6
Quality Control

The primary goal of a population-based cancer registry is to determine the incidence of cancer within its geographical population. It is therefore of the utmost importance that the registry data be reliable and of good quality. Coverage of the population should be as complete as possible and information gathered, especially on essential items, should be complete, consistent and accurate.

Quality control is the mechanism by which the quality of data can be assessed. This may be either a formal on-going programme incorporated in the standard operating procedures of the registry or it may be an ad hoc survey to assess completeness and consistency of casefinding, abstracting and coding, as well as the accuracy of reporting. Quality control programmes can show the level of errors, and can also include feedback mechanisms to improve accuracy and consistency. Less formal quality control involves the critical scrutiny of the data as they are used. To some, this is the best form of quality control.

This chapter presents various methods for monitoring the quality of data in the cancer registry. More detailed information can be found in the IARC/IACR Technical Report *Comparability and Quality Control in Cancer Registration* (Parkin et al., 1994).

1. *Completeness of cover*

The population-based registry aims to record all cancer cases occurring within its defined geographical area. It is therefore essential that all the data sources for the registry be covered completely. That is, casefinding and abstracting should include all hospitals within the catchment area of the registry. All data sources within these hospitals should likewise be covered in order to avoid under-reporting.

On the other hand, each abstract must be carefully checked soon after it arrives in the registry to determine if the case is eligible for registration:

- Is the diagnosis included in the registry's reportable list?
- If the registered cases are limited to residents of the catchment area, is the patient's residence inside the boundaries of this area?
- Is the incidence date on or after the reference date (the date from which cases are recorded) of the registry?
- Are the essential items of information in the registry abstract complete?

The registry must also endeavour to avoid duplication of patients or multiple registrations as these will artificially inflate the incidence rates. The items of identifying information must be sufficient to ensure that a patient who has been registered previously can be recognized as the same person, should he be reported again to the registry. Multiple primaries occurring in the same patient, either at the same time or at different times, should also be identified. A case with more than one primary can be identified by, for example, a primary registration with a sequence number of 2 or more (see section 4.2.7 (10), sequence number).

How to assess completeness of cover

The use of death certificates as a source of information provides a very useful method of evaluating completeness.

(a) The proportion of cases which come to the attention of the registry for the first time through a death certificate (death certificate notification – DCN) should be monitored. This type of case may be easily identified if 'Method of detection' (see section 3.3.4) or a similar variable is recorded by the registry. The registry should allow a defined time period (which would vary according to how frequently the registry has access to the death certificates) to elapse before matching death certificates against

the file of registered cases. If many cases are being found via death certificates, it is certain that registration is incomplete. For every cancer patient who dies before being discovered by the registry, there is probably one who does not die, and is thus never identified. If a high proportion of cases is identified by death certificates, the registry should determine the reasons for this and implement measures to improve completeness. The cases should be carefully followed back to the hospitals where they died or they may be traceable through the physician who signed the death certificate. In this way it should be possible to establish where in the hospital (or other medical resources in the community) the registry is failing to find cases.

A distinction must be made between a DCN and a Death Certificate Only (DCO) case. Only the DCN cases which cannot be traced or followed back to hospital or the certifying physician are registered as 'Death Certificate Only' (DCO) cases (see section 3.3.3). A registration which is made on the basis of a death certificate alone, with no other documentation, is likely to be less accurate than a diagnosis supported by histological, or at least clinical, confirmation.

(b) Monitor the incidence of each site annually and compare the latest year of incidence with previous years. Any marked change should be investigated. Under-reporting may be site-specific, e.g., in hospitals where research on a particular cancer is being carried out the medical records on patients involved in the study may be taken out by the researcher so that they cannot be located by registry staff.

(c) Whenever possible, monitor the difference in incidence rates for the subdivisions of the registry's geographical area. For example, note the difference in incidence rates by municipality within a province. The change in incidence rates in a particular area is usually slow. Any marked change should be investigated to identify areas of under-reporting which may need action.

(d) Another method to assess completeness of cover would be to sample patient attendance at a specialized clinic and later check if they are included in the registry. Sometimes case series of cancers have been collected by doctors for research purposes. The registry should use these independent lists to check what proportion of the cases had been found.

Assessment of completeness of cover should be carried out constantly.

2. Completeness and accuracy of details

All incoming reports or registry abstracts should be checked rapidly upon arrival to ensure that at least all the essential items of information are complete. In this way, errors can be corrected while the hospital records are still available. Errors in detail may arise while abstracting or during transcription, coding or key punching.

How to assess accuracy

(a) The best method of assessing completeness and accuracy of detail in a record is by 'blind' re-abstraction and recoding without reference to the original registration. The initial and the reprocessed registrations are later compared to determine the error rates for each item. Since accuracy is a matter of degree, a 'scale of error' for each item may have to be established. Ideally this quality control programme should be built into the registry system, with a percentage of registrations re-abstracted and re-coded.

Duplicate coding of essential items, such as diagnosis (primary site and histology) and most valid basis of diagnosis, ensures consistency between coders.

It is preferable that all the members of the registry staff share in the responsibility for quality control, since this increases their awareness of the need for high quality data. However, if the registry is under-staffed and in need of funds, *ad hoc* exercises would be more practical.

(b) In computerized registries data quality can be checked using automated routines:

(i) Validation checks: these checks are carried out on each data item to ensure that there are no invalid codes fed into the data base. All computerized registries should have computer files containing the valid codes for each item (coding control files). Every incoming code is checked against the control file and if this is invalid, the code will be rejected.

(ii) Consistency checks: these checks compare the concordance of specified data items against other recorded items, for example.

- check for cancer of the cervix uteri in males or prostatic cancer in females.
- the sequence of dates should be checked so that the sequence of date of birth, date of diagnosis and date of death is preserved (a patient's incidence date cannot be earlier than his date of birth nor later than his date of death).
- check that certain site-specific morphology terms have the correct topography codes assigned (e.g., nephroblastoma which arises from the kidney should have a topography code of C64.9 (189.0), and hepatoblastoma which arises from the liver should have a topography code of C22.0 (155.0).

In certain instances, attention may be drawn to possible errors and warnings issued. Cases of rectal carcinoma or chronic myelogenous leukaemia in children can be signalled, not because they are necessarily wrong but because these cases do not usually occur in children. A review of the case is warranted to rule out any error.

Computer checks may either be done at the time the data are being entered (on-line) or as a part of a batch operation (off-line). A 'scale of errors' may be set up in the system such that major errors result in complete rejection of a registration, while less serious ones may be recorded and added to the database. The latter cases should be flagged to indicate that they contain an error. The most serious errors should be corrected first.

The IARC-CHECK Program (available with the publication *Comparability and Quality Control in Cancer Registration* (Parkin *et al.*, 1994)) checks data for validity and for consistency. The data items checked by the program are:

- registration number
- date of incidence
- age (or date of birth)
- sex
- site
- histology
- basis of diagnosis

3. *Pre-requisites for quality control*

(a) Rules and documentation

The registry should have a set of rules covering its different functions and activities, with a rigid definition of the data items to be collected and other associated terms. These rules and definitions should be written down and kept on file as a ready reference for the registry personnel. They may be kept in the form of a procedural manual which should be applied consistently over time as changes occur in the registry. If there are any changes in the rules or definitions, these should be documented for accurate usage and interpretation. In certain cases where subjective judgement is necessary, the senior members of staff should be consulted and the reasons for the decision should be documented as a guide to solving similar situations in the future.

(b) Good coding systems

In a good coding system, only one code is allocated for each appropriate term. If there is any change in the coding system, there should be documented rules as to the time period under which a given set of codes operates.

(c) Standards

The registry should have standards under which to operate. Maximum tolerable error rates should be set for major data items (for example 5% for the three-digit level of the ICD-O, or 0.5% for sex). If these rates are exceeded, corrective action should be taken to reduce the errors to acceptable standards.