DATA QUALITY AND INDICES OF RELIABILITY

The objective of this chapter is to provide an assessment of the quality of the data and the completeness of coverage of cases in a given hospital.

Newer HBCRs

The data of the newer HBCRs that are being reported for the first time in this report are Guwahati and Chandigarh. Care has been taken to ensure that these registries have complied with quality of data in terms of actual data collation from various departments of the hospital, duplicate elimination and the characteristics of the data submitted (Parkin *et al*, 1994).

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 IARC norms. All the checks are built into the HBCRDM 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 data base. Tables 14.1 to 14.5 provide an insight into the quality of the data of 7 HBCRs after such corrections have been done on the data.

Age Unknown

The number and proportion of cancers with age being unknown in each of the 7 HBCRs is given in Table 14.1. Most of the HBCRs do not have any cases with age unknown. Nonetheless, all the HBCRs are unable to ascertain the date of birth in the vast majority of cases.

Unspecified or Unknown Duration of Stay

The number and proportion of cancers with unspecified/unknown duration of stay in each of the 7 HBCRs is given in Table 14.2.

Microscopic Verification

The proportion of microscopically verified cases (Table 14.3) is an internationally accepted indicator of data quality. Higher the proportion of microscopically verified cases the 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.

Other and Unspecified Site (O&U)

The sites of cancer that were categorised as "Other and Unspecified Sites (O&U)" as per ICD-10 were C26, C39, C48, C75, C76, C77, C78, C79, C80, C97 (WHO 1994). The relative proportion of cancers that fell into this group (Table 14.4) was less than 5% in all HBCRs.

There is a need for registry abstractors to diligently track these cases to the concerned physician/ pathologist and find the information on the exact primary site of tumour. Timeliness is extremely important and this should be done at initial abstraction itself which in turn should be as close as possible to the date of diagnosis.

Unspecified Sub-site

Anatomical sites of cancer are generally considered as one complete entity for overall expression of number of cases. However, bearing in mind embryological development and in terms of identifying risk factors, there is a need for sub-site classification of at least some important pertinent sites of cancer such as tongue, oesophagus, stomach and colon. Sub-site identification is also an indicator of the meticulousness of the registry staff and the extent of detail of data availability vis-à-vis clinical-pathology records. The registry-wise proportion of unspecified sub-site for these four sites of cancer is given in Tables 14.5 to 14.8. Suffice to state that sub-site categorisation is uniformly low across all HBCRs. Even those with small numbers are unable to obtain information on sub-site in a substantial proportion of cases. Like for "Other and Unspecified Sites" awareness by the abstractor on the need to collect such information where available and pursuing with the concerned clinician/pathologist where not available. Timeliness in both abstraction and pursuit is once again the key in getting such data.

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. Tables 14.9, 14.10 and 14.11 give the proportion of cancers of selected sites (stomach, lung and ovary) where histology was "Not Otherwise Specified".

Clinical Extent of Disease Before Treatment

This is an important item of information to know the spread of the disease. For the group of cases that are treated only at the Reporting Institute and not received any cancer directed treatment (excludes patients previously treated) the percentage of cases not having information on CLE (Table No. 14.12) is less in all HBCRs except in Mumbai (43.6%).

Table 14.1: Age Unknown - Both Sexes
Number (#) and Relative Proportion (%)

Registry	Total	Age Unknown	
		#	%
Mumbai	41108	91	0.2
Bangalore	24398	9	0.0
Chennai	33230	-	-
Thir'puram	38028	106	0.3
Dibrugarh	5171	-	-
Guwahati	11482	3	0.0
Chandigarh	4735	7	0.1

Table 14.2: Unspecified (Unsp.)/Unknown Duration of Stay (DOS) - Both Sexes
Number (#) and Relative Proportion (%)

Registry	Total	DOS Unsp./Unknown	
		#	%
Mumbai	41108	41108	100.0
Bangalore	24398	15117	62.0
Chennai	33230	9209	27.7
Thir'puram	38028	38028	100.0
Dibrugarh	5171	2891	55.9
Guwahati	11482	2	0.0
Chandigarh	4735	1	0.0

Table 14.3: Microscopic Verification (MV) - Both Sexes
Number (#) and Relative Proportion (%)

Registry	Total	MV	
		#	%
Mumbai	41108	38433	93.5
Bangalore	24398	23078	94.6
Chennai	33230	27395	82.4
Thi'puram	38028	36058	94.8
Dibrugarh	5171	4744	91.7
Guwahati	11482	10513	91.6
Chandigarh	4735	4639	98.0

Table 14.4: Other and Unspecified Site (O&U) - Both Sexes
Number (#) and Relative Proportion (%)

Registry	Total	O&U	
		#	%
Mumbai	41108	2012	4.9
Bangalore	24398	1087	4.5
Chennai	33230	1517	4.6
Thi'puram	38028	1384	3.6
Dibrugarh	5171	254	4.9
Guwahati	11482	567	4.9
Chandigarh	4735	213	4.5

Table 14.5: Unspecified (Unsp.) Sub-Site - Tongue (ICD10: C01-C02) - Both Sexes
Number (#) and Relative Proportion (%)

Registry	Total	Unsp. Sub-Site	
		#	%
Mumbai	2055	463	22.5
Bangalore	917	237	25.8
Chennai	1366	34	2.5
Thi'puram	1762	801	45.5
Dibrugarh	215	18	8.4
Guwahati	607	20	3.3
Chandigarh	230	76	33.0

Table 14.6: Unspecified (Unsp.) Sub-Site - Oesophagus (ICD10: C15) - Both Sexes
Number (#) and Relative Proportion (%)

Registry	Total	Unsp. Sub-Site	
		#	%
Mumbai	1555	609	39.2
Bangalore	1553	371	23.9
Chennai	1451	116	8.0
Thi'puram	1091	536	49.1
Dibrugarh	617	80	13.0
Guwahati	1507	527	35.0
Chandigarh	270	83	30.7

Table 14.7: Unspecified (Unsp.) Sub-Site - Stomach (ICD10: C16) - Both Sexes
Number (#) and Relative Proportion (%)

Registry	Total	Unsp. Sub-Site	
		#	**%**
Mumbai	1116	912	81.7
Bangalore	900	509	56.6
Chennai	2092	821	39.2
Thi'puram	1201	901	75.0
Dibrugarh	330	194	58.8
Guwahati	640	262	40.9
Chandigarh	79	59	74.7

Table 14.8: Unspecified (Unsp.) Sub-Site - Colon (ICD10: C18) - Both Sexes
Number (#) and Relative Proportion (%)

Registry	Total	Unsp. Sub-Site	
		#	**%**
Mumbai	631	268	42.5
Bangalore	250	118	47.2
Chennai	468	104	22.2
Thi'puram	628	236	37.6
Dibrugarh	109	23	21.1
Guwahati	147	55	37.4
Chandigarh	90	28	31.1

Table 14.9: Unspecified (Unsp.) Histology - Stomach (ICD10: C16) - Both Sexes
Number (#) and Relative Proportion (%)

Registry	Total	Unsp. Histology	
		#	**%**
Mumbai	1201	115	9.6
Bangalore	923	193	20.9
Chennai	2122	656	30.9
Thi'puram	1233	183	14.8
Dibrugarh	330	81	24.5
Guwahati	642	136	21.2
Chandigarh	81	6	7.4

Table 14.10: Unspecified (Unsp.) Histology - Lung (ICD10: C33-C34) - Both Sexes

Number (#) and Relative Proportion (%)

Registry	Total	Unsp. Histology	
		#	%
Mumbai	2343	449	19.2
Bangalore	1014	293	28.9
Chennai	1821	808	44.4
Thi'puram	3374	1257	37.3
Dibrugarh	138	41	29.7
Guwahati	645	165	25.6
Chandigarh	303	7	2.3

Table 14.11: Unspecified (Unsp.) Histology - Ovary (ICD10: C56)

Number (#) and Relative Proportion (%)

Registry	Total	Unsp. Histology	
		#	%
Mumbai	939	253	26.9
Bangalore	779	145	18.6
Chennai	922	303	32.9
Thi'puram	954	132	13.8
Dibrugarh	200	72	36.0
Guwahati	229	37	16.2
Chandigarh	168	19	11.3

Table 14.12: Unspecified/Unknown Clinical Extent of Disease (CLE) (Excludes Patients Previously Treated)- Both Sexes

Number (#) and Relative Proportion (%)

Registry	Total	CLE Unknown	
		#	%
Mumbai	33205	14470	43.6
Bangalore	20538	564	2.7
Chennai	28414	687	2.4
Thi'puram	27732	1154	4.2
Dibrugarh	4955	68	1.4
Guwahati	9580	4	0.0
Chandigarh	3967	22	0.6