CHAPTER 2
Processing and presentation of the data

PROCESSING OF THE DATA

The data used to create the tables presented in this book were extracted from the database of the African Cancer Registry Network (AFCRN). A list of individual anonymized cases with the following variables was extracted for each contributing registry:

- registration number that identifies the patient or the case
- sex
- ethnic group or race (optional)
- age
- date of incidence
- anatomical site of the tumour
- morphology of the tumour
- behaviour of the tumour
- basis of diagnosis.

All data had been coded according to the *International Classification of Diseases for Oncology, 3rd edition* (ICD-O-3) (Fritz et al., 2000). They were processed by the IARC software packages DEPedits and IARCcrgTools (Ferlay et al., 2005) for validation. After validation, the records were converted to *International Statistical Classification of Diseases and Related Health Problems, 10th revision* (ICD-10) coding (WHO, 1992) for presentation purposes. Because all contributing registries used the CanReg system, a software program developed at IARC for use by population-based cancer registries, the data had already undergone the same edits as those performed by the IARCcrgTools programs. This simplified and expedited the data validation process.

PRESENTATION OF THE DATA

The main sets of tables in this book present data on age-specific and age-standardized incidence, either by covered population (i.e. by cancer registry) or as summary tables by anatomical site.

Tables of incidence by registry

Average annual population at risk: The AFCRN database contains data on the population at risk by sex and age for each registry for as many years as possible. A denominator corresponding to the period of the incident cases (person-years at risk) was estimated based on this information, using intercensal estimates and postcensal projections as necessary. For each registry, the average annual population at risk for the period analysed is presented in a population pyramid within the description of the registry. The numbers (by 5-year age group) are also shown at the foot of the incidence tables for each registry.

The age-specific incidence tables: The numbers presented in the body of these tables are the numbers of cancer cases registered during the corresponding period by sex, anatomical site, and age group, along with summary rates of incidence. An example is shown in Fig. 2.01 (p. 5). The column headings are defined as follows:

Site: A shortened version of the full ICD-10 title describing each anatomical site or site grouping

All ages: The total number of cases (i.e. in patients of all ages) by site

Age unk: The number of cases in patients whose age is unknown; these cases are included in the total (i.e. all-ages) number of cases and in the calculation of the crude average annual incidence rate; they are also taken into account in the computation of the world age-standardized and cumulative incidence rates

MV%: The percentage of microscopically verified cases; i.e. the proportion of cases known to be diagnosed by a microscopic method (either histology or cytology), expressed as a percentage of all cases registered, including cases with unknown patient age or unknown basis of diagnosis

Age group (years): The number of cancer cases registered by age group, using the 5-year age bands indicated (0 to < 5 years, 5 to < 10 years, etc., up to 70 to < 75 years and 75+ years)

Crude rate: The crude average annual incidence rate, calculated by dividing the total number of cases (including those with unknown patient age) by the corresponding population at risk (i.e. all males or all females), expressed per 100 000 person-years

%: The proportional frequency of cases occurring at each site compared with the total number of cases at all sites except C44 (other skin)

CR 74: The cumulative incidence rate up to and including the age of 74 years, expressed per 100 000 person-years; this is the sum of the age-specific incidence rates over each year of age from birth through the age of 74 years; this cumulative rate is computed using 5-year age bands (0 to < 5 years, 5 to < 10 years, etc., up to 70 to < 75 years) and adjusted to account for cases with unknown patient age (Parkin et al., 1997)

ASR (W): The world age-standardized incidence rate, expressed per 100 000 person-years; this is calculated by the direct method, using the world standard population and 5-year age bands (0 to < 5 years, 5 to < 10 years, etc., up to 70 to < 75 years and 75+ years) and adjusted to account for cases with unknown patient age (Bray & Ferlay, 2014)

ICD-10: The ICD-10 code(s) corresponding to the anatomical site or site grouping listed in the left-most column

Average annual population: These numbers, listed at the foot of the table, indicate the average annual population at risk (in each 5-year age group) for the

period analysed. To calculate the incidence rate per 100 000 person-years for a particular age group, anatomical site, and sex, divide the pertinent number of cancer cases by the corresponding average annual population and the number of years for which the applicable data are presented, and multiply the result by 100 000.

Data quality indicators tables

The data quality indicators tables included in this volume (Chapter 5, p. 114) present the values (by anatomical site, covered population, and sex) for two indicators of data quality:

1. the percentage of cases that were microscopically verified (MV%); and
2. the percentage of death-certificate-only cases (DCO%), i.e. cases for which no information source other than a death certificate mentioning cancer could be found. Only a few of the cancer registries included in this volume have access to routinely processed death registry data, which is noted where applicable in the registry descriptions in Chapter 4 (p. 13).

Summary tables

The summary tables presented in Chapter 6 (p. 136) summarize the numbers of cases, the world age-standardized incidence rates, and the cumulative incidence rates up to and including the age of 74 years, by anatomical site, covered population, and sex.

Fig. 2.01. An example of an age-specific incidence table

Country, registry (year)

Number of cases by age group and summary rates of incidence: males

Site	All ages	Age unk	MV %	DCO %	Age group (years) 0-	5-	10-	15-	20-	25-	30-	35-	40-	45-	50-	55-	60-	65-	70-	75+	Crude rate	%	CR 74	ASR (W)	ICD-10
Mouth	212	0	89	0	-	9	4	6	6	2	12	11	15	22	29	27	26	15	12	16	2.6	5.2	0.91	**7.5**	C00-06
Salivary gland	25	0	80	4	-	-	1	-	1	-	1	3	2	5	2	-	2	2	4	2	0.3	0.6	0.15	**1.0**	C07-08
Nasopharynx	203	0	91	1	-	-	13	17	24	16	9	21	21	32	16	11	11	9	2	1	2.5	5.0	0.46	**4.2**	C11
Other pharynx	53	0	57	26	-	-	-	2	1	-	3	5	3	5	6	3	11	4	3	7	0.7	1.3	0.24	**2.1**	C09-10, C12-14
Oesophagus	319	0	77	7	-	-	-	3	2	5	13	22	30	18	44	34	31	39	28	50	4.0	7.9	1.70	**14.1**	C15
Stomach	225	0	71	12	1	1	1	2	1	6	8	6	27	28	30	24	23	28	19	20	2.8	5.6	1.22	**9.2**	C16
Colon	223	0	68	9	-	1	1	1	4	7	16	12	18	33	18	31	26	20	13	22	2.8	5.5	1.00	**8.3**	C18
Rectum	136	0	79	1	-	-	1	1	1	6	4	19	8	15	16	12	20	12	11	10	1.7	3.4	0.68	**5.2**	C19-20
Anus	7	0	100	0	-	-	-	-	-	-	-	-	1	1	-	-	2	1	-	2	0.1	0.2	0.03	**0.4**	C21
Liver	196	0	72	17	1	-	3	6	1	4	16	15	15	25	18	22	23	13	16	18	2.4	4.8	0.91	**7.2**	C22
Gallbladder etc.	35	0	77	9	-	-	-	-	-	-	1	2	1	-	6	2	5	6	7	5	0.4	0.9	0.30	**2.0**	C23-24
Pancreas	85	0	55	9	-	-	1	-	-	-	3	3	3	9	7	15	12	7	8	17	1.1	2.1	0.45	**4.0**	C25
Larynx	81	0	83	1	-	-	-	-	-	-	-	3	4	10	8	14	11	10	8	13	1.0	2.0	0.49	**3.9**	C32
Trachea, bronchus, and lung	116	0	72	9	-	-	-	-	1	3	2	3	6	14	11	16	17	16	12	15	1.4	2.9	0.72	**5.5**	C33-34
Bone	123	0	73	2	2	5	12	22	16	11	5	6	5	7	8	5	10	4	3	2	1.5	3.0	0.30	**2.7**	C40-41
Melanoma of skin	18	0	94	6	-	-	-	-	1	-	-	1	1	3	3	1	-	4	2	2	0.2	0.4	0.12	**0.8**	C43
Non-melanoma skin	112	0	90	4	1	2	1	3	5	6	11	11	10	9	14	8	10	6	8	7	1.4		0.44	**3.5**	C44
Mesothelioma	1	0	100	0	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	0.0	0.0	0.00	**0.0**	C45
Kaposi sarcoma	175	0	85	15	1	5	5	2	2	15	27	31	31	24	9	7	4	4	2	6	2.2	4.3	0.30	**3.2**	C46
Connective and soft tissue	83	0	82	5	6	10	6	5	11	8	5	5	5	5	3	3	6	1	3	1	1.0	2.0	0.19	**1.7**	C47, C49
Breast	72	0	74	8	-	1	-	2	-	-	4	2	8	12	6	5	15	5	5	7	0.9	1.8	0.35	**2.9**	C50
Penis	11	0	91	0	-	1	-	-	-	1	-	1	2	2	-	-	1	-	-	3	0.1	0.3	0.01	**0.4**	C60
Prostate	683	0	75	6	-	-	-	-	1	-	4	1	5	15	44	66	101	111	133	202	8.5	16.8	5.57	**43.9**	C61
Testis	10	0	80	0	1	1	-	1	-	3	1	1	-	-	1	-	-	-	-	1	0.1	0.2	0.01	**0.2**	C62
Kidney and renal pelvis	43	0	84	7	8	2	-	-	3	-	2	2	3	5	2	6	1	-	7	2	0.5	1.1	0.21	**1.4**	C64-65
Bladder	70	0	79	9	1	-	1	-	-	-	-	-	3	5	4	11	9	7	17	12	0.9	1.7	0.60	**4.0**	C67
Ureter and other urinary	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.0	0.0	0.00	**0.0**	C66, C68
Eye	136	0	90	1	28	7	3	3	7	7	10	7	23	14	6	4	4	4	4	5	1.7	3.4	0.29	**2.8**	C69
Brain and nervous system	94	0	66	7	7	3	5	7	3	5	8	3	5	6	8	5	8	9	4	8	1.2	2.3	0.35	**3.0**	C70-72
Thyroid	20	0	85	0	-	-	-	-	1	-	4	1	3	2	1	2	1	2	2	1	0.2	0.5	0.10	**0.7**	C73
Hodgkin lymphoma	56	0	100	0	2	15	6	8	2	4	3	2	6	2	2	1	-	-	1	2	0.7	1.4	0.07	**0.9**	C81
Non-Hodgkin lymphoma	208	0	94	4	7	14	14	13	5	10	16	18	20	21	20	9	16	10	8	7	2.6	5.1	0.63	**5.3**	C82-85, C96
Multiple myeloma	54	0	93	7	-	-	-	1	-	-	1	4	3	9	10	5	10	4	4	3	0.7	1.3	0.28	**2.1**	C90
Lymphoid leukaemia	65	0	100	0	12	17	11	2	-	1	1	2	3	2	2	1	8	2	1	-	0.8	1.6	0.15	**1.4**	C91
Myeloid leukaemia	43	0	98	2	3	5	4	2	3	2	3	2	4	5	3	-	2	1	2	2	0.5	1.1	0.11	**1.0**	C92-94
Leukaemia, unspecified	29	0	69	31	1	1	4	1	3	2	2	-	1	3	3	1	3	-	1	3	0.4	0.7	0.07	**0.8**	C95
Other and unspecified	144	0	56	17	4	5	7	3	7	1	4	12	10	6	14	23	11	10	12	15	1.8	3.6	0.66	**5.3**	O&U
All sites	4166	0	78	7	86	105	104	113	112	125	199	237	305	374	375	374	440	366	362	489	52.0		20.05	**162.6**	C00-96
All sites except C44	4054	0	78	7	85	103	103	110	107	119	188	226	295	365	361	366	430	360	354	482	50.6	100.0	19.60	**159.2**	C00-96 exc. C44
Average annual population					199381	151900	119951	115772	211089	234596	182623	134459	89109	65901	41682	24304	15061	7358	4462	5772	1603420				

For definitions and explanations of the terms and abbreviations used in this table, see the corresponding text in this chapter.