

## Adjusting the German cause-of-death statistics: redistributing “non-informative” ICD-codes

Annelene Wengler

---

### Background

The German cause-of-death statistics are often utilized to draw conclusions about the health status of the population and the significance of certain diseases. Within the project BURDEN 2020 - *The burden of disease in Germany and its regions* – at Robert Koch Institute in Germany we use the cause-of-death statistics to calculate the years of life lost (YLL). The YLL measure the remaining life expectancy that would have been expected for the person at the time of death.

Unfortunately cause-of-death statistics – not only in Germany – often include a relatively high proportion of *non-informative* ICD-codes. *Non-informative* means that an ICD-code does not sufficiently reflect the underlying cause of death. For example this may be the case when the indicated ICD-code is intermediate (e.g. heart failure) or non-specific (e.g. unspecified cancer). Among medical staff and scientists alike there are different opinions about what should be considered an underlying cause of death and what not. The setting in which the ICD-codes are used and the reason why they are analysed must also be kept in mind. An ICD-code which is non-informative as an underlying cause of death may well be informative and valid in a different context, e.g. when specifying the causal chain till death.

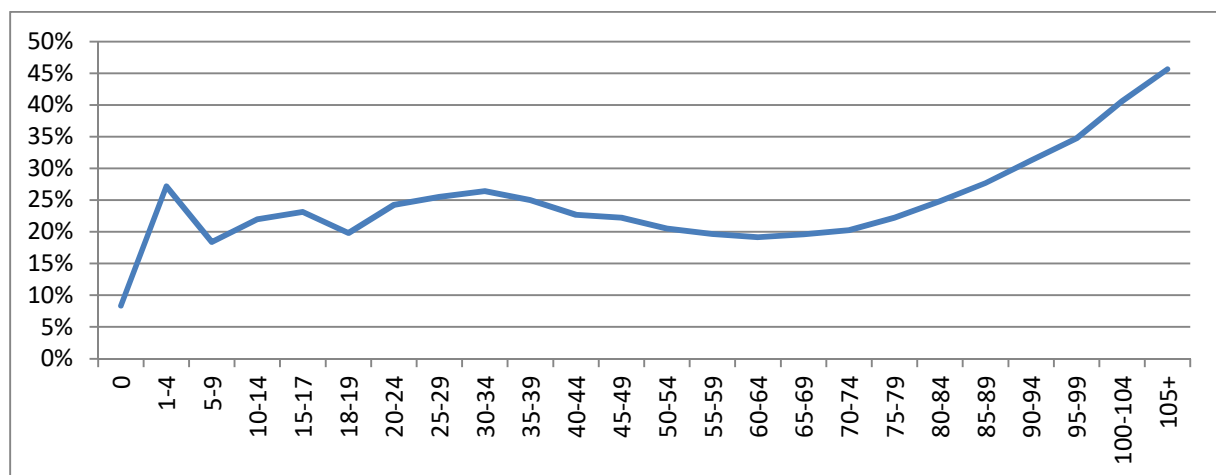
### Method

We use the framework of the Global Burden of Disease Study (GBD) carried out by the Institute for Health Metrics and Evaluation (IHME) to differentiate between non-informative and informative ICD-codes for the underlying cause of death. We analyse how many deaths with non-informative codes can be found in the cause-of-death statistics in Germany in absolute and relative terms and how they are distributed by age, sex and region. After identifying non-informative ICD-codes it is necessary to further process those codes. We do so by redistributing non-informative codes to valid ICD-codes. Additionally we present how the cause-of-death statistics change during the redistribution process and compare the number of deaths for specific causes before and after redistribution.

## Results

Using the most current data for Germany (2017) and the framework of the Institute for Health Metrics and Evaluation (IHME) in the Global Burden of Disease Study (GBD) the proportion of non-informative codes is 26.0 %. In the past, we have also looked at a different framework to identify non-informative ICD-codes. According to the WHO framework, 15.6 % of the 925,200 deaths in Germany in 2015 have a non-informative ICD-code. Respectively the share was 26.6 % in 2015 when using the framework of the Global Burden of Disease Study (GBD). In general the share of non-informative ICD-codes is lower using the WHO framework but we see similarities when looking at the distribution patterns across age. As figure 1 indicates the share of deaths with non-informative ICD-codes is highest for the oldest age groups. In general it can be constituted that the share of deaths with non-informative ICD-codes seems to be decreasing in the recent years.

*Figure 1* Share of deaths with non-informative ICD-codes (of all deaths) in 2015

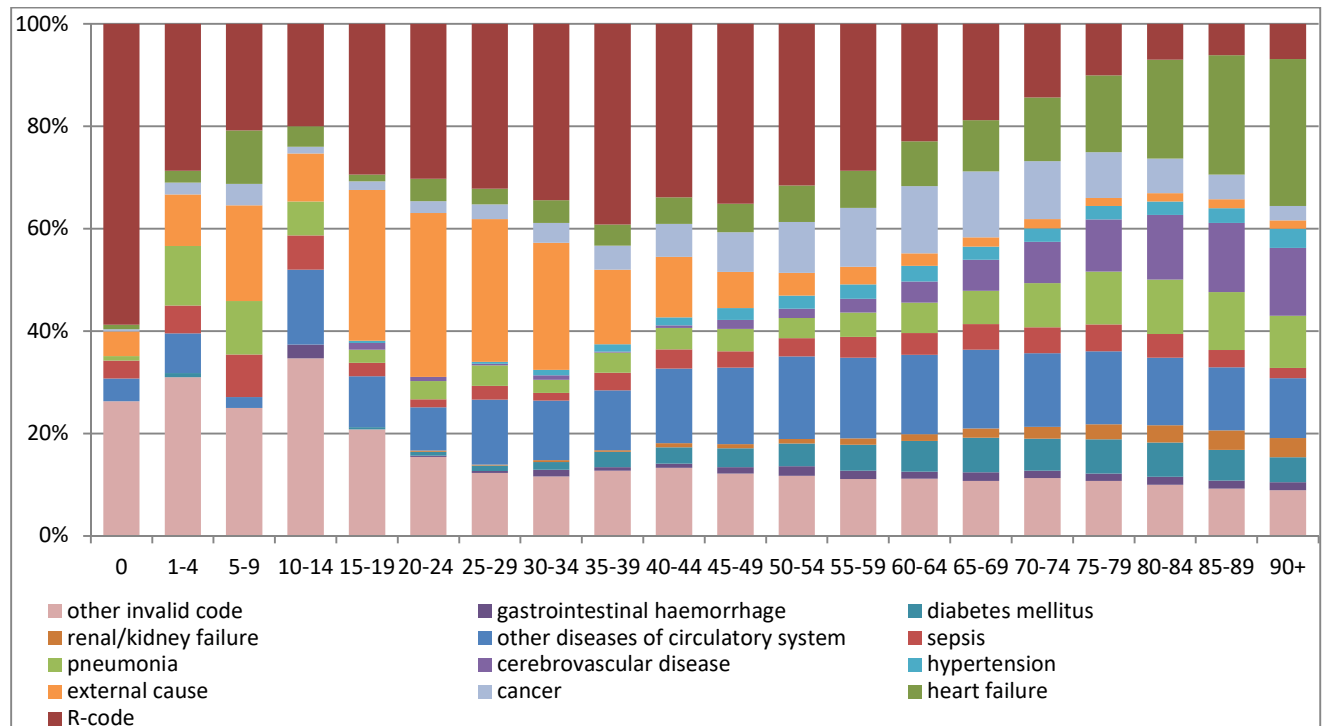


Source: German cause-of-death statistics, 2015 (own calculations)

In figure 2 we grouped the non-informative ICD-codes (e.g. R-codes or specific cause-groups). R-codes seem to be especially relevant in the younger age groups. Non-informative ICD-codes in the field of external causes take up a large share in the middle ages and heart failure is an important non-informative ICD-code in the older age groups.

Considerable differences exist between the federal states, with shares of non-informative codes between 16 % and 35 % (IHME framework) (results not shown). The differences between the federal states can only partially be explained by different electronic data processing. Due to the further dissemination and improvement of the electronic data collection, a higher cause-of-death statistics quality in Germany can be expected in the future.

Figure 2 Distribution of non-informative ICD-codes across different groups, by age in 2015



Source: German cause-of-death statistics, 2015 (own calculations)

For some diseases (e.g. diabetes) we see that the number of deaths manifolds after the adjustment process. This has strong implications for drawing conclusions from cause-of-death statistics and hence for public health research in general. The cause of death can be seen as the final health outcome of a person. Knowing precisely what people die off and in what number is of great importance when analyzing the health of a society.