Comparability and Quality Control in Cancer Registration; Karachi (Data Monitoring 1995-2001)

Y. Bhurgr (Karachi Cancer Registry)
Y. Bhurgr,S. H. Hasan (Sindh Medical College)
Y. Bhurgr (Aga Khan University Hospital)
A. Bhurgri (Dow Medical College, Karachi.)

Abstract

Introduction: Comparability and Quality Control in Cancer Registration remains an essential component of standardisation of cancer incidence data. The information provided by registries is used as a base-line data for cancer control and cancer research. As such misleading information will encourage fictitious cancer patterns with over or under registration of some cancer sites.
Methodology: The registration in Karachi started in 1995 and was confined to a single district of Karachi, the South District. Since 1998 the registration has extended to Karachi Division. For both the data sets, recommendations of the International Agency for Research on Cancer (IARC) and the International Association of Cancer Registries (IACR) have been followed to assure comparability and Quality Control of the results.
Results: The 2 data sets from Karachi show well-recorded demographic details of cases. Age was not a true indicator of validity in our case, as we have approximate ages of patients aggregated around quintiles and decades. All the cancer cases were reported from a minimum of 3 sources. Microscopic verifications (MV) were high. Cancers with the highest histological verification (HV) were those located in easily accessible sites e.g. breast, oral cavity and cervix. The age specific incidence rates (ASR) for Karachi South, all sites (1995-1999) was 148.1males and 175.5 females. The same for Karachi Division (1998-99) was 132.4 males and 133.0 females.
Conclusion: The cancer data in 1995 during the initial phase of registration was incomplete and suffered from inaccuracies associated with a lack of awareness in early years of registration. By the year 2001, the data (1995-1999) showed improved completion and accuracy. The incidence rates have shown no dramatic changes during the past 2 years. The malignancies also show a stable pattern, despite the possibility of under-registration of clinically diagnosed cases. The microscopic verification is high compared with the world region and similar to those observed in developed countries. This may be due to the widely available diagnostic facilities in Karachi and should therefore be considered an indicator of validity of registered data (JPMA 52:301 ;2002).

Introduction

Comparability and Quality Control in Cancer Registration remains an essential component of standardisation of cancer incidence data. A cancer registry is a source of information and unreliable information is worse than no information. Quality control procedures are therefore instituted to identify the areas and degree of imperfection and thus assist in the interpretation of the data1. The information provided by registries is used as a base-line data for cancer control and cancer research. As such misleading information will encourage fictitious cancer patterns with over or under registration of some cancer sites. This is more likely to occur in developing countries where a lack of awareness of precise

data compounded with illiteracy may result in intentional and unintentional errors in reporting. It is here in these low resource settings where limited finances make it all the more necessary to organise a targeted cancer control program. Cancer registry data is also used to monitor cancer control programs and inaccurate data will further aggravate the existing discrepancies in the program. The International Agency for Research on Cancer (IARC) and the International Association of Cancer Registries (IACR) have laid down specific recommendations for improving the quality and comparability of registered data.

Completeness and accuracy of registered data are the decisive factors in the Quality Control of registries. Completeness of registration may be defined as the extent to which all the incident cancers occurring in a target population are included in the registry data base[2]. Completeness of detail refers to recording of essential items of information which could be related to the patient's identity, referral and to the details of malignancy.

Accuracy of detail, reporting and interpretation are also essential to maintain the quality of data. Consistency checks are able to detect some errors in detail, but vigilance and regular rechecks of data are required as an on-going program along with occasional ad-hoc surveys to completing tasks without awareness of the importance of databases. Tumour coding errors are also frequent. Moreover, staff that do not have first-hand knowledge of the intricacies of the computer file carries out programming of enquiries. Under these circumstances, reporting errors are quite likely to occur and unless they give rise to totally unexpected results, may well go undetected[1].

The methodologies implemented for evaluating completeness are based on data sources, independent case ascertainment and historical data methods[2]. The number of sources/notifications per case, death certificate method and microscopic verification of diagnosis can assess the completeness of data sources. Independent case ascertainment by checking independent data sets and rescreening of reported cases helps in the assessment of completeness and accuracy of registered data. Finally the historical data methods i.e. stability of incidence rates, comparison of incidence in different populations, age-specific incidence curves and childhood cancers help confirm the completion of data. The entire exercise will be futile if the denominator i.e. the population of the group under study is not precise and a recent census is essential for this.

The registration in Karachi started in 1995 and was confined to a single district of Karachi, the South District. Since 1998 the registration has extended to Karachi Division. Based on the above recommendations, we present here the factors of Comparability and Quality Control used to standardise the incident cancer data in Karachi, with stress on the more complete and mature Karachi South data. Data monitoring for both data sets has been conducted till 31st December 2001.

Methodology

Two cancer data sets are being analysed in the city, Karachi South and Karachi Division. The methodology remains identical for both. The larger data sets are provided in a computerised format by the hospitals concerned. Visits to the hospitals, laboratories and clinics are still conducted for abstracting data from records and interviewing the patients, re-check data and for the case-control studies. The information on cancer cases, is checked, coded and computerised. The number of notifications or sources recorded and the cases updated as further notifications are received.

The definition of 'incident date' followed, as per recommendation of Jensen et al[1] are: 'Date of consultation at, date of admission to, a hospital, clinic or institution for the cancer in question. Or (for cases not admitted to, or seen at a hospital), date of first diagnosis by a physician or date of the first pathology report (mentioning cancer). Date of death in case of DCOs[1]. This has been done to facilitate comparison of data internationally. As autopsies are not compulsory in Karachi, diagnosis on this basis is not possible.

The IARC and IACR recommendations[2] have been followed for factors influencing comparability i.e. classification, coding and mortality data. The data are classified using the International Classification of Diseases-Oncology (ICD-O) (2nd ed.)[3] and computerised using a customised version of CANREG-3 software. In addition SSPS, SAS and Epi-2000 are used for some analyses. Rules for certifying cause of death in Karachi are those established by WHO[2]. This death registry data is classified using the International Classification of Diseases 10th Revision (LCD 10) [4].
Validity is an essential component in assessing the quality of cancer registry data. The criteria used for evaluating are diagnostic criteria method, missing Information, reabstracting and recoding method and internal consistency method. The two indicators of validity used are cases with histological verification (HV) and cases registered on the basis of information only from a death certificate (DCO). Re-screening of cases at selective hospitals in Karachi is done annually. Duplicate registration files and the name files are checked bi-annually. The incidence tables were based on the census 1998. The age-structure for individual years being prepared at growth-rates calculated by the Federal Census Bureau[5].

Results

The registration system in Karachi at the onset in 1995 was confined to a single district, the South District Ccnsus 1998; estimated populations; growth-rate 1.94.
All the cancer cases were reported from a minimum of 3 sources, the primary-care hospital, laboratory records and at least I treatment centre. In many of the primary care hospitals there were at least 2 or 3 notifications per case.

Table 1. Mid-1997* Population Structure - Karachi South.

Age Group	1997* Male Population	1997* Female Population	1997* Total Population
0 - 4	128059	125509	253568
5 - 9	127876	120965	248841
10 - 14	114753	102243	216996
15 - 19	97892	84145	182037
20 - 24	86773	70026	156799
25 - 29	70551	60606	131157
30 - 34	56790	43503	100293
35 - 39	50501	42722	93223
40 - 44	45033	37885	82918
45 - 49	36193	27432	63625
50 - 54	32366	16589	48955
55 - 59	21429	14639	36068
60 - 64	17510	11908	29418
65 - 69	8487	6226	14713
70 - 74	8578	7696	16274
75+	8573	7994	16567
Total	911364	780088	1658638

*Census 1998; estimated populations; growth-rate 1.94

Table 2. Notifications per case in Karachi.

Notifications	%
One	16.8
Two	46.0
Three or more	37.2
Total	100.0

Table 3. Basis of Diagnosis, Karachi South (1995-99).

	Males %	Females %
0. Death Certificate Only	2.1	1.2
1. Clinical Only	1.1	0.5
2 Clinical Investigation	3.4	2.5
3. Exploratory Surgery	0.0	0.1
4. Biochemical	0.1	0.0
5. Cytology	5.0	7.5
6. Histology of Metastasis	6.8	4.7
7. Histology of Primary	77.5	80.3
8. Haematology	4.0	3.2
9. Unknown	0.0	0.0
Total	100.0	100.0

Table 4 . Most valid basis of Diagnosis by ICD-10 site categories-Karachi South.

Site	DCO	CL	HV	NOS	Total
Lip	-	-	100.0	-	26
Tongue	1.4	0.5	98.1	-	219
Salivary gland	-	4.8	95.2	-	61
Mouth	0.4	2.7	96.9	-	489
Oropharynx	-	-	100.0	-	58
Nasopharynx	1.8	-	98.2	-	53
Hypopharynx	-	-	100.0	-	119
Pharynx	12.1	6.1	81.8	-	33
Oesophagus	0.7	6.8	92.5	-	282
Stomach	0.5	2.5	97.0	-	182
Small Intestine	-	-	100.0	-	14
Colon	0.7	1.4	97.9	-	136
Rectum	-	3.9	96.1	-	124
Gall bladder etc.	0.8	5.0	94.2	-	121
Pancreas	-	13.9	86.1	-	35
Liver	7.1	10.3	82.6	-	179
Nose, sinuses etc.	-	-	100.0	-	18
Larynx	-	3.6	96.4	-	256
Bronchus, Lung	0.6	13.6	85.8	-	492
Oth. Thoracic Organs	-	5.9	94.1	-	12
Bones	-	6.8	93.2	-	116
Connective tissue	0.8	0.8	98.4	-	147
Mesothelioma	-	-	100.0	-	3
Melanoma of skin	-	-	100.0	-	23
Other Skin	-	0.7	98.6	0.7	218
Breast	0.2	2.3	97.5	-	1251
Cervix Uteri	-	0.7	99.3	-	141
Placenta	-	6.7	93.3	-	15
Corpus teril	-	1.0	99.0	-	102
Ovary etc.	-	2.4	97.6	-	208
Other Female Genital	-	-	100.0	-	17
Prostate	0.6	2.4	97.0	-	155
Testis	-	5.4	94.6	-	36
Penis	-	-	100.0	-	5
Bladder	-	1.5	98.5	-	268
Kidney etc.	-	1.4	98.6	-	71
Eye	-	4.0	96.0	-	25
Brain, Nervous System	2.2	5.0	92.9	-	192
Thyroid	-	2.0	98.0	-	130
Other Endocrine	-	-	100.0	-	13
Hodgkin's Disease	2.0	3.4	94.9	-	119
NHL	0.4	3.8	96.8	-	286
Multiple Myeloma	-	4.8	95.2	-	50
Lymphoid Leukemia	1.0	1.0	98.0	-	102
Myeloid Leukemia	1.5	-	98.5	-	132
Leukemia Nos.10.5	-	89.5	-	-	40
Other and Nos.10.0	6.1	83.8	0.1	-	799
Total					7575

Microscopic verifications (MV) include histological and cytological verifications. In the Karachi South data the histological verification (HV) for the years 1995-1999 was 89.3 % males and 92.5% females. The MV for the same was 93.3% males and 95.7% females. The clinically diagnosed cases were 4.6%

males and 3.1% females; the DCOs were 2.1% males and 1.2% females. The cancers with the highest HV were those located in easily accessible sites e.g. breast, oral cavity and cervix. Sites difficult to access e.g. lung and pancreas had the lowest percentage of histological verification. Less invasive techniques for diagnostic purpose, such as fine needle aspiration cytology (FNAC) and biopsy (FNAB) contributed to the high degree of MV, accounting for 20% of all such cases in the series.

The proportion of cases in Karachi South, with an unknown primary site for the years 1995-1999, were 494 out of the total of 3891 cases males and 305 out of the total of 3684 cases females, i.e. 12.7% in the males and 8.3% in the females. The cases with primary site unknown were coded as C80.9, C76, C39 and C26. The cases were detected either as lymph nodes with secondary malignancies or body fluids like ascitic fluid with malignant cells. Cancer series of selected consultants and re-screening of cases at hospitals in Karachi gave a 'missed case' rate of< 1% in June 2001. The incidence data for the year 1995 is slightly higher than in other years.

The Karachi Division data though accurate and comparable is not complete. There is a deficiency of 15-20% in the data. The HV for the years 1998-1999 was 75.3% males and 77.1% females. The MV for the same was 85.9% males and 91.3% females. The clinically diagnosed cases were 9.0% males and 4.3% females; the DCOs were 0.2% males and 0.1% females. FNAC and FNAB accounted for 10.6% and 14.2% of the cases in the males and females respectively. The proportion of cases with an unknown primary site for the years 1998-1999, were 1295 out of the total of 7396 cases males and 950 out of the total of 6847 cases females, i.e. 17.5% in the males and 13.9 % in the females. The age specific incidence rates (ASR) for Karachi South, all sites (1995-1999) was 148.1males and 175.5 females. The same for Karachi Division was 132.4 males and 133.0 females.

Table 5. Year-wise Incidence Data, Karachi South (Males).

Site	ICD-10	1995	1996	1997	1998	1999	1995-99
Bronchus, lung	C33-C34	20.9	16.3	25.9	20.6	15.9	20.0
O.Cavity	C00-C06	17.2	12.2	15.8	20.7	16.3	16.6
Larynx	C32	8.2	9.4	8.6	10.2	9.7	9.4
Bladder	C67	11.7	11.1	6.0	9.0	9.1	9.4
Pharynx	C09-C14	7.9	6.7	6.4	4.9	9.5	7.9
Oropharynx	C09-C10	2.8	1.9	1.2	1.5	2.4	2.0
Nasopharynx	C11	2.1	0.7	0.8	0.5	1.2	1.1
Hypopharynx	C12-C13	7.0	4.1	4.4	2.9	5.9	4.8
Lymphoma	C81-85, C96	9.9	6.4	7.4	7.8	7.4	7.9
Hodgkin's disease	C81	4.6	2.6	1.3	1.5	1.4	2.3
NHL	C82-C85,C96	5.3	3.8	6.1	6.3	6.0	5.6
Prostate	C61	6.4	5.9	7.3	10.1	7.9	7.6
Oesophagus	C15	5.2	7.5	6.8	4.4	6.8	6.2
Colo-rectum	C18-C21	5.8	5.6	4.6	5.3	8.2	5.9
Liver	C22	5.4	6.1	7.0	4.4	2.3	5.0
Stomach	C16	3.7	3.7	4.5	6.2	4.6	4.6
Leukaemia	C91-C95	3.3	3.9	3.9	4.5	4.1	4.0
Thyroid	C73	0.7	1.5	1.6	0.9	0.7	1.1
Testis	C62	0.5	1.1	0.6	1.4	0.9	0.9
Breast	C50	0.6	0.9	0.3	1.7	0.5	0.8
Skin Melanoma	C43	0.1	0.0	0.7	0.4	0.9	0.4
Others		46.1	45.6	47.2	54.1	46.0	39.8
All sites		150.3	137.9	147.5	158.2	141.8	148.1

1995 data - higher incidence rate due to prevalent cases.

1999 data - lower incidence rate due to under coverage.

Sites with a slowly increasing incidence rate are oral cavity, stomach, colo-rectum, prostate, non-Hodgkin's lymphoma and leukemia.

Table 6. Year-wise Incidence Data, Karachi South (Females).

Site	ICD-10	1995	1996	1997	1998	1999	1995-99
Breast	C50	62.0	48.8	47.9	59.6	63.7	56.6
O.Cavity	C00-C06	12.1	19.2	13.0	11.1	17.5	15.4
Ovary	C56	15.6	8.0	10.2	5.3	8.9	9.6
Cervix	C53	5.1	8.9	6.4	9.7	6.0	7.4
Oesophagus	C15	9.7	5.6	5.6	7.0	6.5	7.0
Lymphoma	C81-85, C96	4.7	3.3	6.1	6.4	7.3	5.7
Hodgkin's disease	C81	0.8	0.6	0.9	1.2	0.7	0.9
NHL	C82-85, C96	3.9	2.7	5.2	5.2	6.6	4.8
Uterus, all	C53-55	11.1	14.8	13.9	14.0	10.2	5.6
Colo-rectum	C18-21	9.4	4.1	3.3	3.4	4.9	5.0
Thyroid	C73	2.0	3.9	6.1	3.0	4.0	4.0
Leukemia	C91-95	3.6	3.9	4.6	3.8	3.9	3.9
Liver	C22	4.0	2.7	4.6	2.6	3.0	3.4
Stomach	C16	4.2	1.8	2.4	4.3	3.4	3.2
Pharynx	C09-14	2.4	5.0	2.6	3.1	1.8	3.1
Oropharynx	C09-10	0.7	1.1	0.1	0.5	0.2	0.5
Nasopharynx	C11	0.4	1.6	0.5	0.5	0.3	0.7
Hyopharynx	C12-13	1.3	2.3	2.0	2.1	1.3	1.9
Bladder	C67	3.9	3.0	3.6	2.8	2.2	3.1
Bronchus and Lung	C33-34	3.4	3.3	2.0	2.7	2.5	2.8
F.genital, other	C51-52; C57	2.2	0.2	0.2	1.2	0.6	0.9
Skin Melanoma	C43	0.8	1.0	0.2	0.4	0.0	0.5
Others		41.5	36.5	35.8	42.3	36.6	38.3
All sites		190.5	166.0	174.9	177.7	177.4	175.5

1995 data - higher incidence rate due to prevalent cases.

1999 data - lower incidence rate due to under coverage.

Sites with a slowly increasing incidence rate are oral cavity, oesophagus, stomach, uterus, Hodgkin's and non-Hodgkin's lymphoma.

Discussion

The data of Karachi South has been studied as a sample population of the country, in the absence of a

nationwide registration system. The cancer data of Karachi Division is more representative of the city of Karachi with a predominant migrant (mohajir) population from India. The cancer data in 1995 during the initial phase of registration was incomplete and suffered from inaccuracies hospices for cancer patients, an additional data source has emerged. The discrepancies in the data reported from more than one centre were cleared by direct interviews of the patients.

The inaccuracies were routine discrepancies of data collected by different investigators from multiple centres and especially when reported from more than one centre, by the centre itself. Some common names with lack of associated demographic data made the duplication very high. At present completeness of registration is improving and almost all the cancer cases are reported from a minimum of 3 sources, the primary-care hospital, laboratory records and at least one treatment centre. In many of the primary care hospitals there are at least 2 or 3 notifications per case per hospital. Usually the patients are admitted to the medical units for preliminary diagnosis and subsequently shifted to the surgical units and the specialities for subsequent surgical and oncology treatment. Thus the likelihood of missing a case, which enters the health-care system, is minimised. Lately with the development of as in all the registries in developing countries the Karachi Division data may suffer from some under-registration. The data of Karachi South has over the past few years stabilised and the incidence rates have shown no dramatic changes during the past 2 years. The malignancies also show a stable pattern. Despite this a possibility of under-registration of clinically diagnosed cases remains. Advanced cases of malignancies clinically diagnosed in outpatient departments and inoffices of clinicians are less likely to be subjected to intense diagnostic and therapeutic procedures. Records may not be available for some of these patients in hospitals or clinics. Consequently, some of these might not be registered. As such the cancer cases, which may not have come into contact with the health-care system, could have remained undiagnosed.

The section of the death certificate pertaining to cause of death though completed by medical practitioners is rather vague and non-specific. Therefore death certificates still remain of questionable value as indicators of completeness of data. Furthermore, inadequate death registration procedures, particularly in terms of cause-specific mortality, means that death certificate notifications do not constitute a major source of data. Despite this the DCOs were more frequent in the early years of functioning of the registry, especially 1995. The persons had cancer that in all probability was diagnosed before the registry became operational. The DCOs rate is gradually falling.

The two indicators of validity used are cases with histological verification and cases registered on the basis of information only from a death certificate. The MV in Karachi South is high compared with the world region and similar to those observed in developed countries. This may be due to the widely available diagnostic facilities in Karachi or to the insistence on part of the clinicians of establishing MV in almost all instances[6]. The high HV could well be considered a positive indicator of validity of information. This pattern has persisted over the past 6 years. This can be attributed to the similar qualitative methods of registration or to the characteristics of the health-care system or to a developed city in a developing country. Despite the high MV, it has not been possible to use the pathology data of all laboratories as primary data due to a lack of addresses.

The incidence data for the year 1995 reflects the registration of prevalent cases. Increase in the passive registration methodology has slowed down the registration process but helped develop good quality hospital-based data. The hierarchial order of the ten-most common tumours remains stable.

Age-specific incidence curves indicate an increase in the incidence with age and a slight decrease in the incidence in the oldest age group (over 70) has been observed. The latter change could be a consequence of a life-expectancy around ‘60’, or an actual decrease as many people may be dying of other competing cases of mortality or of a less efficient case ascertainment as the older age groups may not be entering the health-care system. Post-mortems not being done in all the cases, confirmation is not possible.

Conclusion

The cancer data in 1995 during the initial phase of registration was incomplete and suffered from inaccuracies associated with a lack of awareness in early years of registration. By the year 2001, the data (1995-1999) showed improved completion and accuracy. The incidence rates have shown no dramatic changes during the past 2 years. The malignancies also show a stable pattern, despite the possibility of under-registration of clinically diagnosed cases. The microscopic verification is high compared with the world region and similar to those observed in developed countries. This may be due to the widely available diagnostic facilities in Karachi and should therefore be considered an indicator of validity of registered data.

References

1.Jensen OM. Parkin DM. MacLcnnan R, et at, (Eds.). Cancer Registration: principles and Methods; IARC Scientific Publications No. 95 Lyon: IARC,1991.
2.Parkin DM, Chen VW, Ferley 3, at at. (Eds.). Comparability and Quality Control in Cancer Registration, IARC Technical Report No. 19, Lyon: IARC.1994.
3.World Health Organization. International Classification of Diseases for Oncology, 2nd ed. Geneva: WHO., 1990.
4.World Health Organization. International Classification of Diseases, 10th Revision, Geneva, WHO., 1992.
5.Census Bulletin-!: Government of Pakistan population and housing census of Pakistan 1998, Islamabad: Population Ccnsus Organisation, Statistics division, Federal Bureau of Statistics, 1998.
6.Bhurgn V. Bhurgri A, Hasan SH, eta!. Cancer Incidence in Karachi, Pakistan: first results from Karachi Cancer Registry. Int. J. Cancer,, 2000; 85, 325-329.