



CHAPTER 1

Registry fundamentals

Contents

- I-2 Introduction**
 - I-2 General registration procedures
 - I-3 Population-based registries and special registries
 - I-3 Formalization**
 - I-3 Founding
 - I-3 Relations with AIRTUM
 - I-4 Relations with IARC
 - I-4 Relations with ENCR and IACR
 - I-4 Organization**
 - I-4 Group structure
 - I-4 Training
 - I-4 Defining projects and methods
 - I-5 Registries and mortality data
 - I-5 Accreditation**
 - I-5 Rules
 - I-6 Accreditation procedures and methods
 - I-6 Periodic audits of accredited registries
 - I-7 References**
-

CHAPTER 1

Registry fundamentals

Introduction

The level and quality of the contribution that cancer registries can offer to cancer research is greatly influenced by the choice of data sources and the application of rules of registration, which should be as uniform as possible between the different registries. The essential goals of registries are to achieve the greatest possible precision and accuracy, along with comparability and integration of the data produced; these goals can only be reached through patient research and comparison of the methods employed. The recent evolution in technology (improvement and evolution of diagnostic procedures), management (from computerization of the data sources to automatic registration), and care (reduced hospitalization and favoring of outpatient procedures, new models of care) have made these goals even more crucial.¹

The choice of building a single national database that guarantees adequate control of the quality of the data through analysis of the traditional quality indicators (DCO cases, histocytological confirmation, incidence/mortality ratio) invests AIRTUM with responsibility not only for the quality but also for the completeness and homogeneity of registration.²

General registration procedures

The purpose of this handbook is to provide methodological tools to help achieve adequate standardization in data acquisition, recording, and coding; software and other IT material is also provided, to enable registrars to simplify and optimize the data entry process and to carry out data quality control during registration. One of AIRTUM's tasks consists in making available to the registries the information and training needed to keep activities and procedures adequately up to date.

The standardization of the technical procedures of registration, albeit conditioned by the rapid evolution of problems, aims therefore at providing documented and shared rules for classification systems, with the following objectives:

- ◆ direct or indirect verification of the completeness of case collection;
- ◆ collection for each case of a predefined set of data of the highest quality possible;
- ◆ guarantee of a standardization of methods (and consequently of results) in space (between the various registries) and over time.

Conventions concerning the main aspects of registration are established internationally by the International Agency for Research on Cancer (IARC) and the International Association of Cancer Registries (IACR).³⁻⁵

A further resource is provided by the regular publication of worldwide incidence data (the quinquennial volumes of *Cancer Incidence in Five Continents*), which favors standardization of common rules, described in the preparatory material provided to the registries and in the chapters introducing the final publication.⁶ For the past few years, the process of standardization and publication of data has also had a European reference, the European Network of Cancer Registries (ENCR).⁷

These “institutional” criteria of standardization have been analyzed and in part modified by the registry networks of individual countries, as is the case for SEER and UKACR; it must be stressed that a detailed description of the criteria used, even if they differ from those recommended by IACR and ENCR, is essential in order to understand data and correctly evaluate differences. This means that whereas differences may exist between the networks and the IACR/ENCR, differences are excluded *a priori* within each network.

In Italy, too, each registry must align itself with the technical rules described in this handbook; each registry may have further internal rules - which always need to be explicitly stated - due to particular data flow issues or special local needs. For instance, internal registry rules may call for recording of data fields or information not included in the data set of the national database but considered necessary for activities supporting epidemiological and clinical research activities, in which registries are nowadays more and more involved (screening evaluation, follow-up, hereditary assessment in cancer genetics, etc.). From this point of view the registries, besides providing population data and facilitating health planning, are increasingly taking on an active role in clinical management and research.^{8,9}

Registration is based on **active casefinding**. Active casefinding is a process in which all available sources are used to find new cases and classify them as best possible; registries may also passively collect data from local information systems and then integrate them with active casefinding.

Population-based registries and special registries

Population-based registries cover the entire population of a specific area and all types of cancer; the rules of this handbook are particularly concerned with this kind of registry.

Special registries consider a subset of the population (childhood registries) or a specific disease (organ- or disease- specific registries). In order to be accredited, they must follow the rules contained in this handbook with respect to organization. Coding, data collection, and evaluation of cases, as well as control systems, may follow internal rules based on the recommendation of either international associations (e.g., childhood registries) or national coordination agencies (e.g., mesotheliomas).

Acquisition of data from multiple sources obviously produces an increase in the quality of the information available, but on the other hand it increases the number of multiple reports for each case, which must then be analyzed and discarded. This makes automatic record linkage procedures of data referring to the same individual extremely useful for reducing the workload in this phase.

There are therefore two possible types of operation:

- ◆ manual or semiautomatic registry: after an initial phase of active or passive registration of cases from one or more data sources, multiple reports are then checked using record linkage, with or without the help of computerized procedures;
- ◆ automated registry: data are processed automatically from the initial acquisition phase, with automatic record linkage flagging dubious or discrepant cases on which manual checks are then carried out; usually these registries cover regions or large city areas, where manual procedures would entail an impossible workload.

For a correct reconstruction of prevalent cases, it is essential for newly activated registries to have access to data sources referring to at least the two years prior to when registration started.

Formalization

Founding

In planning the information flow of a registry, data acquisition and management strategies must undergo a feasibility study, analyzing the following issues in depth:

- ◆ the reasons behind the creation of a new cancer registry, its context and territorial relevance;
- ◆ the population involved in terms of size and social and demographic characteristics;
- ◆ the expected information feedback and its contribution for the health system.

The study must also necessarily include an analysis of available resources with respect to:

- ◆ actual availability of the data sources defined as mandatory;
- ◆ working structure for the gathering, recording, coding, and analysis of data;
- ◆ adequacy of IT tools for case registration;
- ◆ health archives searchable on paper and/or with computer access;
- ◆ relationship with municipal registers and whether they may be accessed via computer;
- ◆ quantification of the actual cost (in terms of work, financial costs, and time needed for data acquisition) for every method of registration and every source of information;
- ◆ quality of data and level of data completeness that can be guaranteed;
- ◆ expected timeline for the activation of accreditation procedures, where planned, and for the database to be made available;
- ◆ any data fields recorded in addition to those of the AIRTUM Database, and their use;
- ◆ quality control systems employed and assessment schedules.

These criteria, which are obviously necessary to start a registry and gain accreditation, should also be used and verified at regular intervals to keep organization as cost-effective as possible, taking into account the new information opportunities that health organization and automated data management can offer.

The founding of a new registry must be approved by the local administration that the registry belongs to; the founding of a registry should also be acknowledged by regional administrations, for legal recognition purposes, as well as to enable a new information flow to be set up. This procedure is also designed to ensure stability over time of the workgroup. Before starting operations, every registry must draw a **protocol of operations** adapted to the specific local needs, which, along with this handbook, will provide a technical and organizational reference for workers or future workers of the registry; the registry must ensure that these tools are provided to all of the staff, who must receive adequate information and training.

Relations with AIRTUM

AIRTUM must be informed about the drafting of a feasibility study and the founding of a registry. Operators can become members of AIRTUM at any time and request aid in the form of useful skills and competences in the first stages of validation.

The new registry must then follow accreditation procedures with AIRTUM after a suitable period of registration (a period of at least three consecutive years is suggested). Accreditation is based on an overall evaluation of case records and quality indicators, as



well as on the registry's organization and procedures. AIRTUM will grant accreditation following the criteria listed in its statute.

Relations with the IARC

The IARC (International Agency for Research on Cancer of the World Health organization, based in Lyons) collects approximately every five years the data produced by cancer registries, and after assessing their quality publishes them in *Cancer Incidence in Five Continents*.

Relations with ENCR and IACR

Relations with ENCR and IACR are managed autonomously by each registry, with no involvement by AIRTUM.

It must be noted that registries that have already published their case sets in *Cancer Incidence in Five Continents* have automatic, free membership in the ENCR, whereas no membership is granted simply upon request; association with the IACR, on the other hand, is upon request and is regulated by a specific statute.

Organization

Registry organization is closely correlated with the way activities are structured. In an automated registry, for instance, management of very large, complex data sets and systematic use of electronic record linkage usually requires registry staff to have computer skills, while specialized registries obviously need close integration with medical specialists (pathologists, oncologists, etc.), based on the type of data collected and the registry's set goals. Furthermore, the functional location of the registry (university center, hospital or hospital foundation, Local Health Unit prevention department) is important, as it can lead to further and complementary research activities, favoring integration of specialists from various fields.

Group structure

An organization chart must be drawn up, with a central coordinator and supervisor (registry director) who represents the registry within AIRTUM. The registration group can be supported by external consultants (clinicians, pathologists, computer engineers, statisticians, epidemiologists), based on the registry's special needs. The director assigns functions, based on the level of experience and IT skills, defines the operations protocol and represents the association in relations with institutions, with the agency the registry belongs to, and with all other organizations the registry needs to interact with (regions, municipalities, Local Health Units, hospitals).

The director of the registry is also responsible for staff training, both initial and ongoing, as well as of ensuring data confidentiality.

Training

All employees of the registries must follow a training program and ongoing training sessions specifically targeted to their role, with regards to:

- ◆ computer use;
- ◆ use of databases;
- ◆ procedures to be followed to manage data;
- ◆ registration and coding rules;
- ◆ advances in the field of cancer diagnosis and therapy;
- ◆ descriptive epidemiology;
- ◆ data analysis and quality control software.

As regards, in particular, registration and coding rules, and analysis and quality control software, AIRTUM organizes a yearly meeting for registrars, held each fall. Registrars, in any case, are advised to take part in ENCR and IACR courses whenever possible.

Defining projects and methods

As previously mentioned, before starting operations, registries must draw up a protocol of operations.

Several aspects need to be addressed in the protocol.

- ◆ **Timeline of data collection**, linkage with other health sources, and linkage with mortality data. In this context, mortality data need to be systematically compared only after data collection and linkage have been completed, and before active casefinding of clinical records from hospitals outside the catchment area of the registry, in order to correctly manage and compute DCI cases and carry out a correct trace-back.
- ◆ **Process timelines**. The duration of each phase of work must be established before data collection begins. Should it be necessary for economic or organizational reasons (e.g., due to the need to access a great number of health centers), certain stages, such as follow-up at municipal registers or external clinical record acquisition, may be carried out over more than one year.
 - ◆ **Data set layout**, mandatory and optional data fields.
 - ◆ **How to manage problematic cases** in terms of nosology, coding, and/or assignment of date of incidence.
 - ◆ **How to manage cases *not sure eligibility*** (NSE).
 - ◆ **Quality control and revision** of cases, as well as management of errors discovered.

- ◆ **Electronic and hardcopy archive management**, including back-up procedures and facility security.
- ◆ **Any additional internal rules** that must not be in conflict with ENCR or IACR rules, nor with any rules described in this handbook.
- ◆ **Publication standards**.

Registries and mortality data

Local health units (ASL) collect a copy of ISTAT death records, which contain data that can be managed and used to study the registry catchment area and larger areas (cities, provinces, regions), as well. Availability of these data is essential for every registry, because, coupled with incidence data, they enable registries to:

- ◆ complete the follow-up on the life status of registered patients;
- ◆ verify place and cause of death, especially for NSE cases;
- ◆ find cases previously overlooked by incidence recording and tracing them back (if necessary using any indications on the disease's duration), estimating these losses using DCI (death certificate initiated) and DCO (death certificate only);
- ◆ correct any mistakes in vital statistics;
- ◆ compute survival measures and estimates;
- ◆ compute prevalence measures and estimates.

In any case, the quality of data must be carefully monitored (for instance, through linkage between registry and mortality databases), to ensure that errors in completeness or other coding criteria do not lead to loss of information on the presence of a tumor among the causes of death. Changes to improve and correct the above processes should be implemented if needed.

Due to the considerable differences in disease definition, maximum sensitivity criteria need to be used, including even generic or mis-classified terms, using the list found in the paragraph devoted to NSE cases. For, in part due to age and co-morbidity, some sites tend to be classified poorly with no indication whether the tumor is benign or malignant due to the use of generic expressions, such as “neoplasia” or “tumor” (e.g., prostate and bladder).

Thus, linkage between registries and death certificates also provides a useful tool for estimating both underreporting of mortality (in the event that the cancer cause was not considered among the causes of death by the ISTAT certificate) and its overestimate (mortality **with** cancer and mortality **from** cancer, which might require specific analysis).

Taking up coding of mortality according to ICD-10 might lead to further distortions, particularly in cases with multiple primaries (e.g., if a tumor is associated with a primary of the lung, the latter is considered metastatic;

multiple primaries in independent sites are coded to C97 “multiple primaries”), a close comparison between registries and mortality data is therefore essential.

Cancer registry with internal mortality registry

Whereas in certain regions mortality registration must needs take place within each registry, in other regions registries may decide to register mortality for the area they cover, especially in particular situations (e.g., at registry start up, if the data are available, etc.).

Staff must of course be trained to follow ISTAT rules and gather all specific neoplasms independently of their being the cause of death. It must be stressed that national and international guidelines do not (except for extraordinary cases) consider accreditation of mortality data other than the official ones. In any case, registries must make sure beforehand that institutional sources of mortality data (Local Health Units, regions, ISTAT) are willing to provide the needed data at national and international incidence data publication times.

Cancer registry without mortality registry

These registries must ensure they have access to the individual database of the mortality registry of the area they cover, both in order to build statistics and, in particular, to carry out the above-mentioned trace-back of DCI and NSE cases. To do this, they need to have access to complete vital statistics or to unique and shared indicators (to be verified) to carry out individual record linkage (tax number, etc.).

Registries must therefore:

- ◆ acquire from the Local Health Unit (ASL) a complete listing of vital statistics and causes of death, if available, or at least the record of deceased people recorded by townships, on paper if no electronic file is available.
- ◆ systematically analyze ISTAT death records and extract a copy of them with reference to the presence of cancer or subjects with NSE diagnosis.

Accreditation Rules

Accreditation procedures guarantee the scientific level of the data published by a cancer registry and of any material submitted to larger studies. Data accreditation therefore traditionally precedes all official publications; the quinquennial publication by the IARC of *Cancer Incidence in Five Continents* is for every registry a test of the quality of its work and proof of belonging to the community of internationally accredited cancer registries.

It must be stressed that accreditation procedures concern the data produced (first and foremost, incidence data) and only indirectly concern the structure, organization, or resources of a registry.

Acceptable results can of course only be achieved with a solid work organization and structure, and these are therefore a prerequisite for accreditation, but they only represent the basis for efficient following of registration rules.

Data-oriented accreditation results in a dynamic, ongoing process that concerns all cancer registries and not only the newer ones; it also provides a chance for registries to debate and share issues in the national and international scientific community; this has allowed cancer registries to qualify themselves increasingly over the past few years as a research network and as providers of institutional services.

Accreditation procedures and methods

In Italy, AIRTUM has set up a data accreditation procedure for registries that follows international rules, in order to share data in a single national archive (AIRTUM Database). Registries whose data are not yet included in this database are invited to request accreditation from AIRTUM, which sets up an evaluation committee, according to the procedures listed in its internal rules.

To be evaluated, registries must present their incidence data referring to a significant period of observation (at least three consecutive years). Registries being evaluated need to provide the following material:

- ◆ a computerized copy of the incidence data being evaluated, in a format that follows the standards of the AIRTUM Database for analytical evaluations;
- ◆ mortality file (individual or aggregated data) and population file, as detailed as possible (preferably with yearly age classes from 0 to 90+ by every single year of registration);
- ◆ questionnaire provided by the AIRTUM offices based on the most up to date IARC/IACR and ENCR guidelines;
- ◆ printouts/files of IARCTools and DEPedit outputs. (see Chapter 5);
- ◆ printout/file with number of incident cases, deceased subjects in the time interval, and specific age rates (including subdivision by age classes 0-34; 35-64; 65-74; 75+);
- ◆ printout/file with: total number of cases, crude rates, standardized rates (Italy 1981, world, truncated), cumulative risks at 74 years (following the format of *Cancro in Italia*);
- ◆ printout/file with the percentage of cases with histological diagnosis, cytological diagnosis, clinical diagnosis, and DCO cases, including the subdivision by age classes 0-34; 35-64; 65-74; 75+;

- ◆ printout/file with mortality/incidence ratio;
- ◆ printout/file with distribution of subsites;
- ◆ printout/file with distribution of morphologies by site;
- ◆ printout/file with total number of cases by site, sex, and year of registration;
- ◆ printout/file with total number of DCO cases by site, sex, and year of registration;
- ◆ printout/file with total number of cases with microscopic verification (histology and cytology) by site, sex, and year of registration;
- ◆ printout/file with number of cases with unknown morphology (ICD-O-3 M-8000) and histological basis of diagnosis by site and sex.

These data must be integrated with information mainly regarding:

- ◆ the various independent information sources that the registry regularly refers to (e.g., laboratories, radiology units, health tax exemption archives, etc.);
- ◆ criteria used in the following situations:
 - ◆ in which cases a histological basis (IARC basis of diagnosis code “7”) is taken instead of a cytological one (IARC basis of diagnosis code “5”) from needle biopsy, bone marrow aspiration biopsy, thoracentesis, paracentesis, and so forth;
 - ◆ what the registry's criteria are with respect to assigning a behavior code (/1, /2, /3) when there is a histological report of a urothelial epithelium tumor in which the condition of invasion or the pT is not explicitly reported;
 - ◆ when the registry assigns a specific morphology code based on a diagnostic method that is not microscopic diagnosis (e.g., CNS tumors diagnosed by imaging);
- ◆ procedures for treatment of DCI cases;
- ◆ procedures used for follow-up of cases (sources, criteria, definition of reference dates, active finding and reporting of emigrants);
- ◆ the type of registry establishment (regional, private, other) and staff number.

Once all conditions have been verified, the accreditation process ends with the admission of the registry's data to the national database.

Periodic audits of accredited registries

When registries subsequently submit data to the AIRTUM Database, they must inform AIRTUM of any variations in their operational procedures, without having to repeat the accreditation process; data, instead, are verified with the same quality criteria as in the initial accreditation, and then entered into the national archive.

References

1. Parkin DM. The Evolution of the Population-Based Cancer Registries. *Nat Rev Cancer* 2006; 6(8): 603-12.
2. Black R. Ensuring High-Quality Epidemiological Data on Cancer. *Eur J Cancer Prev* 2005; 14(4): 305-06.
3. Jensen OM, Parkin DM, Mac Lennan R et al. *Cancer Registration, Principles and Methods*. IARC Scientific publications no. 95, Lyon 1991.
4. Esteban D, Whelan SL, Laudico A, Parkin DM et al. *Manual for Cancer Registry Personnel*. IARC Technical report no. 10, Lyon 1995.
5. Parkin DM, Chen VW, Ferlay J et al. *Comparability and Quality Control in Cancer Registration*. IARC Technical report no. 19, Lyon 1995.
6. Parkin DM, Whelan SL, Ferlay J et al. *Cancer Incidence in Five Continents*. Vol. VIII IARC Scientific publications no. 155, Lyon 2002.
7. European network of cancer registries. *EUROCIM User Manual*. 2nd edition. Lambda plus software products, 1995.
8. Glaser SL, Clarke CA, Gomez SL et al. Cancer Surveillance Research: A Vital Subdiscipline of Cancer Epidemiology. *Cancer Causes Control* 2005; 16(9): 1009-19.
9. Goodman MT, Hernandez BY, Hewitt S et al. Tissues from Population-Based Cancer Registries, A Novel Approach to Increasing Research Potential. *Hum Pathol* 2005; 36(7): 812-20.