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 (MPI-DR) in Rostock, Germany, based on 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.

### 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 successive 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 (Barberi, Chung, and Boe, 2008; Camarda, Pechholdová, and Meslé, 2015).
- Setting up an 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 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.

### Types of Data in the HCD

Data availability includes 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.

### 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