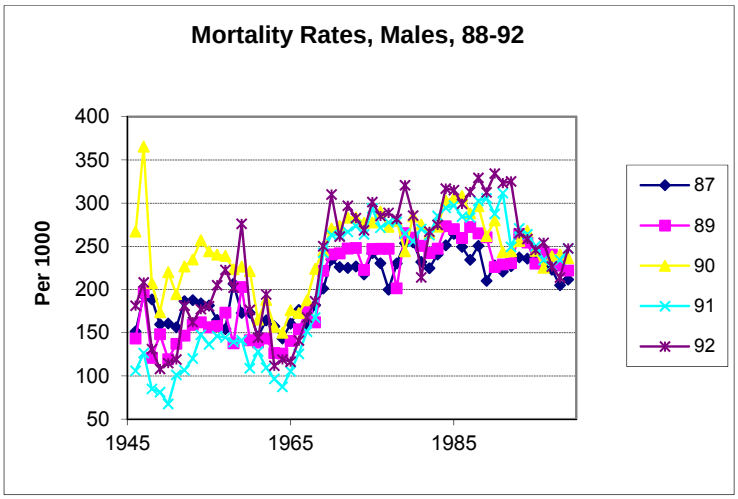


Figure 8

Figures 1-8: Mortality Rates at selected ages, Russia, Males: 1946-1999.

Taking into account this serious age heaping problem, it was decided that the Human Mortality Database should not include the whole Russian data set, but only series for the period since 1970.

2. Unnatural trend in number of deaths at age 99

In the Russian data for the whole period under study, age 99 is the last closed age interval, and 100+ is an open-ended age interval. It is clear that the number of deaths at age 99 in the period 1958-1987 is too high relative to deaths at nearby ages such as 97, 98, and 100+ (see the ratios presented in Figure 9). Given the situation of age heaping at other ages, we would expect the opposite effect for an age ending in 9 (i.e., artificially lower numbers of

deaths). This effect therefore suggests a systemic error of another nature, most likely associated with some rules or special instructions adopted for a registration procedure.

One feasible explanation could be the following; the age of people who die as indicated on the death certificate is a result of simple subtraction of the year of birth (e.g., specified in the passport of the deceased person) from the known year of death. This would mean that a large portion of deaths (more than 50%) would be wrongly classified by age in a cohort-wise manner.

Taking this problem into account, it was decided that 99+ would be used as an open-ended age interval for the period 1970-1988 instead of the 100+ age group that is officially used.

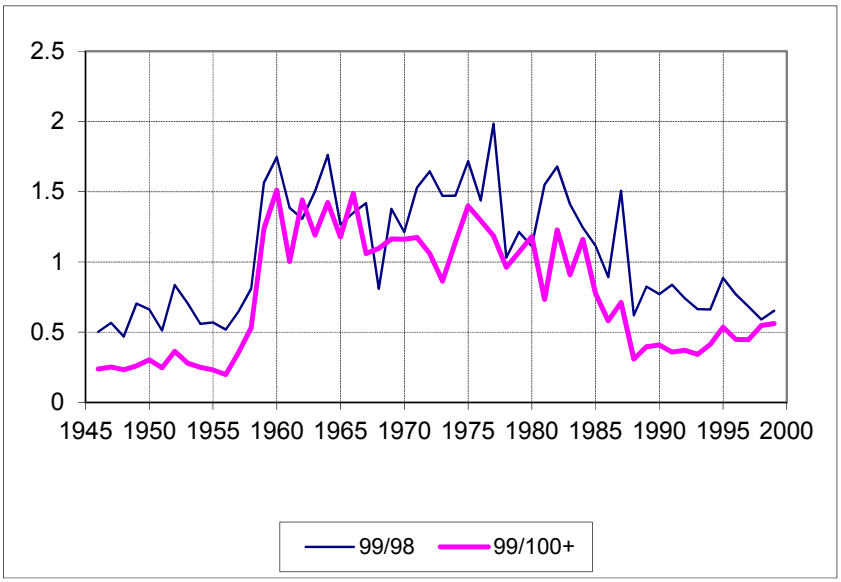


Figure 9. Ratio of the number of deaths at age 99 years to the number of deaths at age 98 years, and to the number of deaths at age 100+ years, Russia, Males: 1946-1999.

CONCLUSION

1. There are serious errors in the official estimation of the population of elderly men and women reported by the Russian Federation's Goskomstat, especially for persons over age 80. As a result, elderly population changes and, correspondingly, mortality rates, become highly irregular and inaccurate. The older the people, the more problematic the official data. These errors are of a methodological character and, most probably, have to do with use of the wrong model of approximation for very old people's migration patterns.
2. Application of the extinct cohort method gave good results. New estimates of the population over age 80 and of age-specific mortality rates re-calculated with the new denominator are much better than corresponding official data with respect to their dynamics over time. At least, the unjustified irregularity of population changes and mortality rates has been attenuated. Inconsistency between age-specific mortality rates has also been reduced.

3. Nonetheless, the problem of death age heaping is still one of the main problems of Russian data. The authors did not intend to solve this problem, and leave it open to researchers who would particularly focus on it. An attempt has been made to mitigate this problem by the data matrix to the period after 1970. Therefore, care should be taken in the use of the Russian data on number of deaths and mortality rates by one-year age groups included in the HMD.

APENDIX III. Survival ratio in 2014

The HMD methodology requires that population at ages 80+ years be systematically re-estimated using a combination of the extinct and almost extinct cohort methods and the survival ratio method. The key factor for estimating population counts at ages 90+ is the survival ratio (R) over the five immediately-older cohorts[3]:

$$R = \frac{P(x,t)}{P(x-5,t-5)}$$

The coefficient *R* is estimated for each cohort using data for the five immediately older cohorts. Population at age *x* in a given cohort is then estimated as:

$$P(x,t) = \frac{R}{1-R}\dot{D} \ ,$$

where $\dot{D}$ denotes the number of deaths which occurred in the cohort over the age interval $[x-5, x)$. Value of coefficient *R* by age for Russia in 2014 is shown in Figure III.1.

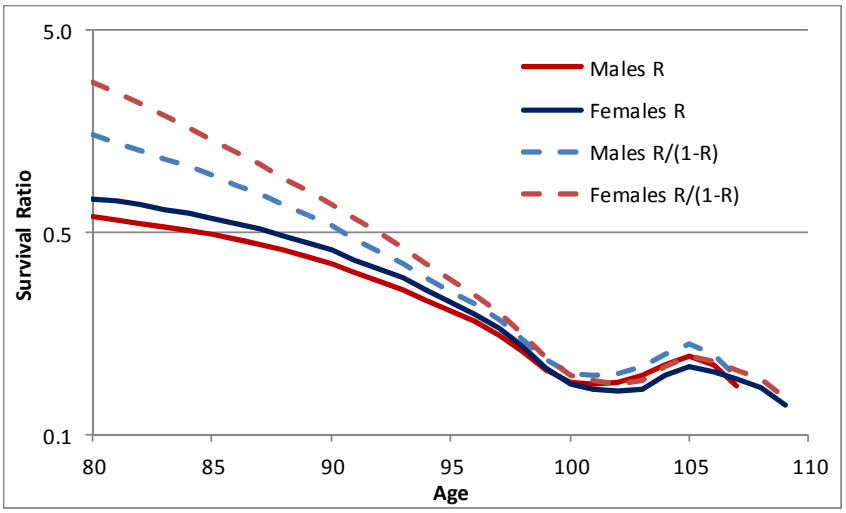


Figure III.1 Survival ratio by age, Russia, males and females, 2014.

The underestimation in the number of deaths at ages 100+ years in Russian statistics leads to artefactual fluctuations in the survival ratio coefficients by age and, as a result, in the death rates at ages 100-109. This affects not only rates for the most recent year, but also those in the previous years, for which population at ages 80+ years is estimated using extinct/almost extinct cohort method. The artificial pattern in the death rates can be observed for the whole period 2005-2014. In earlier years, the influence of population estimates based on the survival ratio method is almost negligible (since most of the estimation relies on the extinct or nearly-extinct cohort method).

[3] See the HMD Methods protocol for details and exact notation.