2 Methodological aspects

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

The usefulness of population-based data on cancer largely depends on the extent to which all new cases of cancer are registered. Since 2010, the German Centre for Cancer Registry Data (ZfKD) has estimated the completeness of the data collected by the epidemiological cancer registries in all federal states. The estimate is made with the help of an internationally recognized indicator, the ratio of mortality to incidence. For a particular cancer diagnosis, this ratio (the M/I index) should be fairly constant from region to region, provided that diagnostics and therapy and thus also the survival prospects of cancer patients do not differ significantly between regions. With the help of the M/I index in a reference region that is assumed to be complete, and using mortality figures from the region in which registration completeness is to be estimated (the study region), the expected incidence for the study region is estimated and compared with the number of cases actually collected there. Cases only identified via death certificates (DCO) were not included in these calculations. The completeness of each registry in the reference region is also estimated by comparison with the respective expected values.

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
▶ Average completeness for all cancers combined over the last ten years above 90% and over 80% for each individual year
▶ Average proportion of DCO cases (cases registered only via death certificate) for all cancers combined below 15% over the last ten years or from the sixth year after the start of registration

The composition of the reference region has been slightly modified several times in recent years based on the data situation; currently it consists of the cancer registries of Bavaria, Bremen, Hamburg, Lower Saxony, North Rhine-Westphalia, Saarland and Schleswig-Holstein.

According to the principle described above, sex-specific expected values are calculated for six age groups and 16 (for women) or 15 (for men) diagnosis groups.

If age-specific mortality in the study region averaged fewer than five deaths per year, the modelled incidence rate in the corresponding age group in the reference region was used to calculate the expected number of new cases instead of the quotient of incidence and mortality. The estimated completeness for each diagnosis group results from the quotient of the observed and expected case numbers summed up over age group and sex. The completeness for all cancers combined is in turn estimated by summing the observed and expected values for all diagnosis groups and calculating their quotient.

Limitations of the described procedure arise foremost when the cancer-specific mortality is low overall or in relation to the corresponding incidence (testicular cancer, malignant melanoma, thyroid cancer), or when the ratio of mortality to incidence in fact differs between regions. This can be the case, for example, if utilization of cancer screening differs among the federal states or, as in the case of mammography screening, an early detection program is introduced at different times. A regionally varying distribution of tumour stages at diagnosis or different subtypes of a cancer diagnosis (for example in the case of thyroid cancer) can also lead to incorrect estimates.

According to the current completeness estimate for diagnosis year 2018, eight cancer registries in western Germany achieved over 95% completeness. Completeness in two other registries is between 90% and 95%. It was not yet possible to provide sufficient data from the Joint Cancer Registry for the new federal states and Berlin (GKR) for the years 2016 to 2018 (see Chapter 1.2). Therefore, on the basis of the data available to the ZfKD, no reliable statements can currently be made about the completeness of registration in these federal states. However, it can be assumed that in the course of the nationwide conversion to comprehensive clinical-epidemiological cancer registration, a sufficient completeness of cancer registration has already been or will soon be achieved in these federal states as well.

2.2 Estimating cancer incidence in Germany

The nationwide cancer incidence presented here comprise results partly from the counted cases and partly from the results of a mixed Poisson regression model. To estimate nationwide incidence, registries and diagnosis years were identified in which the following quality criteria regarding ›all cancers‹ were fulfilled: State-wide coverage for at least 10 years, and an estimated completeness of at least 90% over

the last five years. In addition, the annual DCO proportion in these five years must not exceed 15%. In years in which registries fulfilled these quality criteria, incident cases registered there were used as reference data for the regression and were directly counted for the nationwide incidence estimate. In years in which registries did not meet the quality criteria, incident cases were estimated using the regression model.
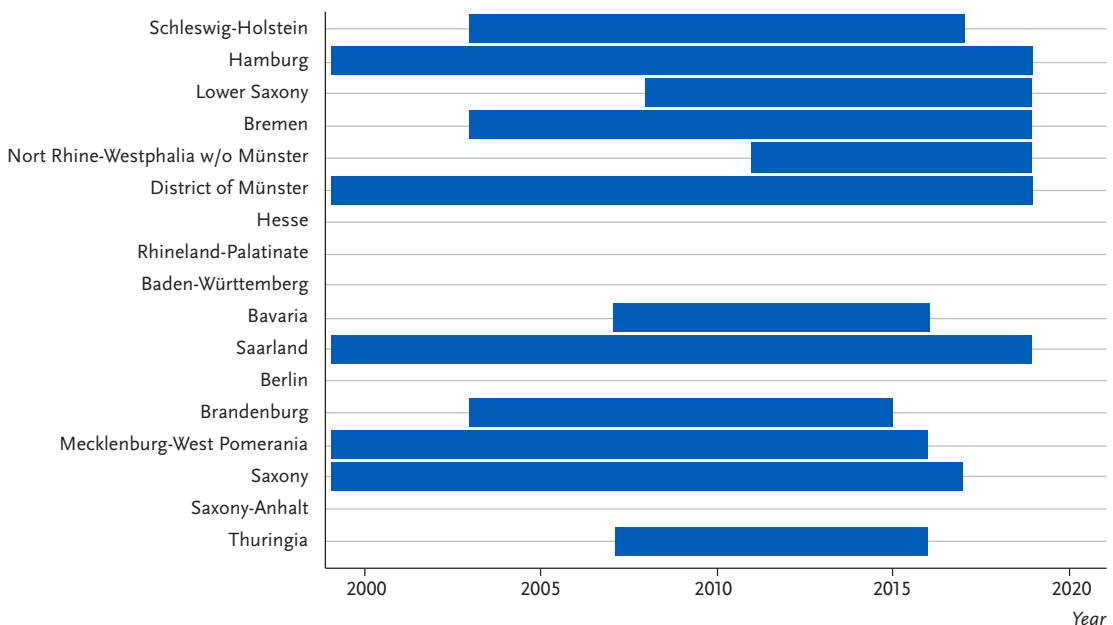
In the regression, incidence was modelled by cancer-specific mortality, population size and year of diagnosis. In addition, differences between incidence rates in the registries were accounted for by registry-specific parameters (axis intercepts as random effects) in the model. The regression was fitted to the reference data, stratified by sex, diagnosis and age group. Four registries did not yet meet the above quality criteria for any year: Baden-Württemberg, Berlin, Hesse and Saxony-Anhalt. For these registries, the incidences were first estimated for the years 2011 to 2013 using the ratio of incidence to mortality in the reference data and the mortality in the respective federal state. As with the reference data, these estimated incidences were used to fit the regression parameters so that specific characteristics of the registries could be considered.

The nationwide annual incidence is thus the sum of the counted cases from registries that fulfilled the quality criteria in the respective diagnosis year and the estimated incidence in the other registries. The proportion of counted cases in the nationwide incidence will increase as more registries meet the criteria for completeness, allowing a smooth transition from »estimating« to »counting« with an improvement in data quality.

The incidence of non-melanoma skin cancers (ICD-10 C44) was estimated with the same approach. However, due to the low mortality, there are no estimates for the completeness of coverage for these diagnoses. The reference region consisted of six registries (Schleswig-Holstein, Lower Saxony, North Rhine-Westphalia, Rhineland-Palatinate, Saarland and Mecklenburg-Western Pomerania) whose data were considered to be complete, at least for one period of time. Due to the data situation, nationwide incidence estimates were limited to the period 2006 to 2018. Estimated non-melanoma skin cancer incidence is associated with overall more uncertainty than the results for other cancers. Incidence of all cancers combined (Chapter 3.1), does not include non-melanoma skin cancer, as is customary internationally.

Figure 2.2.1
Time intervals for which cancer registries provided reference data on incidence for all cancer sites except non-melanoma skin cancer (C44)

2.3 Indicators and data presentation

The measures and graphical representations used in the results chapters are explained below.

Age-specific rates

The age-specific rate is calculated by dividing the number of cancer cases or deaths from cancer in a certain age group by the corresponding number of women or men of that age in the population. The graphical representation of these rates, stratified by sex, shows the correlation between age and incidence or mortality. The age-specific incidence rates are given as annual incident cases per 100,000 residents of the respective age group.

Age-standardised rates

As the presentation of age-specific incidence rates for women and men in this report shows, cancer incidence rates generally increase considerably with increasing age. Therefore, if one wants to compare incidence or mortality in different countries and regions or in the same population at different points in time, differences in the age structure of the populations to be compared must first be compensated for with the help of age standardisation. For this purpose, the observed age-specific rates are first weighted according to the proportion of a selected (fictitious) ›standard population‹ in the respective age groups. Then the weighted rates are summed up across all age groups. The age-standardised rate calculated in this way indicates how many incident cases or deaths per 100,000 persons would be observed in a population if it had the same age structure as the selected standard population. In this report, the ›old European standard population‹ was used.

Incidence and mortality risks

Age-specific incidence and mortality rates can also be interpreted as a measure of the age- and sex-specific risk of developing or dying from a specific cancer within one year. In order to make this form of risk communication more descriptive, the 10-year and lifetime risks of developing or dying from a certain cancer were calculated as a function of age and sex. In addition to the usual representation in percent, the results are also given as one per N persons of the same age and sex. So-called »competing risks« were included, i.e. it was taken into account that, for example, a 75-year-old man may, with a certain probability, die due to a cause other than cancer within the next ten years. Similarly, the lifetime risk was calculated, i.e. the risk of developing cancer within a person's remaining expected lifetime. Only current incidence and mortality rates as well as overall life expectancy are included in the calculations; possible future developments of these values were not considered.

Furthermore, these results are to be seen as average values for the population in Germany; individual risks can differ considerably due to the presence or absence of relevant risk factors. The programme ›DevCan‹[1], developed by the National Cancer Institute in the USA, was used to calculate these risks.

International comparison

In order to assess the estimated cancer incidence and cancer mortality in Germany in an international context, current age-standardised incidence and mortality rates from Germany's neighbouring countries as well as from England, Finland, Sweden and the USA were used. References for data sources can be found in the appendix (Chapter 5.5), where any deviating time periods are also noted. International data were included without checking plausibility or completeness; an underestimation of some rates seems possible, especially regarding incidence. Some countries group ICD-10 codes somewhat differently from groups used in Germany, which may limit comparability (see corresponding footnotes).

Median age at diagnosis and death

Median ages at cancer diagnosis and at death were calculated using an approximation formula from the incidence estimates and from the official cause of death statistics of the Federal Statistical Office, which are only available for 5-year age groups.

Mortality

Cancer mortality is based on the annual number of deaths due to cancer according to the official cause of death statistics. For this purpose, each death is assigned an underlying cause of death, and nationwide age- and sex-specific statistics are published. The mortality rates are calculated by dividing the annual number of deaths by the size of the population and are presented per 100,000 persons. In this report, the absolute number of deaths as well as crude and age-standardised mortality rates (old European standard) are reported from 1999 to 2019. The data source is the official cause of death statistics of the Federal Statistical Office (www.gbe-bund.de).

Projected cancer incidence in 2022

In order to estimate the number of cancer cases for the year 2022, current trends in age- and sex-specific incidence rates were analysed for each diagnosis presented in this report. For this purpose, the joinpoint method was used, in which regression models are used to identify points in time at which a statistically significant change in temporal trends (›joinpoints‹) occurs. For the projections, the average annual change since the last trend change was extrapolated to 2022. The age- and sex-specific rates for 2022 determined in this way were multiplied by

the corresponding figures from the 14th coordinated population forecast of the Federal Statistical Office (Variant 2) in order to calculate the projected numbers of incident cases.

Regional comparison

The estimated age-standardised incidence rates (old European standard) in the federal states for the period 2017 to 2018 are shown in comparison with the corresponding estimates for Germany; for the reference regions the rates reflect the reported incidence (see Chapter 2.2). For the same period, age-standardised mortality rates by diagnosis and sex are presented for all federal states in comparison to nationwide mortality.

Crude rates

For a given cancer diagnosis and population, the crude rate of incidence or mortality is calculated by dividing the total number of cancer cases (incidence) or the total number of deaths due to cancer (mortality) in a given time period by the total number of all women or men in the respective population (here: resident population of Germany). The result is given as the number of cases or deaths per 100,000 residents per year. In contrast to the age-standardised rate, the crude rate is strongly dependent on the age structure of a population.

Survival rates

The results of the survival analyses in this report describe average survival prospects after a cancer diagnosis for people over 15 years old at the time of diagnosis. Absolute and relative survival rates from 1 to 10 years after diagnosis were calculated. Absolute survival rates represent the proportion of patients who are still alive a certain length of time after their diagnosis. For example, an absolute 5-year survival of 80 % means that 80 out of 100 people with a certain type of cancer survived the first five years after their diagnosis.

Relative survival rates, on the other hand, depict cancer-related mortality by calculating the quotient of the absolute survival of cancer patients and the expected survival of persons the same age and sex in the general population. A relative 5-year survival of 100 % thus means that within five years of diagnosis, just as many persons with cancer have died as would have been expected without this diagnosis. The relative survival is always higher than the corresponding absolute survival. The expected survival was calculated with the so-called Ederer II method using the German period mortality tables of the Federal Statistical Office.

On the basis of previously defined data quality criteria, data from Schleswig-Holstein, Hamburg, Lower Saxony, Saarland as well as from the adminis-

trative district of Münster (North Rhine-Westphalia) were included in the current survival time calculations.

Relative 5-year survival rates by tumour stage (and sex) are also presented. For these analyses, only those cases were included whose tumour stage had been coded according to the seventh edition of the TNM classification. For some diagnoses (e.g. leukaemias and lymphomas), other characteristics were chosen for stratification.

In order to estimate as up-to-date survival prospects as possible, the so-called period method was used. This method considers the survival of persons who lived during a certain period of time (here: 2017 to 2018).

The presented ranges of 5- and 10-year survival show the respective lowest and highest survival in the individual regions included. For these results, only regions with a standard error of estimated survival of less than 7.0 were considered. If this criterion was met by fewer than four regions, the range was not shown. This range is most likely only to a very small extent indicative of potential differences in the quality of care: Differences in data quality or in the proportion of DCO cases may play a role, as may fluctuations due to chance, especially in the smaller federal states. Methodological differences between the registries, especially the trace-back of DCO cases (›follow-back‹), which is not carried out everywhere, can also influence the results. Substantially fewer persons with cancer are alive 10 years after diagnosis than 5 years after diagnosis. For this reason, registry-specific 10-year survival has a greater statistical uncertainty than 5-year survival. Therefore, the values in the ranges of relative 10-year survival may be slightly higher than the corresponding values for 5-year survival.

Overall, it can be assumed that survival rates for Germany are slightly overestimated, at least for cancers with an unfavourable prognosis, although this probably also applies to most internationally published results.

Distribution of tumour stages

The spread of solid malignant tumours at diagnosis in 2017 to 2018 by sex was evaluated using the TNM classification. Since the 8th edition of the TNM classification was published in 2017 and contains far-reaching changes for some diagnoses, stage distribution was shown separately for the 7th and 8th editions. The proportion of cases with missing stage and the proportion of DCO cases cannot be reliably assigned to a specific edition, so these proportions were calculated using all cases diagnosed in 2017 and 2018; only cases with valid stages were included in the figures. In addition to the size or spread of the primary tumour (T), the UICC stages (I to IV) shown also consider the lymph node status (N) and any distant metastasis (M). UICC stages were assigned using

the SEER TNM Registrar Staging Assistant Version 1.9 (https://staging.seer.cancer.gov/tnm/home/1.9/). Here, missing information on M is evaluated as M0 (no metastases), while missing information on N leads to a missing UICC stage in most cases. The proportion of cases with missing stage also includes those cases for which no TNM or UICC stage is intended due to the histology; this pertains to sarcomas, among other cancers. For the distribution of tumour stage, the data from all registries except the Joint Cancer Registry of the new federal states and Berlin were included.

Prevalence (up to 25 years after diagnosis)

Prevalence refers to the number of people who are alive at a given point in time (here 31 December 2018) and have previously been diagnosed with cancer. For example, the 5-year prevalence only considers people who have been diagnosed with cancer within the last 5 years. The prevalence was estimated according to Pisani's method[2] using the incidence estimate for Germany (see Chapter 2.2) and survival rates derived according to the Kaplan-Meier method. To estimate the 25-year prevalence, the incidence estimate had to be extended to the year 1993. Survival rate estimates excluded data of persons who had most likely died, but whose death or date of death was not known to the cancer registry (e.g. after moving to another federal state). Furthermore, survival rates were corrected for the proportion of incident cases for which information was only available from death certificates (Death Certificate Only or DCO cases)[3]. Due to the corrections, all registries except Berlin and Saxony-Anhalt could be included in survival estimates.

Further analyses

Additional analyses, for example a breakdown of incidence by histology or more precise tumour location, can be found for some cancers in this report or on the website of the Centre for Cancer Registry Data (www.krebsdaten.de/english). In this edition, these analyses are based on data from all population-based cancer registries except the Joint Cancer Registry of the new federal states and Berlin (GKR).

References

1 DevCan: Probability of Developing or Dying of Cancer Software, Version 6.7.9. Surveillance Research Program, Statistical Methodology and Applications, National Cancer Institute, 2012. http://surveillance.cancer.gov/devcan/

2 Pisani, P., Bray, F., & Parkin, D. M. (2002). Estimates of the world-wide prevalence of cancer for 25 sites in the adult population. Int J Cancer, 97(1), 72–81. doi:10.1002/ijc.1571

3 Dahm S., Bertz J., Barnes B. & Kraywinkel K. (2018) A mixed linear model controlling for case underascertainment across multiple cancer registries estimated time trends in survival. Journal of Clinical Epidemiology, 97C, 2018, pp. 123–133