

Chapter 11

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 HBCR that is being presented for the first time in this report. Care has been taken to ensure that this registry 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 database. Tables 11.1 to 11.5 provide an insight into the quality of the data of eight 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 eight HBCRs is given in Table 11.1. 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 eight HBCRs is given in Table 11.2.

Microscopic Verification

The proportion of microscopically verified cases (Table 11.3) 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 11.4) was less than 5% in all HBCRs except TMH - Mumbai and BBCI - Guwahati.

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 11.5 to 11.8. 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 11.9, 11.10 and 11.11 give the proportion of cancers of selected sites (stomach, lung and ovary) where histology was “Not Otherwise Specified”.

Clinical Extent of Disease Before Treatment (CEDBT)

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 CEDBT (Table No. 11.12) is less in all HBCRs except in TMH - Mumbai (51.9%) and KMIO - Bangalore (23.2%).

Table 11.1: Age Unknown - Both Sexes*Number (#) and Relative Proportion (%)*

Registry	Total	Age Unknown	
		#	%
TMH	25541	-	-
KMIO	8687	-	-
CI (WIA)	8554	-	-
RCC - TVM	23206	50	0.2
AMC	3587	-	-
BBCI	22562	3	0.0
PGIMER	17182	2	0.0
BRAIRCH	8039	3	0.0

Table 11.2: Unspecified (Unsp.)/Unknown Duration of Stay (DOS) - Both Sexes*Number (#) and Relative Proportion (%)*

Registry	Total	DOS Unsp./Unknown	
		#	%
TMH	25541	25541	100.0
KMIO	8687	1757	20.2
CI (WIA)	8554	62	0.7
RCC - TVM	23206	23206	100.0
AMC	3587	-	-
BBCI	22562	3	0.0
PGIMER	17182	2	0.0
BRAIRCH	8039	-	-

Table 11.3: Microscopic Verification (MV) - Both Sexes*Number (#) and Relative Proportion (%)*

Registry	Total	MV	
		#	%
TMH	25541	24062	94.2
KMIO	8687	8228	94.7
CI (WIA)	8554	7170	83.8
RCC - TVM	23206	22241	95.8
AMC	3587	3170	88.4
BBCI	22562	19168	85.0
PGIMER	17182	16928	98.5
BRAIRCH	8039	7914	98.4

Table 11.4: Other and Unspecified Site (O&U) - Both Sexes*Number (#) and Relative Proportion (%)*

Registry	Total	O&U	
		#	%
TMH	25541	1443	5.6
KMIO	8687	326	3.8
CI (WIA)	8554	331	3.9
RCC - TVM	23206	485	2.1
AMC	3587	109	3.0
BBCI	22562	1298	5.8
PGIMER	17182	499	2.9
BRAIRCH	8039	171	2.1

Table 11.5: Unspecified (Unsp.) Sub-Site - Tongue (ICD10: C01-C02) - Both Sexes*Number (#) and Relative Proportion (%)*

Registry	Total	Unsp. Sub-Site	
		#	%
TMH	1309	316	24.1
KMIO	347	192	55.3
CI (WIA)	406	3	0.7
RCC - TVM	971	253	26.1
AMC	124	18	14.5
BBCI	679	9	1.3
PGIMER	721	289	40.1
BRAIRCH	361	98	27.1

Table 11.6: Unspecified (Unsp.) Sub-Site - Oesophagus (ICD10: C15) - Both Sexes*Number (#) and Relative Proportion (%)*

Registry	Total	Unsp. Sub-Site	
		#	%
TMH	813	334	41.1
KMIO	523	267	51.1
CI (WIA)	324	27	8.3
RCC - TVM	601	164	27.3
AMC	403	45	11.2
BBCI	2962	1298	43.8
PGIMER	1127	584	51.8
BRAIRCH	223	186	83.4

**Table 11.7: Unspecified (Unsp.) Sub-Site - Stomach (ICD10: C16)
- Both Sexes**

Number (#) and Relative Proportion (%)

Registry	Total	Unsp. Sub-Site	
		#	%
TMH	727	562	77.3
KMIO	412	338	82.0
CI (WIA)	571	192	33.6
RCC - TVM	880	618	70.2
AMC	222	77	34.7
BBCI	1184	509	43.0
PGIMER	293	237	80.9
BRAIRCH	160	106	66.3

**Table 11.8: Unspecified (Unsp.) Sub-Site - Colon (ICD10: C18)
- Both Sexes**

Number (#) and Relative Proportion (%)

Registry	Total	Unsp. Sub-Site	
		#	%
TMH	502	211	42.0
KMIO	125	77	61.6
CI (WIA)	129	9	7.0
RCC - TVM	404	104	25.7
AMC	96	16	16.7
BBCI	314	118	37.6
PGIMER	280	135	48.2
BRAIRCH	125	79	63.2

**Table 11.9: Unspecified (Unsp.) Histology - Stomach (ICD10: C16)
- Both Sexes**

Number (#) and Relative Proportion (%)

Registry	Total	Unsp. Histology	
		#	%
TMH	730	46	6.3
KMIO	409	48	11.7
CI (WIA)	469	59	12.6
RCC - TVM	872	62	7.1
AMC	203	6	3.0
BBCI	965	43	4.5
PGIMER	299	17	5.7
BRAIRCH	163	27	16.6

**Table 11.10: Unspecified (Unsp.) Histology - Lung (ICD10: C33-C34)
- Both Sexes**

Number (#) and Relative Proportion (%)

Registry	Total	Unsp. Histology	
		#	%
TMH	1424	330	23.2
KMIO	333	53	15.9
CI (WIA)	318	51	16.0
RCC - TVM	1997	511	25.6
AMC	69	15	21.7
BBCI	968	68	7.0
PGIMER	1401	12	0.9
BRAIRCH	475	1	0.2

Table 11.11: Unspecified (Unsp.) Histology - Ovary (ICD10: C56)

Number (#) and Relative Proportion (%)

Registry	Total	Unsp. Histology	
		#	%
TMH	576	117	20.3
KMIO	233	40	17.2
CI (WIA)	224	39	17.4
RCC - TVM	607	95	15.7
AMC	105	15	14.3
BBCI	408	25	6.1
PGIMER	578	39	6.7
BRAIRCH	265	48	18.1

**Table 11.12: Unspecified/Unknown Clinical Extent of Disease (CEDBT)
(Excludes Patients Previously Treated) - Both Sexes**

Number (#) and Relative Proportion (%)

Registry	Total	CEDBT Unknown	
		#	%
TMH	19932	10354	51.9
KMIO	7493	1741	23.2
CI (WIA)	7427	-	-
RCC - TVM	17737	-	-
AMC	3344	24	0.7
BBCI	19240	37	0.2
PGIMER	15644	657	4.2
BRAIRCH	5754	97	1.7