2 Methodological Aspects

2.1 Estimating the completeness of case registration by the epidemiological cancer registries

The utility of population-based data on cancer occurrence largely depends on the level of completeness with which new cancer cases are registered. Since 2010, the German Centre for Cancer Registry Data (ZfKD) has annually estimated the completeness of the data collected by the epidemiological cancer registries in all federal states. This estimate is conducted with the help of an internationally recognized standard indicator: the ratio of mortality to incidence. If there are no significant differences in the survival prospects of cancer patients, for example due to diagnostics or treatment, this ratio (the M/I index) can be assumed to remain largely constant in a particular region for a specific cancer diagnosis. The expected incidence, therefore, is estimated using the M/I index in a reference region (assumed to have complete coverage) and the mortality data from the region being examined. The resulting estimates are then compared with the data reported from the region’s registry. These calculations exclude cases that have only been identified via death certificates (DCO cases). The completeness of the data from the reference region is also estimated using this method.

The following criteria were established a number of years ago for the selection of registries for the reference region:

- Comprehensive cancer registration for at least ten years
- An average completeness for total cancers above 90% over the last ten years and above 80% for each individual year
- An average of under 15% DCO cases over the last ten years, excluding the first five years of registration.

These criteria are currently met by Saarland, Hamburg, Schleswig-Holstein and Lower Saxony, as well as by four of the seven registry districts in Bavaria, and the eastern region of North Rhine-Westphalia (Westfalen-Lippe).

The calculations described above are carried out with data stratified by age group, sex, and 16 or 17 diagnosis groups.

If mortality is particularly low in the area being tested (an average of fewer than five deaths per year), the number of new cases for a particular stratum is estimated using the ratio of incidence to population in the reference region instead of the incidence-to-mortality ratio. Estimated completeness for each diagnosis group is calculated as the ratio of the sum of observed cases to the sum of estimated cases from all age groups for men and women combined. The completeness estimates for “total cancers” are calculated based on the sums of the observed and expected values from all diagnosis groups.

The limitations of this method become apparent particularly when deaths for a certain cancer are either generally low, low in relation to incidence (this applies to testicular cancer, malignant melanoma and thyroid cancer), or if the real ratio of mortality to incidence varies between regions. The latter may be the case, for example, if screening rates differ between federal states or if screening was introduced in different years, as was the case with breast cancer screening. Even a difference in the regional distribution of tumour stages or different cancer subtypes (such as in thyroid cancer) can distort the figures.

Current estimates for the last two available years show that ten federal states achieved an estimated completeness of at least 90% compared with the reference registries mentioned above; five exceeded 95% (see Figure 2.1.1). As the Joint Cancer Registry was unable to provide data for 2016 from any of the federal states encompassing former eastern Germany and the city of Berlin, estimates of completeness for these states are based on data from 2014/2015. As most of the 2015 data from Bremen still needs to be processed, data from 2014 and 2016 were used instead. The results published here, therefore, are less robust than those published in the previous issue of “Cancer in Germany”. As described above, the nationwide rollout of clinical cancer registration and the associated reorganisation of existing registries has led to substantial delays in data processing in some registries. Nevertheless, once the current restructuring phase has been completed, epidemiological reporting will benefit from the nationwide expansion of clinical cancer registration, and the current data gaps for 2015 and 2016 are also expected to be closed.
2.2 Estimating cancer incidence in Germany

Since the last issue of »Cancer in Germany«, the circumstances described above have prompted a change in the methods for estimating cancer incidence in Germany. The cancer incidence figures in this report are derived from the results of a mixed Poisson regression model. The model accounts for the respective cancer-specific mortality, population size and year of diagnosis, with variations in incidence between registries modelled as random effects. In order to apply the regression model, data were stratified by sex, registry, diagnosis and age group. The regression model was fitted using data from registries that met the following quality criteria: at least 10 years of operation, an estimated completeness of ≥90 % for »all cancers« for at least five years, and fewer than 15 % DCO cases per year. Data from the first five years of coverage were excluded because they contain a substantial number of DCO cases that correspond to cancers diagnosed before registration began, inclusion of which would result in overestimated incidence. For those federal states that did not meet these criteria for inclusion in the reference region for any diagnosis year (Baden-Württemberg, Berlin, Hesse and Saxony-Anhalt, Rhineland-Palatinate, Saarland, Thuringia, Saxony and Saxony-Anhalt), data were included only for those years in which all quality criteria were met.
Anhalt), incidence was estimated for 2011 to 2013 using the ratio of incidence to mortality in the reference region and the mortality figures in each excluded federal state. These data were then incorporated into the data for the reference region for model fitting. The resulting regression parameters were used to estimate incidence for the various diagnoses and age groups according to the corresponding registry, mortality figures and year of diagnosis. Nationwide incidence was calculated as the sum of these state-specific results. In contrast to the previous procedure, DCO cases were treated in the same manner as regularly reported cases.

Analyses demonstrate that the new model largely compensates for the substantial decline in recorded case numbers as well as for the gaps in the data in some registries. As such, the results appear plausible, particularly with regard to general trends. For diagnosis years 2002 through 2005, the new method resulted in slightly lower estimates, probably because Bavaria and Lower Saxony, both of which have large populations, established their registries during this period.

The estimates for the most recent diagnosis years are based on a more extensive data pool and, therefore, can generally be regarded as more robust than results for earlier years.

Although the incidence of non-melanoma skin cancers (ICD-10 C44) was calculated using the same method, no estimates of completeness are available due to the low mortality associated with this type of cancer. In this case, the reference regions encompassed data from the six federal states used to calculate estimates published in the previous two editions of »Cancer in Germany«: Schleswig-Holstein, Lower Saxony, North Rhine-Westphalia, Saarland, Mecklenburg-West Pomerania (through 2015) and Rhineland-Palatinate. For the first time, this edition presents temporal trends in nationwide incidence of non-melanoma skin cancers in a separate chapter (Chapter 3.14). Due to data limitations, incidence of non-melanoma skin cancers was estimated for a shorter period (2006 to 2016) than for other types of cancer. Furthermore, in line with international conventions, non-melanoma skin cancers are not included in the estimated incidence of »all cancers« (Chapter 3.1).

2.3 Indicators and data presentation

The following provides more information about the statistics, graphs and tables presented in subsequent chapters.

Age-specific rates

Age-specific rates were calculated by dividing the number of cancer cases or cancer deaths in a given age group by the corresponding number of women or men from that age group in the population. The graphs depicting these rates demonstrate the relation between age and disease frequency by sex. Age-specific incidence rates are reported as the number of new cases per year per 100,000 residents in a particular age group.

Age-standardised rates

As the age-specific incidence rates in this report demonstrate, cancer risk usually increases substantially with age. Therefore, in order to compare incidence and mortality in different federal states, different regions or within the same region at different times, age standardisation should be used to account for differences in the age structure of these different populations. To this end, the observed age-specific rates are weighted according to the age structure of a (fictitious) »standard population«, and the weighted rates from all age groups are then summed. This results in an age-standardised rate that demonstrates the number of new cases or deaths per 100,000 individuals expected in a population with the age structure of the selected standard population. The »old European standard population« was used to calculate the figures in this report.

Cancer incidence and mortality risks

Age-specific incidence and mortality rates can be interpreted as a measure of the age-specific risk of developing or dying from cancer within a particular year. In order to ensure that this risk is communicated clearly, this report presents the age- and sex-specific risk of developing or dying from cancer within the next ten years or within the remaining expected lifetime. This information is represented both as percentages of age- and sex-specific populations as well as one case per N persons in the corresponding population. These figures also account for »competing risks«, for example, the risk that a 75-year-old man has of dying within the next ten years for reasons other than cancer. The »lifetime risk« – the overall risk of developing cancer throughout a person’s lifetime – was calculated in a similar manner. Only the current
rates (incidence, mortality and general life expectancy) are included in these calculations. No prognoses are provided on how these rates may change in the future. In addition, these rates should be treated as average values for the German population; individual risks may differ substantially due to the presence or absence of cancer risk factors. The DevCan software program, developed by the National Cancer Institute in the US, was used to perform the calculations.

International comparison
Current age-standardised incidence and mortality rates from Germany’s neighbouring countries as well as from England, Finland, Sweden and the US provide an international context for Germany’s cancer statistics. The appendix (Chapter 5.5) lists the sources for these data and whether the data refers to a different calendar period. As the international data were not checked for plausibility or completeness, it is possible that they underestimate rates, particularly regarding incidence. In addition, some countries report slightly different diagnosis groups from those used in this report, thus further limiting comparability (footnotes are provided in these cases).

Average age at diagnosis and death
The median age at cancer diagnosis and death are based on the incidence estimates and the official cause of death statistics available from the Federal Statistical Office. Since these are only available for five-year age groups, an approximation formula was used to calculate the median ages.

Mortality
Cancer mortality was calculated using the annual number of deaths from cancer reported in the official cause of death statistics. These statistics report the underlying cause of death stratified by age group and sex. Mortality rates are calculated as the number of annual deaths per 100,000 residents. This report provides the absolute number of deaths as well as crude and age-standardised mortality rates (using the old European standard population) from 1999 through 2017. The official cause of death statistics are produced by the Federal Statistical Office (www.gbe-bund.de).

Projected cancer incidence in 2020
In order to forecast the number of cases in 2020, current diagnosis-specific incidence rate trends were analysed by sex and age group. The trends were identified using the Joinpoint method, which uses regression models to locate points in time at which statistically significant changes in temporal trends occur (=joinpoints=). The projection was produced by extrapolating the average annual change since the last joinpoint to the year 2020. The resulting age- and sex-specific rates for 2020 were converted into absolute numbers using population figures from the 13th coordinated population projection (variant 4) produced by the Federal Statistical Office. Current trends in prostate cancer and in breast cancer among women aged 50–74 are strongly influenced by use of screening tests (Mammography screening for breast cancer and the PSA-test for prostate cancer). Therefore, it is unrealistic to expect current trends for these diagnoses to continue into the year 2020. Instead, age-specific incidence was assumed to remain constant in these cases, and projections of incident cases between 2016 and 2020 consider only the effects of demographic change.

Regional comparison
Age-standardised mortality rates (old European standard population) for 2015 and 2016 by federal state are presented in graphs alongside nationwide mortality rates. Age-standardised incidence rates for 2015 and 2016 are also depicted by federal state and are based on the number of recorded cases. For comparison, the graphs show corresponding nationwide incidence estimates for Germany. As stated in Chapter 2.1, different periods have been used to depict incidence in the federal states in eastern Germany, Berlin, and Bremen. Incidence from states with estimated completeness below 90% (<80% for malignant melanomas and <70% for thyroid cancer) are shown with more lightly coloured bars.

Crude rates
Crude incidence and mortality rates are calculated by dividing the total number of new cancers in a given period (incidence) or the number of deaths from cancer (mortality) by the total number of women or men in the respective population (in this case, Germany’s resident population). The result is given as the number of cases or deaths per 100,000 residents per year. In contrast to the age-standardised rate, this rate is strongly dependent on a population’s age structure.

Survival rates
This report describes the average survival rate beginning at cancer diagnosis for those 15 years and older at diagnosis. Absolute and relative survival rates up to 10 years after diagnosis are reported. Absolute survival rates represent the proportion of patients who are still alive at a particular time after diagnosis. An
absolute 5-year survival of 80%, for example, means that 80 out of 100 people with a specific cancer were still alive five years after diagnosis.

Relative survival rates, on the other hand, provide an estimate of how much the cancer diagnosis impacts survival. These are calculated as the ratio of absolute survival among cancer patients to the expected survival rate in the general population of the same age and sex. A relative 5-year survival rate of 100%, for example, means that during the first five years after diagnosis, the same number of deaths occurred among people with cancer as would be expected in a similar population without cancer. Relative survival rates are always higher than absolute survival rates. Expected survival rates were calculated using the Ederer II method and are based on data from the life tables published by the Federal Statistical Office.

Predetermined quality criteria were fulfilled by the registries in Hamburg, Lower Saxony, Brandenburg, Mecklenburg-West Pomerania, Saxony, Thuringia, Saarland, and the administrative district of Münster (North Rhine-Westphalia). Data from these regions were included for the calculation of current survival rates.

Relative 5-year survival rates by tumour stage (and sex) are reported here for the first time. Data from the registries stated above – with the exception of Saarland – were used for these calculations. Survival analyses were stratified by other factors for certain diagnoses, such as leukaemias and lymphomas. The period method was used to ensure that estimates of survival were as up-to-date as possible. This method takes into account the survival of people alive during a certain period of time (in this case, between 2015 and 2016).

The given ranges of 5- and 10-year survival represent the lowest and highest cancer survival estimates from the regions included in the calculations. Only those regions with a standard error for estimated survival of less than 7.0 were considered for determining the range. No range is reported when this criterion was met by less than four regions. Current findings suggest that only a small amount of the variance indicated by this range is due to differences in quality of care. In addition to random fluctuations, differences in data quality and proportion of DCO cases may play a larger role in these differences, especially in smaller federal states. Furthermore, methodological differences between the registries can also have an impact on the results; this particularly applies to DCO follow-back, which was not undertaken by all registries.

10-year survival rates are based on substantially smaller case numbers than 5-year survival rates. For this reason, registry-specific 10-year survival rates show a greater level of statistical uncertainty than 5-year survival rates. As a result, in some cases, the values within the range of relative 10-year survival rates may be slightly above the corresponding values for 5-year survival.

Overall, the estimated survival rates in Germany, at least in the case of cancers with unfavourable prognoses, may be slightly overestimated, although this is probably the case with most results that have been published internationally.

Tumour stage distribution
The extent of solid malignant tumours diagnosed between 2015 and 2016, as well as survival by cancer stage, were assessed using the TNM classification (7th edition). UICC stages (I to IV) are presented in this publication for the first time. In addition to the size or extent of a primary tumour (T), these figures also take into account lymph node status (N) and any distant metastasis (M). The UICC stages were assigned using the SEER TNM Registrar Staging Assistant (version 1.9) (https://staging.seer.cancer.gov/tnm/home/1.9/). Missing information for M was treated as M0 (no metastases), whereas missing information for N generally resulted in no reported UICC stage. To evaluate incidence by stage, data from all registries except Saarland were used.

5- and 10-year prevalence
The 5-year and 10-year prevalence estimates reflect the number of people alive at a particular time (in this case, 31 December 2016) who had developed a new case of cancer within the previous five (or ten) years, i.e., between 2012 (or 2007) and 2016. Prevalence estimates (by age, sex, cancer type and calendar year) were calculated from nationwide incidence estimates using the Pisani method, with absolute survival rates calculated using the Kaplan-Meier method with data from the regions listed under »survival rates«.

Additional analyses
Additional analyses are available for some cancer sites in this report and on the website of the German Centre for Cancer Registry Data (www.krebsdaten.de/english). These include figures on histology and more precise tumour sites. Unless otherwise stated, these evaluations are based on data from all registries.