CANCER REGISTRIES: A WORLDWIDE ENDEAVOUR
Cancer occurrence has been documented for centuries by clinical and autopsy reports, and at the population level by data on causes of death. As Johannes Clemmesen, the founder of the first nationwide cancer registry, in Denmark, put it, “Favoured by its accessibility, and consequently by good conditions for diagnosis, mammary carcinoma was among the first neoplasms subject to valid statistical observation.” Clemmesen produced a very interesting comparison of the frequency of new cases of cancers of the breast and uterine cervix in Copenhagen, Denmark, in 1943–1957 with the frequency of deaths from the same two cancers as recorded by the city medical officer in Verona, Italy, in 1760–1839, the first known example of a proper statistical treatment of cancer data. The frequency in Verona was much higher than that in Denmark (perhaps due in part to an underestimate of the population), but the shapes of the two sets of curves show some interesting similarities.

In general, one would expect that cancer incidence rates and cancer mortality rates would be almost the same if cancer leads almost invariably to death, as was unfortunately the case until the middle of the 20th century. However, as effective treatments are introduced, mortality rates no longer reflect the incidence rates, and complete registration of new cases becomes indispensable to obtain a clear picture of the cancer burden in a population. Indeed, the development of cancer registration roughly paralleled the first tangible, although uneven, successes of cancer therapy – surgical, radiological, and drug treatments.
EARLY STEPS IN CANCER REGISTRATION, NATIONAL AND INTERNATIONAL

In Europe, the first register of cancers was established in Hamburg, Germany, as early as 1927, and similar regional initiatives were undertaken in the 1940s and 1950s in the United Kingdom and other countries. The nationwide registry in Denmark was instituted in 1942. In the USA, a series of surveys of cancer incidence were carried out in the late 1930s, and registration started in Connecticut and New York State in 1940.

Immediately after the Second World War, a group of European specialists interested in cancer statistics met in Copenhagen and recommended the implementation of cancer registration systems and the establishment of an international body to foster uniformity in terminology and classification as well as the correlation of the data obtained in each country. Based on this recommendation, in 1950 the World Health Organization (WHO) set up a subcommittee on the registration of cases of cancer as well as their statistical presentation. Clearly, the
time was ripe for such initiatives: in the same year, UICC organized a week-long symposium in Oxford dedicated to the geographical pathology and demography of cancer. The Oxford gathering led to subsequent meetings on this topic, and to the establishment of an ad hoc committee that in 1966 produced a technical report, the first volume of the Cancer Incidence in Five Continents series (see “Cancer Incidence in Five Continents, Volume I”).

CANCER INCIDENCE IN FIVE CONTINENTS, VOLUME I

Richard Doll, Peter Payne, and John Waterhouse were responsible for the first compilation of international statistics on the incidence of cancer, published in 1966 under the aegis of UICC. This extract from the preface to Volume I of Cancer Incidence in Five Continents by the three editors gives an insight into the background, state of the art, and collaborative climate of the day.

“The suggestion that cancer incidence rates for different parts of the world should be brought together in a single volume arose in discussion among members of the Geographical Pathology Committee of the International Union Against Cancer during a symposium in Mexico in 1964. [Note: Another topic discussed at that meeting was the “French proposal” for a new cancer organization, highlighted in the chapter “The birth of IARC”.] That there was a need for such a volume rapidly became apparent when the directors of cancer registries were asked for their opinion. Of those approached, all but one responded enthusiastically and immediately agreed to contribute. In the event, data have been collected from 32 cancer registries in 24 countries, and 39 scientists have contributed personally by describing the character of their registry and by collecting and submitting figures in a standard way.

“The form in which the book appears was suggested by a committee of 15 members, which met at the Ciba Foundation in London in May 1965, and the editors have been guided in their work by the results of the discussions that took place at that meeting. In a few instances, it has not been possible to follow the Committee’s advice, for reasons of finance, and the text which was written by the editors may, in some places, have inadvertently misrepresented the Committee’s views. The editors, therefore, take full personal responsibility for all defects in both style and scientific presentation.

“It is a pleasure to acknowledge the help that has been given in the collection of information about individual registries by the Cancer Unit of the World Health Organization, and particularly by Dr A. Tuyns, who took part in the meeting of the Editorial Committee.”
CANCER REGISTRIES, AN ESSENTIAL TOOL FOR RESEARCH AND PUBLIC HEALTH

A cancer registry can be defined as an organization – with its own premises, resources, and personnel – for the systematic collection, storage, analysis, interpretation, and reporting of data on people with cancer.

Hospital-based cancer registries are concerned with the recording of information on the cancer patients seen in a particular hospital. Their main purpose is to contribute to patient care by providing readily accessible information on the patients with cancer, the treatment they received, and the outcome of the treatment. The data are used mainly for administrative purposes and for reviewing clinical performance.

Population-based cancer registries seek to collect data on all new cases of cancer occurring in a well-defined population. Their main objective 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. A population-based cancer registry has at least three main uses.

The first is to describe the extent and nature of the cancer burden in the community and to assist in establishing public health priorities. These include preventive measures to reduce the observed burden of new cases as well as the provision of adequate health services, general and specialized, for the care of the expected number of cases.

The second main use is to be a source of material for epidemiological investigations of the causes of cancers – a necessary step for prevention. Registered cancers can be the cases in case-control studies (e.g. a comparison of the dietary habits of stomach cancer cases and controls), or a registry can be used to follow up a cohort of people (e.g. recording the lung cancer incidence in workers exposed to an airborne pollutant).

The third main use of a population-based cancer registry is to assist in monitoring and assessing the effectiveness of cancer control activities, in particular by examining area-based survival of treated cancer patients. Unlike hospital-based statistics, which are strongly influenced by the type and degree of severity of the cancer cases a hospital deals with, the area-based survival statistics provide an unbiased picture of the actual survival of all patients in a population as a result of the operation of the totality of the locally available health services.

IARC AND CANCER REGISTRIES AROUND THE WORLD

One might argue that the international variation in types and frequencies of cancers was the single observation that most defined the initial research focus of IARC. Certainly this emphasis was behind the creation in the late 1960s of IARC Regional Centres in populations where patterns of cancer were striking, such as in Singapore (see “Calum Muir”, and also the chapter “The birth of IARC”). A starting point for these centres was to establish reliable cancer registries. It was also not coincidental that John Higginson had been studying “geographical pathology” in Africa before being appointed as the first IARC Director, in 1966.
CALUM MUIR

Calum Muir joined IARC in 1966 as head of the Unit of Epidemiology. Aged 37, he had a sound background in human pathology, built up over more than a decade at the Department of Pathology of the University of Malaya (now the National University of Singapore). It was while in Singapore that Muir played an instrumental role in the creation of the IARC Regional Centre there, which started cancer registration in 1968. He had soon recognized the excellent opportunities to document and investigate differences in the incidence of disease among the various ethnic groups in Singapore, publishing several papers on heart disease and on a range of cancers. This line of research had convinced Muir of the essential importance of accurate registration of cancer cases, a task that became his first priority at IARC. He played an active role in the work leading to the publication of the first edition of the *International Classification of Diseases for Oncology* (ICD-O), in 1976, and he was an editor of the second edition, published in 1990.

As a pathologist, Muir did much to assemble the new morphological terms and the latest classifications for lymphomas, leukaemias, and brain tumours. He developed close contacts with cancer registries worldwide, through personal visits and meetings, and he was instrumental in founding the International Association of Cancer Registries in 1966, serving as deputy secretary from 1972 to 1990 and as president from 1992 until his death in 1995. (IARC has provided the secretariat for this nongovernmental organization since 1974.) In many countries, IARC became known – and achieved renown – thanks to Muir’s career-long efforts to expand the network of cancer registries and stimulate their engagement in epidemiological research. He contributed to this research himself, in particular with publications on the cancer patterns in different ethnic groups and migrant populations, on cancer trends, and on the proportion of cancers attributable to various causes. In the second half of the 1960s, IARC was a newborn institution, and Muir, an affable and kind person, will be remembered for contributing his critically needed competence and strength during the early years of epidemiology at IARC.
This focus on cancer registration was firmly established by the time the second volume of *Cancer Incidence in Five Continents* was published, in 1970. By then, IARC and UICC were jointly coordinating the project. IARC’s commitment to this field has been guided by two purposes: promoting a high and uniform quality of existing registries, and helping to establish new registries in the many areas of the world where none yet exist. The efforts have been focused on several areas: classification of cancers, data collection procedures, quality control procedures, education and training of personnel, worldwide expansion of registries’ coverage, and publication of international cancer incidence data. In addition, IARC conducts analyses of registry incidence data on a worldwide scale (see the chapter “Cancer patterns, trends, and burden”).

**Classification of cancers**

As early as 1968, IARC was asked by WHO to make recommendations about the content and structure of the neoplasm chapter in the ninth revision of the *International Classification of Diseases, Injuries, and Causes of Death* (ICD-9). Ever since then, IARC has been the key responsible party in the continuous process of updating ICD. In 1976, this role led to the publication of the *International Classification of Diseases for Oncology* (ICD-O), an expanded offshoot of ICD for the purposes of cancer classification. ICD-O classifies each tumour based on four criteria: the site of origin (e.g. the lung), the microscopic appearance (histology) of the tumour, the degree of differentiation of the tumour (e.g. a lung cancer may have a poorly differentiated “squamous carcinoma” histology), and the biological behaviour (i.e. whether the tumour is biologically aggressive and malignant or is instead benign).

The third edition of ICD-O (ICD-O-3, published in 2000 and dedicated to Calum Muir) uses the nomenclature of the WHO Classification of Tumours series (the “Blue Books”; see the chapter “Education and training of cancer researchers”). As these classifications are revised and new histological terms are introduced, ICD-O-3 is updated with new or modified codes and terms; the first revision was published in 2013. ICD-O has been published in a wide range of languages, and conversions from one version of a classification system to the next, or from one system to another (e.g. from ICD-O to ICD), can often be performed using dedicated software.
A bird’s-eye view of the complex interrelations of different schemes of classification and coding of tumours as they have evolved over time, from the sixth revision of the International Classification of Diseases (ICD-6) in 1948 (top left corner) until the current edition of the International Classification of Diseases for Oncology (ICD-O) and the Systematized Nomenclature of Medicine (SNOMED) in 2000 (bottom right corner).
Data collection procedures

Cancer registries collect information from three main sources: hospitals, laboratory services, and death certificates. Computerized hospital information systems or manual indexes of hospital discharges are the primary sources of data on patients and their diagnoses. Private hospitals and clinics in the area covered by the registry should be included among the information sources, as should hospices and palliative care services. The second key source of information is pathology laboratories, which provide the definitive histological diagnosis of a tumour. The third source is death certificates. These documents are important because they enable the identification of cancer cases that may have been missed by the other two sources, and make it possible to investigate the survival of cancer patients.

Which information should a cancer registry abstract from the records of these sources? IARC provides guidance mainly through two books: Cancer Registration: Principles and Methods, published in 1991, and Planning and Developing Population-Based Cancer Registration in Low- and Middle-Income Settings, published in 2014. Both publications contain relevant references to practical tools for cancer registration, in particular software packages for data processing and storage. The minimum set of information that should be reported by a registry includes data identifying the person and the key characteristics of the tumour.

Quality control procedures

Cancer registries are permanent structures, and population-based cancer registration is a continuous process. Built-in quality control procedures are necessary to ensure that the completeness and high quality of registry data are maintained over time. “Intelligent” computer terminals for data entry perform a series of automatic checks to ensure that inconsistent data are not accepted (in these cases, the registry refers back
to the data source for clarification and correction). An example of such an inconsistency is a reported date of tumour diagnosis that is earlier than the patient’s date of birth.

Once the data have been accepted, a second series of checks on the overall information make it possible to verify the completeness and accuracy of the data collected. For example, a high percentage of cases reported only when the person dies (i.e. via a death certificate) is indicative of poor quality of registration, whereas a high percentage of cases for which the diagnosis was based on microscopic verification of a tissue specimen is usually an indicator of good quality. Throughout the years, IARC has been instrumental in introducing and popularizing among registries data quality indicators, some of which have become criteria for inclusion of the incidence data from a registry in the publication *Cancer Incidence in Five Continents*.

**Education and training of personnel**

Since 1967, when the first IARC course on methods for cancer epidemiology was organized (see the chapter “Education and training of cancer researchers”), IARC has been deeply committed to the education and training of registry personnel, through courses and workshops, site visits by IARC scientists, and publications and software packages. Often, these activities have benefited from a close collaboration with the International Association of Cancer Registries.
Although all IARC courses on epidemiological methods include the topic of cancer registration, about 50 courses, held in Lyon and at sites around the world, have been dedicated to cancer registration. Since 2005, the first module of the IARC Summer School in Cancer Epidemiology, held in Lyon each year, has been devoted to cancer registration. The teaching of this particular module is now being transferred to the newly established IARC Regional Hubs for Cancer Registration.

CanReg5 is one of the software packages developed by IARC for cancer registries. CanReg5 is an open-source tool specially designed to input, store, check, and analyse population-based cancer registry data. For registries using other software, a conversion program is available. CanReg5 is available in Chinese, English, French, Portuguese, Russian, and Spanish, with online help in English. Training sessions on CanReg5 are systematically included in IARC’s cancer registration courses, which are presented face-to-face and, more recently, also in distance learning mode.

**Worldwide expansion of registries’ coverage**

The expansion of the global network of cancer registries ranks high among IARC’s achievements. Over the past 50 years, the number of registries and populations covered by registration has increased steadily, by a factor of 10. This success is an example of the rewards of the shared dedication and enthusiasm of very many participants, motivated by a common priority. This has been the case for the collaboration of individual registries (and often a single devoted
person at a registry), and of the registries collectively through the International Association of Cancer Registries, with IARC. Regional activities have also furthered the expansion of registries; a good example is the GRELL Network (www.grell-network.org), created in 1975 at the initiative of IARC epidemiologist Albert Tuyns and Luc Raymond of the Geneva Cancer Registry.

A defining feature of IARC’s approach to cancer registration has been, and continues to be, working directly alongside cancer registrars all over the world, assisting with the small practical details as well as the larger strategic issues. This has undoubtedly underpinned the willingness of registries to share their hard-won data with IARC and thus to permit the global picture of cancer occurrence to emerge.

Despite this remarkable progress, much remains to be done: only about 21% of the world’s population is covered by population-based cancer registries, with particularly sparse registration coverage in Asia (8% of the total population) and in Africa (11%). Even within these underrepresented regions, there are further disparities, with rural areas far less well covered than urban areas. However, to put these figures,
Schematic diagram of the steps involved in the production at IARC of Cancer Incidence in Five Continents (CI5), Volume X, from the call for data to multiple checks and through to the publication in a uniform format of the statistical data of each collaborating registry.
unsatisfactory as they are, into perspective, it is useful to remember that even for the certification of cause of death, a procedure that has a centuries-long history, only about one third of the world’s population is adequately covered.

To further boost cancer registration, IARC has moved into a new era by establishing a series of IARC Regional Hubs for Cancer Registration in Africa, Asia, the Caribbean, Latin America, and the Pacific Islands. These Regional Hubs in liaison with IARC will develop specific tools in support of registries, including the assessment of registry quality, the publication and presentation of data, the coordination of research projects, and continental data analyses and reports based on all the registries in a Hub region. This approach of devolved responsibility for the development of cancer registries harnesses local expertise, supported by the international standards and expertise of IARC. The IARC Regional Hubs form part of the Global Initiative for Cancer Registry Development (gicr.iarc.fr), launched by IARC in 2011 to “make cancer data count” in driving national cancer control policies.

**Publication of international cancer incidence data**

The publication of the first volume of *Cancer Incidence in Five Continents* (or, as the series has become known, CI5; ci5.iarc.fr) has been followed by nine further volumes. Volume X reports incidence data for the period 2003–2007 from 290 registries in 68 countries. Even in this computerized era, the acquisition of data from every registry, the processing of the data, and the production for each main cancer type of standard tables subdivided by sex and age group is a complex job, involving several cycles of communication between IARC and registry professionals to verify the completeness and accuracy of the data. On average, the preparation of a volume of *Cancer Incidence in Five Continents* takes about three years.

Childhood cancers, usually defined as those occurring before the age of 15 years, are distinct entities from cancers in adults, and IARC has a special project (*International Incidence of Childhood Cancer*, iicc.iarc.fr) dedicated to the acquisition and dissemination of data on cancer incidence in children. Collected incidence data on childhood cancer were published as *International Incidence of Childhood Cancer* volumes in 1988 (IARC Scientific Publication No. 87) and in 1998 (IARC Scientific Publication No. 144); a third volume is in preparation. In addition, the book *Epidemiology of Childhood Cancer*, by Julian Little (IARC Scientific Publication No. 149) was published in 1999.
Rationale

Cancers occurring in children are marked by low incidence rates, typical histologies, favourable survival in resource-rich countries, long-term survivorship issues and a range of ethical, psychological and societal concerns. These characteristics, combined with data presentation requirements unique to these cancers, make a strong case for oncologists and epidemiologists to study childhood cancers separately from other cancers.

As a part of its mission of collection and dissemination of data on cancer, IARC recognises the need for a specific approach to collecting and disseminating childhood cancer data. Following the publication of the two volumes of International Incidence of Childhood Cancer in 1988 (IIIC-1) [10] and 1998 (IIIC-2) [8], IARC is launching, in collaboration with IARC, the third monograph of the series. The new publication will fill the gap in the availability of international data on childhood cancer incidence, after the two previous volumes covering roughly the 1970s and 1980s, respectively.