Chapter 17
The role of cancer registries

17.1 Aims of cancer registries

The cancer registry is an organization for the systematic collection, storage, analysis, interpretation and reporting of data on subjects with cancer. There are two main types of cancer registry: hospital-based and population-based cancer registries.

Hospital-based cancer registries are concerned with the recording of information on the cancer patients seen in a particular hospital. The main purpose of such registries is to contribute to patient care by providing readily accessible information on the subjects with cancer, the treatment they received and its result. The data are used mainly for administrative purposes and for reviewing clinical performance. Although these data may be used, to a certain extent, for epidemiological purposes (see Section 17.7), these registries cannot provide measures of the occurrence of cancer in a defined population because it is not possible to define their catchment populations, that is the populations from which all the cases arise.

Population-based cancer registries seek to collect data on all new cases of cancer occurring in a well defined population. Usually, the population is that which is resident in a particular geographical region. As a result, and in contrast to hospital-based registries, the main objective of this type of cancer registry is to produce statistics on the occurrence of cancer in a defined population and to provide a framework for assessing and controlling the impact of cancer in the community. Thus, the emphasis is on epidemiology and public health.

The uses of population-based cancer registration data may be summarized as follows:

(1) They describe the extent and nature of the cancer burden in the community and assist in the establishment of public health priorities.

(2) They may be used as a source of material for etiological studies.

(3) They help in monitoring and assessing the effectiveness of cancer control activities.

Some of these functions can be fulfilled using mortality data derived from vital statistics systems. Cancer registration data, however, provide more comprehensive, more valid and more detailed information on patient characteris-

tics than can be obtained from death certificates. Moreover, reliable cause-specific mortality data are available in most developed countries but in only a few developing countries. Thus, cancer registries may be the only way of obtaining information on the burden and patterns of cancer in developing countries, as well as providing a focus for research into etiology and prevention.

The discussion in the rest of this chapter will focus on population-based cancer registries unless otherwise specified.

17.2 A brief history of cancer registration

The first serious efforts to estimate the number of new and existing cancer cases in a given population were made at the turn of the century in various European countries. In Germany, an attempt was made in 1900 to register all cancer patients who were under medical treatment. Questionnaires were sent to every physician in the country to record the prevalence of cancer on 15 October 1900 (Anon., 1901). The same approach was adopted between 1902 and 1908 in Denmark, Hungary, Iceland, the Netherlands, Portugal, Spain and Sweden. These efforts were not very successful, however, mainly due to poor collaboration by the physicians. Similar surveys were conducted in the United States of America.

The first population-based cancer registry was set up in Hamburg (Germany) in 1926. Three nurses visited hospitals and medical practitioners in the city at regular intervals. They recorded the names of new cancer patients and transferred data to a central index in the health department. This index was compared once a week with official death certificates. Other population-based cancer registries were set up in subsequent decades, so that by 1955, almost twenty had been established in various countries (Table 17.1).

At present, more than 200 population-based cancer registries exist in various parts of the world. They cover about 5% of the world's population, but the proportion is much greater in developed countries than in developing ones. Moreover, in developing countries, registries are more likely to cover urban areas, where access to diagnostic and treatment services is better.

Nationwide cancer registration operates in some countries such as England & Wales, Scotland, the Nordic countries, Canada, Australia, New Zealand, Israel, Cuba, Puerto Rico and The Gambia. The Danish Cancer Registry, founded in 1942, is the oldest functioning registry covering a national population. In most countries, however, population-based cancer registries cover only a proportion of the population (e.g., Colombia, India, Italy, United States). Some specialized registries that cover only the registration of specific age-groups (e.g., childhood cancers in Oxford, UK) or particular cancer sites (e.g., gastro-intestinal cancers in Dijon, France) have also been established. In addition, hospital-based cancer registries have been set up in a large number of hospitals worldwide.

The International Association of Cancer Registries (IACR) was formed in 1966. The main objective of this association is to develop and standardize the collection methods across registries to make their data as comparable as possible.

A more detailed account of the history of cancer registration is given in Wagner (1991).

Table 17.1.
Population-based cancer registries established before 1955.[a]

Country (region)	Year of establishment	Notification
Germany (Hamburg)	1929	Voluntary
USA (New York State)	1940	Compulsory
USA (Connecticut)	1941 (registered cases retrospectively back to 1935)	Compulsory (since 1971)
Denmark	1942	Compulsory (since 1987)
Canada (Saskatchewan)	1944	Compulsory
England and Wales (SW Region)	1945	Voluntary
England and Wales (Liverpool)	1948	Voluntary
New Zealand	1948	Compulsory
Canada (Manitoba)	1950	Voluntary
Slovenia	1950	Compulsory
Canada (Alberta)	1951	Compulsory
USA (El Paso)	1951	Voluntary
Hungary (Szabolcs, Miskolc, Vas)	1952	Compulsory
Norway	1952	Compulsory
Former USSR	1953	Compulsory
Former German Democratic Republic	1953	Compulsory
Finland	1953	Compulsory (since 1961)
Iceland	1954	Voluntary

[a] Reproduced with permission from Wagner (1991).

17.3 Cancer registration methodology

The aim of a population-based cancer registry is to collect information on every case of cancer identified within a specified population over a given period of time. To ensure this, it is necessary to guarantee that the following basic requirements are fulfilled before setting up a population-based cancer registry:

(*a*) Clear definition of the catchment population. The registry should be able to distinguish between residents of the area and those who have come from outside and it should be able to register cases in residents treated outside the area.

(*b*) Availability of reliable population denominators from the census or other statistical offices.

(*c*) Generally available medical care and ready access to medical facilities, so that the great majority of cancer cases will come into contact with the health care system at some point in their illness and, therefore, will be correctly diagnosed.

(*d*) Easy access to case-finding sources such as hospitals, pathology departments, death certificates and other sources of clinical data within the catchment area and in the surrounding areas.

17.3.1 Data collection

The way in which a registry operates depends, inevitably, on local conditions and on the material resources available. Usually, the main sources of information of a population-based registry include: (1) information from treatment facilities, such as cancer centres and major hospitals (and sometimes, if appropriate, private clinics, hospices, homes for the elderly and general practitioners); (2) information from diagnostic services, especially pathology departments, but also haematological, biochemical and immunological laboratories, X-ray and ultrasound departments, and other imaging clinics; (3) death certificates from the death registration system (if they are available).

The information is collected from these sources by either active collection or passive reporting. *Active collection* involves registry personnel actually visiting the different sources and abstracting the data on special forms. This is the usual method in registries in developing countries. *Passive reporting* involves health-care workers completing the notification forms developed and distributed by the registry, or sending copies of discharge abstracts to the registry. A mixture of both procedures, with an emphasis on the latter, is followed in most registries in developed countries. In certain countries, notification of cancer cases is compulsory, although this does not necessarily ensure completeness.

The data items to be collected by a registry are dictated by the purpose for which the registry has been established, by the method of data collection used and by the resources available to the registry. However, the emphasis should be on the *quality of the data collected rather than on the quantity*. It is advisable that registries in developing countries should start by attempting to collect only information on the basic items listed in Table 17.2.

A unique *registration number* (cancer registry number) is assigned by the registry to each patient. If a patient has more than one primary tumour, the same number is given to each tumour. Multiple primaries are then distinguished on the basis of their incidence date and their topography and morphology.

Other *identification items* such as name, sex and date of birth (or, approximate age, if the date of birth is not known) are important to avoid multiple registrations of the same patient or tumour, to obtain follow-up data and to conduct any type of record linkage. Patient's usual address is essential for establishing the residence status, to exclude all non-residential patients, to conduct analysis by area of residence and for follow-up of the patients. Data on ethnicity is important in populations containing distinct ethnic groups.

The *incidence date* is primarily the date of first consultation or admission to a hospital or clinic for cancer, as this is a definite, consistent and reliable point in time which can be verified from records. This date is chosen as the anniversary date for incidence calculations and as the starting date for survival analyses (see Section 17.6.2). If this information is not available, the incidence date should be taken as the date of first diagnosis by a physician or the date of the first pathological report. A special problem arises when

Table 17.2.
Basic data items to be collected by population-based cancer registries.[a]

Item	Comments
The patient	
Personal identification	
Registration number	Assigned by the registry
Name	According to local usage
Sex	
Date of birth or age	Estimate if not known
Demographic	
Address	Usual residence
Ethnic group	If relevant
The tumour	
Incidence date	
Most valid basis of diagnosis	Non-microscopic or microscopic
Topography (site)	Coded using ICD-O[b]
Morphology (histology)	Coded using ICD-O
Behaviour	Coded using ICD-O
Source of information	Type of source: physician, laboratory, hospital, death certificate or other Actual source: name of physician, laboratory, hospital, etc. Dates (e.g. dates of relevant appointments, hospital admissions, diagnostic procedures)

[a] Modified from MacLennan (1991).
[b] *International Classification of Diseases for Oncology* (Percy *et al.*, 1990).

cancer is first ascertained from a death certificate and attempts to follow back are unsuccessful. The date of death of such 'death certificate only' (DCO) cases should be taken as their incidence date.

Information on the *most valid basis of diagnosis* is of great interest in assessing the quality of the registration data. The minimum requirement of a cancer registry is to discriminate between tumours that were microscopically verified and those which were not. If possible, further information should be obtained to distinguish neoplasms that were diagnosed on the basis of a clinical history only, clinical history plus other investigations (e.g., X-ray), exploratory surgery, autopsy, cytology, etc. For future checking purposes, it is important that the registry collects data on the source(s) of case-finding (e.g., name of physician, hospital, laboratory), dates of relevant medical events (e.g., hospital admission, biopsy) and any other details that will help to trace the patient's medical records (e.g., hospital number, biopsy number, laboratory reference number).

Inclusion of data items other than those listed in Table 17.2 increases the complexity and cost of the registration process and, hence, should be done only if justified by local needs and if the necessary resources are available. A list of optional items is given in Table 17.3; the most relevant ones are clinical extent of disease before treatment (stage at presentation) and follow-up data.

The data from the various case-finding sources are usually abstracted by using a standard registration form developed according to the needs of the

Table 17.3.
Optional items of information which may be collected by population-based cancer registries.[a]

The patient
Identification
Personal identification number (e.g., national identity number or social security number)
Demographic and cultural items
Place of birth
Marital status
Age at incidence date
Nationality
Religion
Occupation and industry
Year of immigration
Country of birth of father and/or mother
The tumour and its investigations
Certainty of diagnosis
Method of first detection
Clinical extent of disease before treatment
Surgical-cum-pathological extent of disease before treatment
TNM system
Site(s) of distant metastases
Multiple primaries
Laterality
Treatment
Initial treatment
Follow-up
Date of last contact
Status at last contact (alive, dead, emigrated, unknown)
Date of death
Cause of death
Place of death

[a] Modified from MacLennan (1991).

registry. Two main considerations should be kept in mind when developing a registration form:

(1) The information on cancer cases should be collected and classified so that it accords with the data available from the census or other statistical offices. This is fundamental to ensure comparability between the numerators (i.e., numbers of cancer registrations) and the relevant denominators (i.e., population figures) in the calculation of incidence rates.

(2) Although data should be collected (and reported) according to local needs and interests, an effort should be made to ensure that comparisons with data from other national and international cancer registries will be possible.

17.3.2 Classification and coding of neoplasms

As mentioned in Appendix 2.2, it is recommended that cancer registries use the *International Classification of Diseases for Oncology* (ICD-O) (Percy *et al.*, 1990) to code the topography (site of primary tumour) and morphology (histological type) of the tumours. The fifth digit in the ICD-O morphology codes describes the behaviour of the tumour—benign, borderline, *in situ*, malignant. The topography of a tumour is the most important data item recorded and provides the main basis of tabulation of registry data.

17.3.3 Data quality

Two main issues should be considered when evaluating the quality of the data in a cancer registry: its *completeness* and its *validity*. A population based-registry should, by definition, register every single case that occurs in its catchment population. However, case ascertainment is rarely complete. Various methods, such as comparisons with death certificates and hospital records, have been used to determine the degree of completeness of registration. It is also important to ascertain the extent to which the registry eliminates registrations of cases from outside the catchment population and avoids multiple registrations of the same person or of the same tumour.

The validity of the data can be assessed in various ways. The proportion of cases with microscopic verification of diagnosis is a very useful index, as is the proportion registered during life (not simply from a death certificate). Cancer registries should develop their own internal quality control checks so that attention is drawn to missing information and inconsistent data. Many registries frequently re-abstract and re-code a sample of cases to assess the quality of their data. A full discussion of quality control methods is given by Parkin *et al.* (1994).

17.3.4 Reporting of results

The collection of information on cancer cases and the production of cancer statistics are only justified if use is made of the data collected. A population-based cancer registry should make its data and findings available in the form of reports and articles in scientific journals. The reports should include background information on the registry, registration procedures, catchment population, degree of data completeness and validity, methods of analysis and findings. Basic statistics should be produced and presented for diagnostic entities mainly according to topography of the tumour. The data should be presented in tabular and graphical form. Examples are given in Figure 17.1 and Table 17.4.

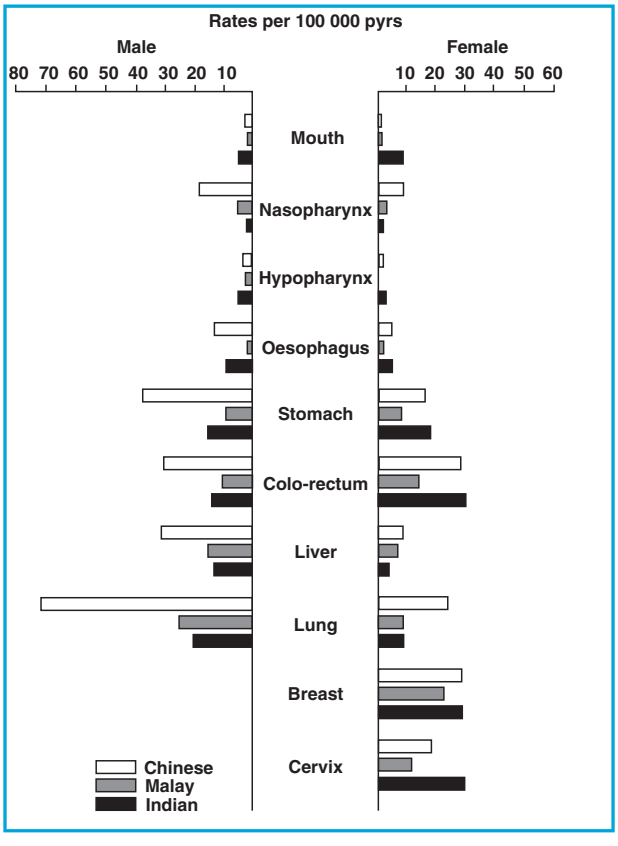

Figure 17.1.
Age-standardized incidence rates (to the world population) for selected cancer sites by sex and ethnic group, Singapore, 1978–82 (reproduced with permission from Lee *et al.*, 1988).

Site (ICD–9)	Number of cases by age group									Total	%	Incidence rate	
	Unknown	0–	15–	25–	35–	45–	55–	65–	75+			Crude	ASR[b]
All sites	10	69	89	255	241	266	362	264	74	1630	100.0	101.1	238.5
All sites but skin	8	69	88	253	236	264	359	259	72	1608		99.8	234.6
Oral cavity (140–145)	1	–	1	1	2	5	2	3	1	16	1.0	1.0	2.5
Nasopharynx (147)	–	–	1	5	1	1	1	4	–	13	0.8	0.8	2.0
Other pharynx (148–149)	–	–	1	–	–	–	2	1	–	4	0.2	0.2	0.6
Oesophagus (150)	–	–	–	1	16	25	63	35	13	153	9.4	9.5	30.4
Stomach (151)	–	–	–	2	10	15	14	20	6	67	4.1	4.2	13.5
Colon (153)	–	–	5	2	6	6	9	9	2	39	2.4	2.4	6.6
Rectum (154)	–	–	2	3	6	8	4	4	1	28	1.7	1.7	3.8
Liver (155)	–	2	10	22	37	41	46	46	9	213	13.1	13.2	34.6
Pancreas (157)	–	–	–	1	2	7	13	7	1	31	1.9	1.9	5.9
Larynx (161)	1	–	–	–	–	4	12	4	2	23	1.4	1.4	4.5
Bronchus, lung (162)	–	–	–	1	6	30	50	32	6	125	7.7	7.8	24.6
Pleura (163)	–	–	–	–	–	–	1	–	–	1	0.1	0.1	0.1
Connective tissue (171)	–	4	4	4	2	2	1	–	–	17	1.0	1.1	1.1
Melanoma of skin (172)	–	–	2	2	2	2	4	1	1	14	0.9	0.9	1.8
Other skin (173)	2	–	1	2	5	2	3	5	2	22	1.3	1.4	4.0
Breast (175)	–	–	–	–	1	3	1	–	–	5	0.3	0.3	0.6
Prostate (185)	3	–	–	–	2	11	37	41	18	112	6.9	6.9	29.2
Penis (187)	–	–	–	–	1	4	3	2	3	13	0.8	0.8	2.8
Bladder (188)	–	1	–	3	5	19	18	16	6	68	4.2	4.2	13.2
Kidney (189)	–	10	1	–	1	2	1	2	–	17	1.0	1.1	1.7
Eye (190)	–	5	1	1	–	1	1	1	–	10	0.6	0.6	0.9
Brain, nervous system (191–192)	–	7	6	4	5	2	4	1	–	29	1.8	1.8	2.4
Thyroid (193)	–	–	1	1	1	1	4	1	–	9	0.6	0.6	1.2
Hodgkin's disease (201)	–	2	1	4	2	2	2	–	–	13	0.8	0.8	1.0
Non-Hodgkin lymphoma (200, 202)	–	12	4	13	11	10	6	2	–	58	3.6	3.6	4.7
Multiple myeloma (203)	–	–	–	1	5	4	8	1	1	20	1.2	1.2	2.7
Lymphoid leukaemia (204)	–	8	3	1	1	1	2	4	–	20	1.2	1.2	2.5
Myeloid leukaemia (205)	–	8	6	6	5	6	2	1	–	34	2.1	2.1	2.7
Other leukaemia (207–208)	–	–	–	–	–	–	1	1	–	2	0.1	0.1	0.6
Kaposi's sarcoma	2	7	28	171	97	44	27	4	–	380	23.3	23.6	24.6
Other and uncertain	1	3	11	4	9	8	20	16	2	74	4.5		

[a] Reproduced, by permission of Wiley-Liss Inc., a subsidiary of John Wiley & Sons Inc., from Bassett *et al.* (1995).
[b] ASR = Incidence rate age-standardized to the world population.

Table 17.4.
Example of the type of table used by cancer registries to report their data. Number of cancer registrations among African men resident in Harare, 1990–92. Harare Cancer Registry, Zimbabwe, 1990–92.[a]

17.4 Cancer registration in developing countries

It might seem that cancer registration should not be regarded as a priority for the health services of a developing country, given all the competing demands upon the limited resources allocated to health. However, cancer is already a significant health problem in many developing countries. More than half of the new cancer cases in the world occur in developing countries (Parkin *et al.*, 1993). The rapid increase in life expectancy (largely because of a reduction in mortality from infectious disease) together with the adoption of western lifestyles suggest that the burden of cancer in these countries is likely to increase in the near future.

Most often cancer registries provide the only opportunity of properly assessing the extent and nature of the cancer burden in developing countries, since very few of them have reliable cause-specific mortality data. Ideally, the objective should be to establish a population-based cancer registry which will be able to estimate the incidence of different tumours in a well defined community. However, because of the relative ease with

which they can be established, cancer registries in developing countries often start on the basis of cases attending certain hospitals or departments of histopathology.

Population-based cancer registries in developing countries usually face enormous logistic problems due to lack of appropriately trained personnel and adequate resources. In addition, their success may be jeopardized by external factors beyond their control.

Lack of basic health services

The functioning of a cancer registry relies heavily on the availability of proper health services for diagnosis and treatment of cancer cases. In many developing countries, however, health facilities are scanty and tend to be concentrated in urban areas. For individuals seeking medical attention, the quality of diagnostic information may be poor and based on clinical examination only.

Lack of proper denominators

Population-based registries require information on the size and the nature of the population served, information which requires the availability of census data. Censuses are particularly difficult to conduct in developing countries, and so they tend to be conducted infrequently, and their results may become available late and with inadequate detail.

The population of many developing countries is particularly mobile because of the increasing tendency to migrate temporarily from rural areas to urban areas and because social and political circumstances may force whole communities to move from one area to another. Inter-censal estimates or post-censal projections of the population size and structure are, therefore, likely to be inaccurate.

These population changes present a special challenge to cancer registries which must make special efforts to distinguish residents from non-residents in their catchment area using, as far as possible, the same definitions as in the census.

Identity of individuals

The ability to distinguish individuals from events (e.g., hospital admissions) is a key feature of a cancer registry. Thus, the registry should have sufficient information on each individual to avoid multiple registrations of the same subject. The most universal and generally used identifier is the name. The utility of using names will vary depending on local custom. For instance, surname (or family name) may not be used—persons may be known only by their first name. Individuals may change their name when they get married or for other social reasons. Variations in spelling of names is a frequent problem, particularly if a large percentage of the population is illiterate. This is aggravated if there is a need to transliterate names to the Roman alphabet, in order to use computerized database systems.

Lack of follow-up

Active follow-up usually means that the registry attempts to contact physicians or patients on a regular basis to see if the patient is still alive. Because this is expensive, many registries rely on passive follow-up, matching with death certificates and assuming patients are alive otherwise. Mixed systems use death certificates plus updating the 'date last known alive' from hospital admissions, consultations, and other sources of data.

Active follow-up of the patients is usually very difficult in developing countries. Few registries have the necessary facilities for regular follow-up of patients. There are also problems with unreliable postal services, unstable addresses and mobility of the population. Passive follow-up is possible only in the few countries where a reliable death registration system exists.

17.5 The role of cancer registry data in epidemiology

Population-based cancer registries are important resources for cancer epidemiologists since they hold information on the distribution of cancer in well defined populations. This information may be analysed without the need for any additional data collection. Cancer site-specific *incidence rates* can be calculated and compared according to many different variables such as age, sex, country of birth, place of residence at the time of diagnosis, etc. Time-trend studies are also possible when data have been accumulated over long periods of time. The methods used in such analyses were discussed in Chapters 4 and 11. Systematic compilations of data from population-based cancer registries from all over the world are published in *Cancer Incidence in Five Continents* (Doll *et al.*, 1966; Waterhouse *et al.*, 1970, 1976, 1982; Muir *et al.*, 1987; Parkin *et al.*, 1992, 1997). These data are of great value for international comparisons.

In addition to incidence figures, population-based cancer registries that conduct adequate follow-up of their patients are able to estimate the *prevalence* of cancer. Prevalence figures give an indication of the burden of the disease in the community. Cancer registries generally assume that once diagnosed with cancer, an individual remains a prevalent case until death. Thus, prevalence may be estimated from data on incidence and survival. When a registry has been in operation for many years, so that all patients diagnosed with cancer before the establishment of the registry have died, the prevalent cases may simply be enumerated from the registry file, provided, of course, that the registry receives information on deaths and emigrations for all registered cases.

The cancer registry provides an economical and efficient method of ascertaining cancer occurrence in *intervention trials* (Example 17.1) and *cohort studies*, as long as the cancer patients are properly identified in their files so that case matching can be performed.

Population-based registries can also provide a source of cases for *case–control studies*. However, in general, cancer registries are not regarded as well suited for the conduct of these studies because of delays in registration. The main value of the registry is rather to evaluate the completeness and representativeness of the case series.

The registry may, however, carry out its own case–control studies using its database, comparing one type of cancer with a selection of the other cancers ('controls') (see Section 11.1.6). The variables usually available for these analysis are limited to those routinely collected by the registry. Registries may supplement these variables with additional information (e.g., smoking, diet,

Example 17.1. *The Gambia Hepatitis Intervention Study is a large-scale vaccination trial in The Gambia, initiated in July 1986, in which about 60 000 infants received a course of hepatitis B vaccine and a similar number did not. New cases of liver cancer will be ascertained through the nationwide cancer registration scheme (Gambia Hepatitis Study Group, 1987).*

Example 17.2. *The importance of some selected risk factors in the etiology of oesophageal cancer in Bulawayo, Zimbabwe, was assessed using data collected by the local cancer registry during the years 1963–77, when an attempt was made to interview all cancer patients using a standard questionnaire. Risk factors for oesophageal cancer were estimated by case–control analysis in which other non-tobacco- and non-alcohol-related cancers were taken as the 'control' group. Table 17.5 shows the analysis for men. There was a strong association with tobacco use, with an apparent dose–response effect. In contrast, alcohol intake appeared to have little effect on the risk of oesophageal cancer in this population (Vizcaino* et al.*, 1995).*

Table 17.5.
Risk factors for oesophageal cancer in men, south-western Zimbabwe, 1963–77.[a]

	Cases	Controls[b]	Odds ratio (95% confidence intervals)[c]
Tobacco use			
Non-smoker[d]	120	947	1.0
Ex-smoker	21	38	3.4** (1.9–6.2)
< 15 g daily	279	542	3.5** (2.7–4.5)
≥ 15 g daily	71	91	5.7** (3.8–8.4)
Not specified	56	116	2.8** (1.8–4.2)
Test for trend			$P < 0.001$
Alcohol intake			
None[d]	144	654	1.0
Occasionally	44	206	0.6* (0.4–0.9)
Weekly	121	387	0.8 (0.6–1.1)
Daily	212	539	0.9 (0.7–1.2)
Not specified	41	68	1.8* (1.1–3.0)

[a] Data from Vizcaino *et al.* (1995)
[b] Formed by all other non-tobacco- and non-alcohol-related cancers (i.e., after exclusion of cancers of the oral cavity and pharynx, liver, larynx, lung and bladder).
[c] Adjusted for age, province, occupation and for the other variable in the table.
[d] Baseline category
* $P < 0.05$; ** $P < 0.001$.

occupation, treatment, etc.) by interviewing samples of patients (Example 17.2), by extracting such information from medical records, or by record linkage with other relevant records.

Cancer registries have been particularly useful in the conduct of case–control studies to investigate the carcinogenic effects of cancer treatments (Example 17.3).

Example 17.3. *A collaborative group of population-based cancer registries and major oncological centres carried out a case–control study to identify reasons for the observed increases in lung cancer risk following Hodgkin's disease. A total of 98 cases of lung cancer were identified in patients who had survived for at least one year following a diagnosis of Hodgkin's disease. For each case, three controls were selected from patients with Hodgkin's disease who did not develop subsequent lung cancer, matched to the case on registry or hospital, sex, year of birth and year of diagnosis of Hodgkin's disease. For both cases and controls, detailed information was abstracted from medical records concerning stage and treatment of Hodgkin's disease. Patients treated with chemotherapy alone had about twice the risk of developing lung cancer compared with those treated by radiotherapy alone or both modalities. There was also an increasing risk of lung cancer with increasing estimated radiation dose to the lung among patients treated with radiotherapy alone (Kaldor* et al.*, 1992).*

17.6 The role of cancer registries in cancer control

The cancer registry is an essential part of any rational programme of cancer control. Its data can be used in a wide variety of areas of cancer control ranging from etiological research, through primary and secondary prevention to health-care planning and patient care. Although most cancer registries are not obliged to do more than provide the basis for such uses of the data, they possess the potential for developing and supporting important research programmes making use of the information they collect.

17.6.1 Planning of cancer control programmes

Accurate information on cancer occurrence is important in fixing priorities and targeting cancer control activities. Population-based cancer registries are in a unique position to provide this information.

The annual numbers of incident cases provide an indication of the resources needed for primary treatment, and the number of prevalent cases describe how many people are in need of regular long-term follow-up (although for certain cancers, no regular surveillance is required beyond the first 5–10 years after diagnosis). Table 17.6 shows the numbers of incident (new) and prevalent (new and old) cancer cases in South-east England in 1992. The ranking of the cancer sites is quite different for incidence and prevalence. This is due to differences in survival. Cancers with a good survival have high prevalence even if their incidence is low, whereas those with

Site	ICD-10	No. of incident cases, 1992[b]	No. of prevalent cases at 31 December 1992[b]	Five-year relative survival (%)[c]
Lung	C33–34	6434 (1)	8201 (4)	9
Prostate	C61	4096 (2)	13 564 (3)	49
Colorectal	C18–21	3492 (3)	14 470 (2)	43
Bladder	C67	2183 (4)	16 538 (1)	70
Stomach	C16	1516 (5)	2407 (7)	13
Non-Hodgkin lymphoma	C82–85	1009 (6)	4582 (5)	51
Oesophagus	C15	858 (7)	805 (10)	8
Pancreas	C25	833 (8)	586 (11)	6
Kidney	C64	644 (9)	2296 (8)	40
Brain	C71	523 (10)	1558 (9)	22
Melanoma of skin	C43	367 (11)	2855 (6)	71
All malignant neoplasms (excluding non-melanoma skin cancer)		**28 732**	**109 637**	**36**

[a] Data from Thames Cancer Registry (1995).
[b] Ranking of sites by decreasing frequency is given in parentheses.
[c] Patients aged 15 years and over, diagnosed during the years 1986–89.

Table 17.6.
Number (and ranking) of male incident and prevalent cancer cases, and five-year relative survival ratios for selected sites. South-east England, 1992.[a]

poor survival have lower prevalence even if their incidence is higher.

Up-to-date cancer statistics provide information on the present burden of cancer to the health care system in a population. To develop long-term programmes for cancer control, it is necessary to predict what the needs will be in the future. In other words, it is necessary to have reliable estimates of the numbers of incident and prevalent cases that will occur in coming years. Cancer registries are an important source of data upon which to base such predictions. The simplest predictions of cancer incidence rates are based on continuing the present age-specific time trends into the future. The forecast can be imp-roved if birth cohort effects can also be taken into account by using age–period–cohort statistical models (see Section 4.3.2). An example is given in Figure 17.2.

An even more sophisticated approach to predic-

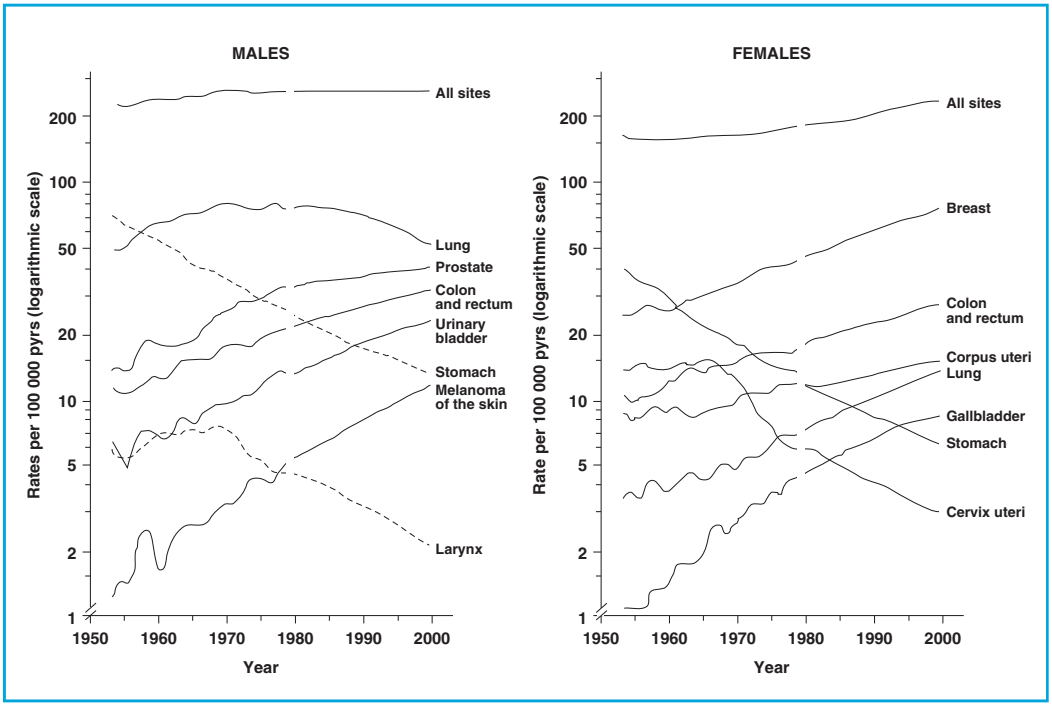

Figure 17.2.
Annual age-adjusted incidence rates of cancers at selected primary sites in Finland: actual rates from 1953 to 1979 and predictions up to the year 2000 based on a statistical model which includes age, period and cohort effects (reproduced with permission from Läärä, 1982).

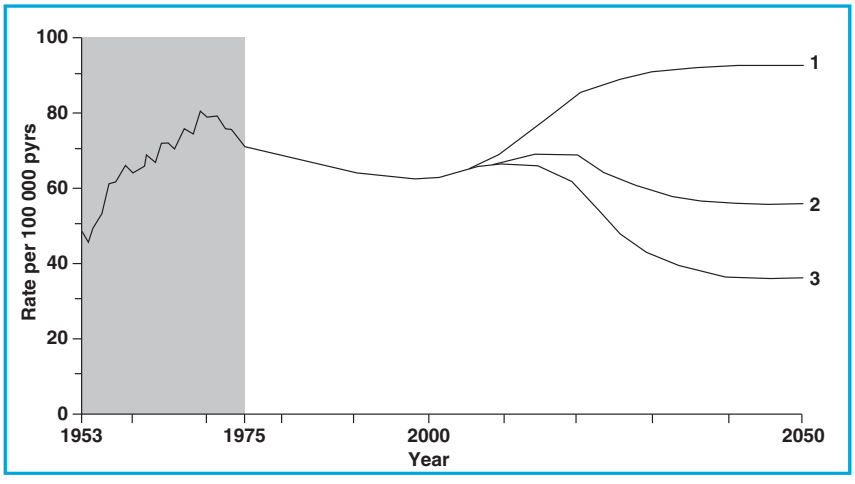

Figure 17.3.
Age-adjusted incidence rates (to the world population) of lung cancer in males in Finland in 1953–75, and forecasts for the rates in 1976–2050 derived from a simulation model based on the following assumptions: in each consecutive five-year period in 1976–2050, 10% of smokers will stop smoking, and one of the following three alternatives (three curves) holds true: (1) 60% of non-smokers aged 10–14 years, 30% of those aged 15–19 and 10% of those aged 20–24 will start smoking; (2) the percentages are 30, 15 and 5, respectively; and (3) the percentages are 15, 7.5 and 2.5, respectively (reproduced, by permission of Oxford University Press, from Hakulinen & Pukkala (1981)).

tion of future cancer incidence rates can be used if information on temporal changes on the prevalence of major risk factors is known and likely future changes can be predicted. This approach has been used to predict lung cancer incidence in relation to prevalence of smoking (as in Figure 17.3) and breast cancer in relation to changes in fertility.

Any predictions should be interpreted with caution, however, since they are based upon a considerable number of assumptions. To assess their robustness, it is advisable to provide forecasts under different possible scenarios, as in the example given in Figure 17.3.

17.6.2 Evaluation of cancer control programmes

Primary prevention

Cancer registries can play an important role in monitoring and evaluating the effectiveness of primary prevention measures. As mentioned in Section 16.2.2, trends in cancer incidence can be related to changes over time in exposure to risk factors. Occasionally, when implementation has been confined to one area, comparisons of the changes in the intervention area versus 'control' areas may be possible. It should be kept in mind when interpreting such relationships that it takes considerable time (generally decades) for the effect of a change in exposure to be reflected in cancer incidence data.

Screening and early detection

Cancer registries can play an important role in the evaluation and monitoring of screening programmes aimed at detecting pre-invasive conditions. Cancer registration data have been used in routine-data-based studies to examine trends in disease rates in relation to screening frequencies within a population and to compare disease rates between different populations with the coverage offered by their screening programmes (see Section 16.3.1). For instance, such studies have supported the hypothesis that regular use of the Pap smear test is effective in reducing the incidence of invasive cervical cancer.

Cancer registries can also contribute to the ascertainment of cancer occurrence in intervention trials and cohort studies designed to assess the value of screening programmes, and as an unbiased source of cases for case–control studies. The main issues to be considered in the design and interpretation of these studies were presented in Section 16.3.1.

When screening programmes are aimed at detecting early invasive cancers (e.g., breast cancer), reduction in mortality rather than incidence

should be the ultimate measure of their effectiveness. However, once a screening programme is known to be effective, cancer registries may help to monitor its performance by providing data on so-called 'intermediate outcome measures'. Absence of a change in such intermediate end-points indicates that the screening has not been effective. Suitable monitoring statistics from cancer registries are:

(*a*) the incidence of interval cancers (i.e., cancers detected between screening tests) as compared to the incidence in the screened population before screening was introduced;

(*b*) the stage distribution of screen-detected cancers compared to the distribution of non-screen-detected cancers. A lack of shift in stage distribution towards early stages indicates that the programme is not effective.

(*c*) if screening is effective, screen-detected cancers should show better survival than non-screened cases.

It should, however, be stressed that *intermediate end-points are subject to several forms of bias* and therefore they may suggest that the programme is effective even though mortality data do not. These issues were discussed in Section 16.3.1.

Tertiary prevention

Survival statistics can be produced by population-based cancer registries that follow up their cases, either actively or passively. Although survival analysis of data from population-based registries cannot evaluate specific treatments (this can be done only in clinical randomized trials), it provides a useful evaluation of cancer care in the area covered by the registry, since all cancer cases will be included regardless of the type of treatment they may have received.

The methods used in survival analyses are those discussed in Chapter 12. The first requirement for the application of these methods is a *clear and well defined case definition*. This should clearly specify the site of the cancer and/or histology, age and sex of the patient and, if available, the extent of disease (stage) at the time of diagnosis. The nature of the cases to be included should also be defined. For example, a decision must be taken on whether to include cases for which the most valid basis of diagnosis is solely clinical. A decision should also be taken regarding cases registered on the basis of a death certificate only (DCO), for whom no information is available on the date of diagnosis of the cancer. The most usual practice is to omit these cases from the analysis, but if they represent a large proportion of registrations, it may be better to present two survival analyses, one including DCO cases and another excluding them. In both cases, the proportion of DCO registrations should be stated in survival reports.

The second requirement is a *clear and well defined starting point*. For population-based cancer registries, the starting date (from which the survival is calculated) is the *incidence date* (see Section 17.3.1).

The third requirement is a *clear and well defined outcome*. Death is generally the outcome of interest, but some registries collect enough data to allow them to conduct analyses using recurrence of tumour, or first recurrence of a particular complication, as the outcome of interest. It is also necessary to formulate clear criteria for deciding who should be considered 'lost to follow-up'. For instance, certain registries would assume that cases for which it was not possible to obtain follow-up data for more than 15 months should be taken as 'lost to follow-up'.

There are several problems in the interpretation of time trends in survival. Firstly, improvements in survival may be due, at least in part, to better ascertainment and recording of incident cases. Secondly, if there has been a trend towards earlier diagnosis (e.g., through introduction of a screening programme), survival may improve but the gain may be due entirely to increased lead time, with no change in mortality rate (see Section 16.3.1).

Despite these caveats, time trends in survival are useful to assess the extent to which advances in treatment have had an effect in the population. For instance, the dramatic improvements in survival observed in clinical trials in the treatment of childhood cancers conducted in the 1960s do seem to have been transposed into the community in many developed countries, as the population-based survival from many of these cancers shows significant increases over time (Table 17.7).

Comparisons of cancer survival estimates derived from population-based cancer registries are increasingly used to compare the effectiveness of cancer treatment across populations. However, survival reflects not only treatment but also prognostic factors such as stage at diagnosis, histological type and other characteristics of the disease. When data on such factors are not available, or when their definition is not properly standardized across registries, the reasons for any variations observed cannot be properly identified.

Table 17.7.
Time trends in five-year survival risk for certain childhood cancers (0–14 years) in Great Britain.[a]

Cancer	Five-year survival risk (%)	
	1962–64[b]	1971–74
Lymphoid leukaemia	14	39
Hodgkin's disease	39	79
Non-Hodgkin lymphomas	20	25
Wilms' tumour	26	57
Malignant bone tumours	22	29

[a] Data from Draper *et al.* (1982).
[b] 1968–70 data for lymphoid leukaemia.

In Example 17.4, survival in Khon Kaen was equal to, or better than, that in the USA for stomach, liver and lung cancers. Thus, improvements in treatment may be of reduced benefit in the control of these cancers in Thailand compared with the potential benefits of primary and secondary prevention (e.g., control

of hepatitis B infection and liver fluke infestation for liver cancer; anti-smoking campaigns for lung cancer). In contrast, survival was lower for Khon Kaen residents than USA white residents for those cancers whose prognosis is associated with early diagnosis (breast, cervix and large bowel), indicating that interventions to promote early detection may provide potential benefits. Survival from leukaemia and lymphoma was also lower for residents in Khon Kaen, probably because of poor access to complex therapeutic regimens.

Example 17.4. *The Khon Kaen Cancer Registry in the north-east of Thailand is one of the few population-based cancer registries in developing countries that collects follow-up data. These data are obtained from clinical records, death certificates and return-paid postcards sent annually to each patient thought to be alive. A total of 10 333 residents of Khon Kaen province registered with cancer during the years 1985–92 were followed up to the end of 1993. Table 17.8 shows five-year relative survival ratios for selected cancer sites. These survival ratios were compared with age-standardized survival data from two developed countries—the USA and Scotland (Sriamporn* et al.*, 1995).*

Table 17.8.
Five-year relative survival ratios in Khon Kaen province (Thailand), USA and Scotland for selected cancer sites.[a]

Site	ICD-9	Khon Kaen, 1985-92	US whites, 1983-88[b]	US blacks, 1983-88[b]	Scotland, 1983-87[b]
Stomach	151	23.4	17.2	19.0	12.8
Large bowel	153–154	41.9	58.9	50.0	42.3
Liver	155	9.2	5.9	3.5	4.2
Lung	162	15.4	14.7	10.3	7.7
Breast (females)	174	48.1	78.4	61.5	66.8
Cervix uteri	180	60.1	69.2	56.8	61.0
Non-Hodgkin lymphoma	200, 202	32.5	56.6	49.7	53.2
Leukaemia	204–208	19.4	45.7	31.7	41.6

[a] Data from Sriamporn *et al.* (1995).
[b] Standardized to the site-specific age distribution of Thai subjects.

17.7 Hospital-based cancer registries

Hospital-based cancer registries are more numerous and widespread than population-based cancer registries. The primary purpose of these registries is to contribute to patient care by providing readily accessible information on the patients with cancer, the treatment they received and its results. The data may also be used for clinical research and, to a certain extent, for epidemiological purposes.

One of the main advantages of hospital registries is that they have ready and instant access to medical records, the primary source of cases. The data items collected by a hospital registry tend to be more extensive than those collected by a population registry. There are, however, several limitations to the data from hospital registries:

(1) They are institution-based and not population-based. This means that no attempt is made to register all cancer cases occurring in any defined population; thus incidence rates cannot be determined. Patients who are hospitalized in more than one hospital are counted more than once in an area's hospital tumour registries. Information may not be shared among hospitals caring for the patient at different times. Changes over time in numbers of any type of cancer or patient characteristics may only reflect shifts by patients (or doctors) from one institution to another. The cancer cases in any one hospital (or group of hospitals) may not be representative of all cancer cases that are occurring in the area. For instance, certain institutions are referral centres for specific types of cancer or for particularly difficult or extensive tumours.

(2) Ascertainment of death is likely to be more incomplete in hospital-based registries than population-based registries because of limited access to, and use of, other sources such as death certificates, and limited sharing of information among hospitals.

(3) In contrast to most population-based cancer registries, hospital registries make little attempt to standardize methods of data collection between them. It is therefore difficult to compare their findings.

Hospital cancer registries produce reports on the numbers of cancers seen in the hospital per year by site, age and sex. These results may be presented as *proportional incidence ratios* (i.e., the frequency of cancers of a particular site in relation to the total number of cancer cases—see Sections 4.3.5 and 11.1.6 for a discussion of this type of measure). They may also provide information on methods of diagnosis, stage distribution, treatment methods, response to treatment, and survival at an institutional level. The hospital registry data may also be used to forecast future demands for services, equipment and manpower in a given hospital.

Although these registries cannot provide incidence rates in the general population, they may be used for epidemiological purposes. For instance, case–control studies may be set up to investigate the etiology of a particular cancer by comparing the characteristics of cases with those of a control group; this control group may be formed by patients with other types of cancer or by other hospital patients. The analysis will be similar to that shown in Example 17.2.

Box 17.1. Key issues

- There are two main types of cancer registry:

  (a) Hospital-based cancer registries record information on all cancer patients observed in a particular hospital. Their main aim is to monitor and plan patient care at an institutional level. However, their data are of limited value for epidemiology, because it is not possible to define the population from which their cases arise.

  (b) Population-based cancer registries seek to collect data on all new cases of cancer which occur in a well defined population. As a result, and in contrast to hospital-based cancer registries, they can provide data on the occurrence of cancer in a particular population and, therefore, they are of particular value for epidemiology and public health.

- Population-based cancer registries play an important role in epidemiology by quantifying the incidence and prevalence of the disease in the community and as a source of ascertainment of cancer cases in intervention, cohort and case–control studies. Their data are also important in planning and evaluating cancer control programmes by helping to establish priorities and forecast future needs; by monitoring cancer occurrence in relation to the prevalence of important risk factors; by helping to assess and monitor the effectiveness of screening programmes; and by evaluating cancer care through survival statistics.

- The data items to be collected by a population-based cancer registry are determined by their aims, the data collection methods to be used, and the resources available. The emphasis should be on the quality of the data rather than their quantity. The completeness and validity of the data should be monitored regularly.

- Population-based cancer registries are particularly useful in developing countries where reliable cause-specific mortality data are rarely available.

Further reading

* Jensen *et al.* (1991) describe in great detail the planning of cancer registries in both developed and developing countries and the uses of registration data in epidemiology and public health planning.

* Parkin *et al.* (1994) provide practical recommendations on how population-based cancer registries can assess and monitor the quality of their data.