The Human Cause-of-Death Database

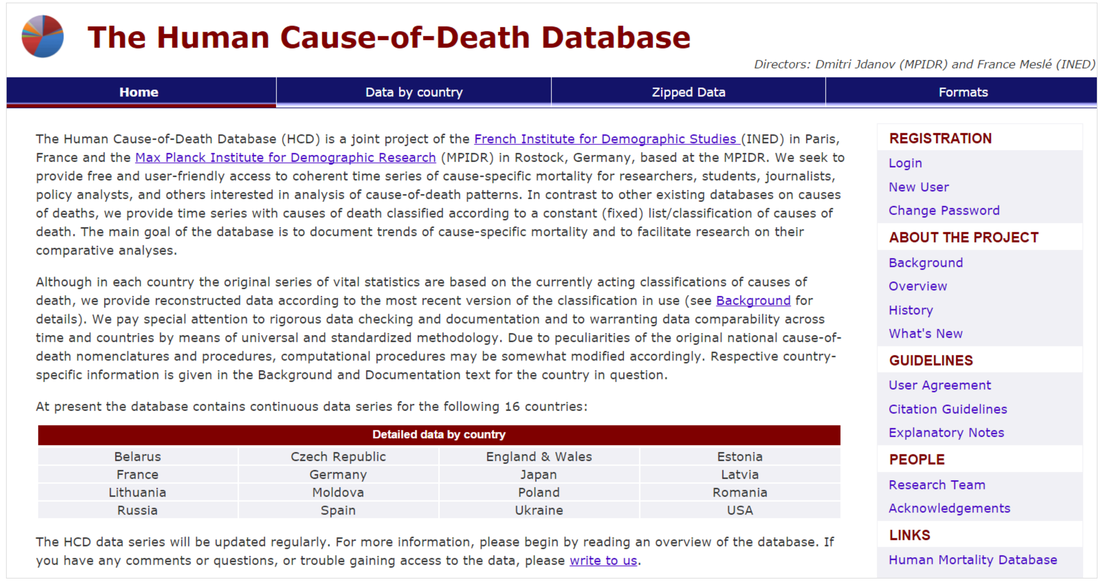
Free access to coherent time series of cause-specific mortality

Inna Danilova, France Meslé, Dmitri Jdanov, Markéta Pechholdová,
Domantas Jasilionis, Vladimir M. Shkolnikov, Jacques Vallin

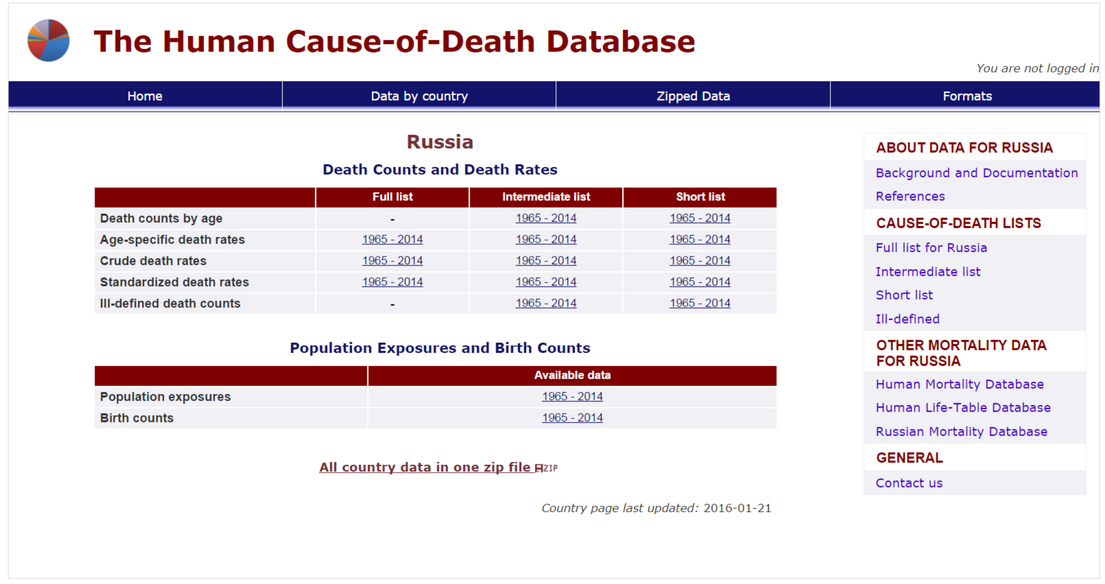
INSTITUT NATIONAL D'ÉTUDES DÉMOGRAPHIQUES

MAX-PLANCK-INSTITUT FÜR DEMOGRAFISCHE FORSCHUNG
MAX PLANCK INSTITUTE FOR DEMOGRAPHIC RESEARCH

www.causesofdeath.org - Main page



www.causesofdeath.org - Country page



BACKGROUND

The Human Cause-of-Death Database (HCD) is a joint project of the French Institute for Demographic Studies (INED) in Paris, France, and the Max Planck Institute for Demographic Research (MPIDR) in Rostock, Germany, based at the MPIDR.

The main goal of the HCD is to provide access to detailed high-quality data on cause-specific mortality to a broad audience of users. In contrast to other existing databases on causes of death, the HCD provides time series with a coherent classification of causes, based on ICD-10. For comparability purposes, we provide mortality data classified according to a short list and an intermediate list of causes of death, identical for all countries. In addition, a detailed list is provided, which varies according to country-specific availability.

The following features make the HCD particularly attractive to its users:

- Continuous data series with coherent cause-of-death classification;
- Availability of basic age-standardized indicators
- Detailed documentation
- Free and easy access to all data
- A uniform and easy to use format of data files

TYPES OF DATA IN THE HCD

Detailed data by causes of death

The HCD includes sex and age-specific death counts and death rates, crude death rates and standardized death rates by causes of death, after redistribution of ill-defined causes of death (according to country-specific methods described in the Background and Documentation for each country). Causes of death are classified according to three lists: short, intermediate and full.

The full list is country-specific; it includes 4- or 3-digit items of ICD-10 or equivalent country-specific lists. For comparability, we also provide mortality data classified according to intermediate and short lists of causes of death, which are identical for all countries. **The intermediate list** consists of 104 items compared to 16 in the **short list.**

Population exposures and birth counts

The source for these data is the Human Mortality Database (www.mortality.org). In some cases we correct infant mortality and, respectively, birth counts

THE METHOD OF RECONSTRUCTION

Cause-of-death time series are disrupted by periodical changes in the disease classifications. This limits mortality analysis and only allows to analyze cause-specific time-trends for a short period (covered by the same classification) or only for broad groups of causes of death.

To reconstruct consistent series, it is necessary to establish transition coefficients between items of two successive classifications, in order to redistribute deaths classified according to the old classification into items of the new classification. When bridge coding (double classification of deaths simultaneously into the old and new classification) has been performed, transition coefficients can be inferred directly from the results, but there are only two countries in the database where this has been done (and only for the transition from ICD-9 to ICD-10), namely England and Wales and the U.S.A. For the other transitions coherent time series are reconstructed by producing ex-post double coding. The method developed at INED in the 1980s is used as a guideline, but the work was tailored to each country independently.

For each classification change, the method comprises three steps (Vallin and Meslé, 1988, 1998; Meslé and Vallin, 1996):

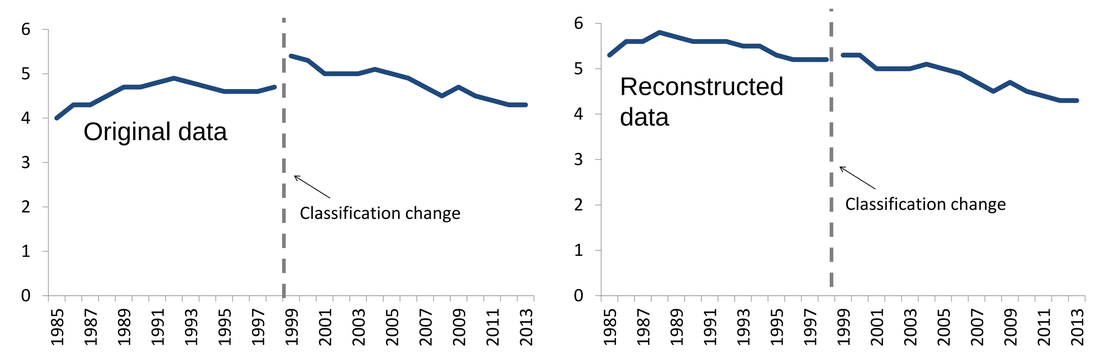
- Setting up a correspondence table which lists, for each item of one classification, all items of the succesive one that are a priori equivalent in terms of medical content.
- Building fundamental associations of items that identify the smallest possible number of items containing the same medical contents in both classifications and testing the consistency of the associations over time using a statistical test (Barbieri, Chung, and Boe, 2008; Camarda, Pecholdová, and Meslé, 2015).
- Setting up ex-post double-coding according to the structure of fundamental associations, to finally obtain transition coefficients.

The results derived from the medical logic of the classification rules **have to be checked statistically**, to detect and solve any remaining breaks in the series. Such checks are carried out by age group and sex.

In addition, national statistical offices introduce occasional changes independent of the official revisions of the classification. To address this problem, the statistical continuity of the series over time is systematically verified and any artificial disruption dealt with appropriately.

Finally country- and time-specific methods are used **to deal with ill-defined causes** (Ledermann, 1955; Vallin and Meslé, 1988).

Figure 1. Leukemia mortality trend in Russia before (left) and after reconstruction (right), age-standardized rates per 100 000



DATA AVAILABILITY

Country	Period
Belarus	1965-2010
Czech Republic*	1994-2014
England and Wales*	2001-2013
Estonia	1955-2012
France*	2000-2013
Germany*	1998-2013
Japan*	1995-2013
Latvia	1956-2012

Country	Period
Lithuania	1956-2012
Moldova	1965-2012
Poland	1959-2013
Romania	1980-2012
Russia	1965-2014
Spain	1980-2012
Ukraine	1965-2013
USA*	1999-2013

*****_Currently, we have several countries with short data series starting with the introduction of ICD-10 in the country. We are working on data reconstruction for these countries, and in the near future these data will be replaced by longer, reconstructed time series with constant classification of causes._

ACKNOWLEDGEMENTS

This project would not have been possible without the enthusiasm and assistance of a large number of colleagues at various institutions. We are grateful for all their contributions.

The work was completed as a part of the AXA project Mortality Divergence and Causes of Death (MODICOD) funded by the AXA Research Fund and the project Diverging Trends in Mortality and Future Health Challenges (DIMOCHA) funded by the Agence nationale de la recherche and the Deutsche Forschungsgemeinschaft. This project also would not have been possible without the encouragement and financial support of the French Institute for Demographic Studies (Paris, France) and the Max Planck Institute for Demographic Research (Rostock, Germany)

REFERENCES

1. Barbieri, M., Chung, R., & Boe, C. (2008). *Automating the redistribution of deaths by cause over ICD changes.* Second Human Mortality Database Symposium, Max Planck Institute for Demographic Research, Rostock, Germany, 13-14 June 2008.
2. Camarda, C.G., Pechholdová, M. & Meslé, F. (2015). *Cause-specific senescence: classifying causes of death according to the rate of aging.* 80th Annual Meeting of the Population Association of America. San Diego (USA), May 2015. http://paa2015.princeton.edu/uploads/153074
3. Ledermann, S. (1955). *La répartition des décès de cause indéterminée.* Revue de l'Institut international de statistique, 23 (1–3), 47–55.
4. Meslé, F., & Vallin, J. (1996). *Reconstructing long-term series of causes of death*. Historical Methods, 29 (2), 72–87.
5. Vallin, J., & Meslé, F. (1988). *Les causes de décès en France de 1925 à 1978* (Travaux et Documents, No.115, 608 p.). Paris: INED/PUF.
6. Vallin, J., & Meslé, F. (1998). *Comment suivre l'évolution de la mortalité par cause malgré les discontinuités de la statistique. Le cas de la France de 1925 à 1993.* In G. Pavillon (Eds.), Enjeux des classifi cations internationales en santé (Questions en santé publique, pp. 113–156, 220 p.). Paris: Éditions INSERM.