Chapter 12

Data Quality and Indices of Reliability

This chapter describes procedures for assessment of the quality of the data and the completeness of coverage of cases in a given registry area.

Checks on Quality of Data

The registry data undergoes several quality checks, both, at the time of data entry and subsequently. These include: Range, Consistency, Unlikely and Family checks as per the International Association of Cancer Registries (IARC) norms. All the checks are built into the PBCRDM 2.1 and the online PBCR data entry application. The list of cases with possible errors is sent back to the respective registries for verification with the original medical records and the corrections received are updated in the registry database. Cancer Incidence in Five Continents (CI5 Vol-XI) published by IARC, has accepted and incorporated the data of 15 Indian registries out of the 22 registries that submitted the data (2008-2012).

Internal consistency

Innovation of different software application at NCDIR-NCRP supports cancer registration in a big way. PBCR softwares (desktop and web based) run the quality checks (consistency, range, unlikely, family), matching and duplicate check to make the data clean and valid. Additionally, a Phonetics software is used to weed duplicate names that sound similar but are spelt differently. Fluctuation in the number of cancer cases over the years from each source of registration is identified using the software for appropriate action.

Some of the specific checks that appear important in this context are:

- % Age Unknown <10%
- % Death Certificates Only <10%
- % Other & Unspecified Sites <10%
- % Microscopic verification (MV) >80% (99-100% is unacceptable).
- Mortality to Incidence Ratio (M:I)%
- Stability of incidence rates (the number of new cases) over time thereby disallowing any abrupt trend.

Age Unknown

Most of the PBCRs do not have any cases with age unknown. The highest proportion of cancer cases with age being unknown was from Delhi PBCR (0.6%).

In HBCRs, out of 58 hospitals, 36 hospitals collected the actual age of the patient. However, only in 154 cases, age was unknown from 22 hospitals.

Microscopic Verification (MV)

Higher the proportion of microscopically verified cases, more accurate is the confirmation as microscopic verification is the most valid basis of diagnosis of cancer. Still, a very high proportion (above 90-95%) of microscopic diagnosis suggests the likelihood that some cancers with a diagnosis based on imaging techniques and solely clinical diagnoses may be missed by the registry. The highest proportion of microscopically verified cases were from Hyderabad district (96.7%) and the lowest was reported from Patiala district (77.1%).

The MV percentage ranged between 90 – 100% in majority of the hospitals. Among microscopically confirmed cases, about 82% were found to be that of histology of primary followed by cytology of primary (6.9%) and bone marrow (5.2%). There were about 5.6% of cases diagnosed from metastatic site of cancer.

Death Certificate 'Only' (DCO) cases

There is a need to follow-back Death Certificate Notifications (DCN), a case notified by death from a death certificate. These DCNs are followed to the last hospital attended to obtain the details of diagnosis to register an incidence. To successfully achieve this, scrutiny of cancer deaths as per death registers/certificates which have not matched with any incidence record is undertaken. This way the exact primary site of tumour and the date of diagnosis is obtained. If no information is obtained it is registered as DCO.

The relative proportion of DCOs should ideally be less than 10% which was found in 26 of the 28 PBCRs. DCO <1% was observed in 9 PBCRs.

Other and Unspecified Site (O&U)

The sites of cancer that were categorised as "Other and Unspecified Sites (O&U)" were as per ICD-10: C26, C39, C48, C75, C76, C77, C78, C79, C80, C97 (WHO 1994). The relative proportion of cancers that fell into this group was more than 10% in the PBCRs at Patiala district, Thiruvananthapuram district, Nagpur and Cachar district. It was less than 5% in 13 PBCRs.

Among 58 hospitals, O&U percentage ranged between 0.1 – 6.8%. In three hospitals, the relative proportion of O&U was more than 10%.

Mortality-Incidence Ratio (MI Ratio)

The mortality-incidence or MI ratio is an indicator of the completeness and accuracy of cancer mortality data. The system of registration of death and certification of cause of death are of major concern. In order to overcome this deficit in cancer mortality data, some PBCRs have used the all-cause mortality data collected from municipal/corporation units of their registry area to match with the incident cases and arrive at the figure of cancer mortality. Barshi rural, Wardha district and Mumbai have M/I% ratios of 67.2%, 59.2% and 56.0, respectively.

Clinical Extent of the Disease before Treatment (CEDBT)

The clinical extent of the disease provides an idea of the degree of the spread of cancer when the patient presents himself/herself to the reporting hospital without receiving any cancer directed treatment earlier. 13 hospitals had the details of CEDBT, 33 hospitals could not ascertain the extent of the disease which was less than 10% of the cases. However, the relative proportion of unknown CEDBT varied from 12.0% to 64.1% in remaining 12 hospitals.

Unspecified sub-site

Anatomical sites of cancer are generally considered as one complete entity for overall expression of number of cases. Registry wise analyses were done for the sites of cancer provided in the section II of this report such as cancer breast, cervix, head & neck, lung and stomach. The percentage variation of unspecified sub-site for each site is given below.

Anatomical Sites of Cancer	Percentage of unspecified sub-site
Cancer Breast	32.4% - 100%
Cancer Cervix	51.0% - 100%
Head & Neck Cancers	10.0% - 77.8%
Cancer Lung	25.0% - 100%
Cancer Stomach	23.2% - 100%

Unspecified Histology

While cancers of different anatomical sites have certain distinctions due to their location, the histological type of cancer in the same site has its own identity in terms of aetiology, prognosis and treatment thereof. Hence, it is important to get information in at least cases where a microscopic diagnosis of cancer is available. The relative proportion of the cases that had unspecified histology is provided below.

Anatomical Sites of Cancer	Percentage of unspecified Histology
Cancer Breast	0.1% - 57.5%
Cancer Cervix	0.1% - 56.1%
Head & Neck Cancers	0.1% - 94.9%
Cancer Lung	0.2% - 36.6%
Cancer Stomach	0.9% - 87.5%

Method of Diagnosis

The proportion of microscopic verification in males varied from 74.3% in Patiala PBCR to 96.2% in Nagaland PBCR.

Among males, clinical diagnosis was the highest in Nagpur at 8.0% and X-ray and Imaging as a form of diagnosis was the highest in Kollam district (14.4%).

Among females, the microscopic proportion varied from 79.5% in Patiala district to 97.3% in Hyderabad district. Among females, clinical diagnosis was highest in Nagpur at 8.2% and X-ray and Imaging as a form of diagnosis was the highest in Cachar district with a relative proportion of 9.2%.

Detailed Microscopic Diagnosis

The proportion of primary histology in males varied from 44.3% in West Arunachal to 84.1% in Aurangabad PBCR.

Among males, cytology was the highest in West Arunachal at 45.8% and bone marrow was the highest in Delhi PBCR (7.2%).

Among females, the primary histology varied from 50.2% in Cachar district to 90.6% in Aurangabad PBCR. Among females, cytology was highest in Papumpare district at 47.3% and bone marrow was the highest in Ahmedabad urban with a relative proportion of 4.7%.