RESEARCH COMMUNICATION

Reliability of the Kuwait Cancer Registry: A Comparison between Breast Cancer Data Collected by Clinical Oncologists and Registry Staff

AA Elbasmi[1*], MS Fayaz[2], S Al-Mohanadi[3], Y Al-Nesf[4], A Al-Awadi[3]

Abstract

Aim: The quality of cancer registration is of great importance and the present study was conducted to assess the reliability of Kuwait Cancer Registry data on breast cancer. Methods: Data from the clinical records extracted by a group of clinical oncologists for another study on 1,235 breast cancer cases diagnosed between 1999 and 2004 were used to audit the data held on these individuals by the Kuwait Cancer Registry (KCR). Only 902 cases met the eligibility criteria. Main measures were sex, nationality, laterality, morphology, stage of disease at time of admission to the center, type of treatment and status at last follow up (alive or dead). Results: Full or high agreement between registry data and clinical oncologists collected data was recorded for sex, nationality and laterality. The rate of agreement for treatment with chemotherapy and status at last follow up was near perfect. Substantial agreement was also noted for morphology, tumor grade, TNM staging, surgical, radiotherapy and hormonal treatment. The majority of minor differences in morphology disagreements occurred when a more specific description was stated by registry staff, while major disagreement occurred due to difference in the codes used. Conclusions: The accuracy of the KCR data seems to be comparable to that found in reviews of other cancer registries. Stage was the hardest variable for the registry to collect accurate information on. KCR data could be improved by improving the quality of information provided to the registry.

Keywords: Cancer registry data - reliability - breast cancer - clinical records

Asian Pacific J Cancer Prev, **11**, 735-738

Introduction

The Kuwait cancer registry (KCR) is a population based register covering about three million Kuwaiti and non-Kuwaiti residents in kuwait. KCR established and in operation since 1971. Work flow and data sources are shown in Figure 1. KCR is a full member of the International Association of Cancer Registries (IACR). KCR is the first Arab and Gulf country to have its results accepted and published in "Cancer in Five Continents" since its fifth edition in 1990 (Curado M et al., 2007). Notification of cancer is compulsory by ministerial regulations. The registry collects information on malignant neoplasm according to the recommendations of the International Agency for Research on Cancer (IARC) (Curado M et al., 2007), as well as mortality data from the Vital and Health Statistics Division, of MOH, Kuwait (Health et al., 2007).

Breast cancer in women ranks as the most frequent site of cancer among both Kuwaiti and non-Kuwaiti females. Between January and December 2007, there were 312 newly diagnosed cases of breast cancer attending Kuwait cancer control center (KCCC) accounting for 21.7% out of 1,439 newly diagnosed malignant cases. Number of breast cancer cases totaled 168 (40.0%) among Kuwaiti females and represented 144 (40.9%) among non-Kuwaiti females. with age standardized incidence rates (ASR world) of 48.2/100,000 and 42.5/100,000, respectively (Elbasmi and Al-Asfour, 2010).

The International Agency for Research on Cancer (IARC) has identified five main areas used while assessing the quality of cancer registry data: completeness of cover, completeness of detail, accuracy of detail, accuracy of reporting, and accuracy of interpretation (Skeet 1991). As Kuwait cancer registry depends greatly on the completeness of the patients' medical records, its accuracy depends on the proper abstracting of data by the registry staff.

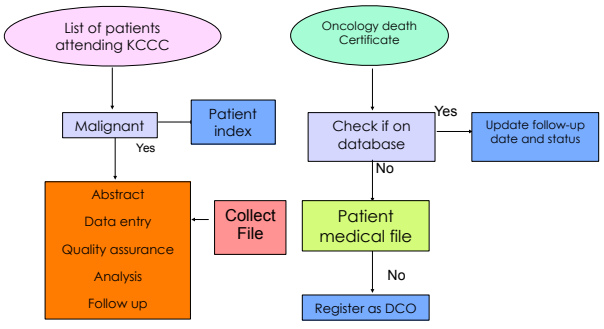

Figure 1. Work Flow at Kuwait Cancer Registry

*[1]Epidemiology and Cancer Registry Unit, [2]Clinical Oncology Department, Breast Cancer Unit, [3]Kuwait Cancer Control Center (KCCC), [4]Ministry of Heath, State of Kuwait, *For Correspondence: elbasmy@yahoo.com*

The aim of the study is to investigate the quality of the Kuwait Cancer Registry, through comparing data supplied by clinical oncologists with data collected by registry staff.

Materials and Methods

Between 1999-2004, 1480 breast cancer cases registered in Kuwait cancer registry, a group of clinical oncologists retrieved back only 1235 cases for another study. The data collected by the clinical oncologists used to audit the same data collected by the registry staff. Only 902 cases of Breast cancer patients were eligible for our study. Eligibility criteria included: 1) Cases of breast cancer diagnosed between first of January 1999 and 31 of December 2004; 2) Primary tumor site is breast: ICD10 C50.0-C50.9; 3) No missing data for any of the items included in the study.

The items included were sex, nationality, laterality, tumor grade, TNM staging, morphology, treatment received (surgery, chemotherapy, radiotherapy and hormonal) and status of last follow up. In order to make the registry data comparable with the clinical oncologists data, nationality variable was summarized as Kuwaiti and non-Kuwaiti, and morphology data was summarized to the three major most commonly seen infiltrating duct carcinoma (IDC), lobular carcinoma, and medullary carcinoma. Other morphological types including (adenocarcinoma, inflammatory carcinoma, and papillary carcinoma etc.) grouped under OTHER. TNM stages grouped according to Data divided into minor and major disagreements according to re-abstracting Centralized cancer patients' data system (CCPDS) study (1985) and summarized in Table 1.

Statistical Methods

Data of Kuwait cancer registry checked against the data collected by the clinical oncologists for the same patients. Clinical oncologists' data regarded as a "Gold standard", Agreement between registry and clinical oncologists' data identified using the positive predictive value as a measure of the percentage of all registry data confirmed by the Clinical oncologists. In addition, registry and clinical oncologists data was compared using $\varkappa$ statistic, which represent the extent of agreement beyond that expected by chance and the 95% confidence interval around this statistic (Hennekens and Buring, 1987). In general, $\varkappa$ statistics ranging from 0.41 to 0.60 have been interpreted to indicate "moderate" agreement, those ranging from 0.61 to 0.80 have been interpreted to indicate "substantial" agreement, and those of $\geq$ 0.81 have been interpreted to indicate "near perfect" agreement (Posner K et al., 1990). Further analysis of disagreement given the CCPDS definitions of coding presented as frequency and percentage. Calculation of all cross-tabulation and study statistics performed using SPSS version 16.

Results

In all, 902 cases of breast cancer registered by the Kuwait cancer registry between 1999 and 2004 met our criteria for inclusion in the study (Table 1), of which 203, 248, 219, 247, 277, 247, cases were registered in 1999, 2000, 2001, 2002, 2003, and 2004 respectively. Histology of primary was used to proof diagnosis in 708 (78.5%) cases, and 189 (21%) were proved by cytological examination.

The highest in exact agreement observed in sex and nationality items 99.8% followed by patient's status at last follow up 99.7%. Morphology, hormonal therapy and stage of disease had the lowest in exact agreement 89.6%, 86.4% and 76.2% respectively (Table 3). Rates of agreement for sex (92.3 and 99.9%) for male and female respectively, and (99.6 %and 100%) for Kuwaiti and non-Kuwaiti. Near perfect agreement between clinical oncologists and the registry for sex and nationality designated by the $\varkappa$ statistic of 0.992 (95% CI, 0.814-1.030) and 0.996 (95%

Table 1. CCPDS Definitions of Coding Disagreements and Adopted Standards

Item		Disagreement		Standard adopted *
		Minor	Major	
1	Sex	-	Any difference	100%
2	Nationality	Kuwaiti vs unknown	Kuwaiti vs non-Kuwaiti	100
3	Laterality	Other differences e.g Rt vs unknown	Rt vs Left	None
4	Morphology	Same first 3 digit with difference in last digit	Difference in fourth digit	96
5	Stage of disease at time of admission to center	Any difference within regional 2 vs.3 vs.4 vs. 5)	Any other difference	88
6	Surgical treatment	Difference not given, given (0,1) vs unknown (9)	Any difference not given vs given (0 vs 1)	96
7	Radiotherapy	Difference not given, given (0,1) vs unknown (9)	Any difference not given vs given (0 vs 1)	96
8	Chemotherapy	Difference not given, given (0,1) vs unknown (9)	Any difference not given vs given (0 vs 1)	96
9	Endocrine (hormonal therapy)	Difference not given, given (0,1) vs unknown (9)	Any difference not given vs given (0 vs 1)	96
10	Patient status at last follow up	Difference dead, alive (1,2) vs unknown (9)	Any difference dead vs alive (1 vs 2)	96

*% of cases free of major disagreements

Table 2. Positive Predictive Value (PPV) and Agreement Levels of Breast Cancer Data, 1999-2004

	KCR	CO	PPV	Kappa (95%CI)
Sex				
Male	13	13	92.3	0.92 (0.81-1.03)
Female	889	889	99.9	
Nationality				
Kuwaiti	460	462	99.6	1.00 (0.99-1.00)
Non Kuwaiti	442	440	100	
Laterality				
Rt	411	426	95.1	0.94 (0.92-0.96)
Lt	457	460	98.7	
Bilateral	16	16	100	
Morphology				
IDC (M-8500)	782	769	96.4	
Lobular (M-8520)	59	67	71.6	0.69 (0.62-0.76)
Medullary (M-8510)	23	24	70.8	
Other	38	42	57.1	
Stage of disease				
I	116	151	67.5	0.64 (0.60-0.68)
II	484	392	88.5	
III	221	300	62.7	
IV	81	59	84.7	
Surgery				
Yes	824	846	96.8	0.70 (0.61-0.80)
No	70	55	81.8	
Unknown	8	1	100	
Radiotherapy				
Yes	682	753	89.9	0.74 (0.68-0.79)
No	175	128	96.1	
Unknown	45	21	100	
Chemotherapy				0.83
Yes	721	742	96.1	(0.78-0.88)
No	151	151	87.4	
Unknown	30	9	100	

IDC, Infiltrating duct carcinoma

CI, 0.989-1.002) in that order.

Regarding laterality, the agreement rates for right come to 95.1%, 98.7% for left and 100% for bilateral (Table 2). This near perfect agreement confirmed by $\varkappa$ statistic of 0.942 (95% CI, 0.920-0.964).

The agreement found between radiotherapy collected data and the data collected by our staff was for morphology, stage of disease, treatment received (surgical, radiotherapy and hormonal) was of substantial value as showed by the $\varkappa$ statistic (Table 2) positive predictive value for those items ranged from 62.7% up to 100%. The rate of agreement of treatment with chemotherapy and status at last follow up was near perfect (Table 2).

Table 3 shows that among the 10 items included in the study of disagreement, the highest rates of major disagreement occur for endocrine (hormonal therapy) 12%, morphology 9% and radiotherapy 6%, whereas the remaining items all have major disagreement rates of 3 % or less. In particular, status patients' status at last follow up had the lowest major disagreement rate 0.3%. The minor disagreement rates for most items represent coding differences of detail. In morphology coding, for example, three out of the 3 of atypical medullary carcinoma (M-8513) were considered medullary carcinoma (M-8510) by clinical oncologists.

Disagreement by Item

Sex, nationality and laterality. In both sex and nationality, disagreement detected in 0.2% of data. As one female case, identified as male by both reviewers. Similar results encountered regarding nationality disagreement. Comparison of the laterality 27 major differences was found. 21/27 (78%) cases were reported by the registry as Lt were reported by the clinical oncologists as Rt and 6 Rt breast cancer cases by the registry were reported as lt by the clinical oncologists.

Morphology. Among the 83 major disagreements, 87% occurred due to conflict between codes (M-8500 infiltrating duct carcinoma, M-8520 lobular carcinoma and M-8522 infiltrating duct and lobular carcinoma). As the clinical oncologists used codes which were consistent with, but less specific than registry codes; 13% involved the coding of carcinoma, NOS adenocarcinoma, NOS, tubular adenocarcinoma, mucinous adenocarcinoma and inflammatory carcinoma as infiltrating duct carcinoma, instead of others. Majority of the minor differences in morphology disagreements 10/11, occurred when a more specific description was stated e.g (M-8500 infiltrating duct carcinoma vs M-8501 comedocarcinoma, NOS or M-8502 secretory carcinoma of breast or M-8503 intraductal papillary adenocarcinoma with invasion). Alternatively, M-8510 medullary carcinoma with M-8512 medullary carcinoma with lymphoid stroma, or M-8513 atypical medullary carcinoma. Or coding lobular carcinoma M-8520, as infiltrating duct and lobular carcinoma M-8522, or infiltrating duct mixed with other types M-8523).

Stage of disease at time of admission to center. The proportion of major disagreement between the two datasets for tumor stage of disease was 2.4% while minor disagreement occurred in 21% (Table 3). There were two main areas: between stage II and III with disagreement by both reviewers (with the KCR data showing a less advanced disease than the clinical oncologists' data.

Therapy. Majority of major disagreements in treatment was related to reporting treatment while none was given by KCR staff. Among 25/30 (83%) major disagreement of surgical treatment, 52/57 (91%) of radiotherapy disagreements, 19/27 (70%) of chemotherapy and 95/105 (91%) of hormonal treatment.

Vital status at last contact. There were only three disagreements as to whether the patient was alive at last contact.

Discussion

Kuwait cancer registry data is used for research as well as policymaking, and health service planning. Which largely depends on the accuracy and the quality of its data, and so it is important to know that data provided by the KCR are reliable in terms of demographic and diagnostic details.

Table 3. Results of the Comparison of the Data of Clinical Oncologists and Cancer Registry Staff

Items	Agreement		Disagreement			
			Minor		Major	
	N	%	N	%	N	%
Sex	900	99.8	-	-	2	0.2
Nationality	900	99.8	-	-	2	0.2
Laterality	875	97.0	-	-	27	3.0
Morphology	808	89.6	11	1.2	83	9.2
Stage of disease at time of admission to center	687	76.2	193	21.4	22	2.4
Surgical treatment	865	95.9	7	0.8	30	3.3
Radiotherapy	821	91.0	24	2.7	57	6.3
Chemotherapy	854	94.7	21	2.3	27	3.0
Endocrine (hormonal therapy)	779	86.4	18	2.0	105	11.6
Patient status at last follow up	899	99.7	-	-	3	0.3

The study found that cancer registration for breast cancer on the KCR was highly accurate with respect to demographic details, but less so for details relating stage of the tumor and health care provided. The accuracy of KCR data appears to be as good as other cancer registries in United States, United Kingdom, and the Netherlands (Lapham R et al., 1992; Gulliford M et al., 1993; Liu WL et al., 1995; Schouten LJ et al., 1997; Middleton RJ et al., 2000; Brewster DH et al., 2002; Thoburn KK et al., 2007) with stage was the most difficult variable for registries to abstract precisely. The key constraint of current study is that it presumes that the clinical oncologists can abstract perfect information from medical records (Schouten LJ et al., 1993). However, they supplied information with less than required detail. Since more detailed information was not directly important for their study. In addition some disagreements were to be expected knowing that, the variables in the two datasets were differently defined. This was particularly the case for Morphology, where the clinical oncologists used incomplete definitions.

Other discrepancies found relate to the time of information abstracted by KCR. As KCR code treatment received within the first 6-12 month after diagnosis. the present study showed the incompleteness of this approach "as regard this item" specially with no follow up for treatment given is done afterward, in contrast with the patients' status at last follow up where follow up by death certificates is done, the low PPV observed is due to the small number of deaths rather than lack of follow up. The differences seen in the staging item can be attributed to the usage of different TNM staging manual editions as the KCR uses the fifth while clinical oncologists used the Sixth edition.

In conclusion, the accuracy of the KCR data seems to be comparable to that found in reviews of other cancer registries. Stage was the hardest variable for the registry to collect accurately. KCR data could be improved by improving the quality of information provided to the

References

Brewster DH, Stockton D, Harvey J, et al (2002). Reliability of cancer registration data in Scotland, 1997. *Eur J Cancer,* **38,** 414-7.

Centeralized cancer patients' data system (1985). definition of major and minor disagreements and standards for reabstracting. In Quality control for cancer registries. Hilsenbeck, S.G., Glaefke, G.S., Feigl, P et al., pp 127-132. National institute of health, Government printing office: Bethesda Md.

Curado MP, Edwards B, Shin HR, et al (2007). Cancer Incidence in Five Continents, Vol. IX IARC Scientific Publications No. 160, Lyon, IARC.

Elbasmi A, Al-Asfour A (2010). Kuwait cancer registry annual report-2007, Kuwait cancer control center, Ministry of health; Kuwait.

Gulliford M, Bell J, Bourne H, et al (1993): The reliability of Cancer Registry records. *British J Cancer*, **67**, 819-21.

Health, Kuwait, edition 44 (2007). Health vital statistics division, department of statistics & medical records, Ministry of health, state of Kuwait.

Hennekens CH, Buring JE (1987). Epidemiology in Medicine. Boston:Little, Brown and Company.

Lapham R, Waugh NR (1992). An audit of the quality of cancer registration data. *Br J Cancer,* **66,** 552-4.

Liu WL, Kasl S, Flannery JT, et al (1995). The accuracy of prostate cancer staging in a population based tumor registry and its impact on the Black-White stage difference (Connecticut, United States). *Cancer Causes and Control,* **6,** 425-30.

Middleton RJ, Gavin AT, Reid JS, et al (2000). Accuracy of hospital discharge data for cancer registration and epidemiological research in Northern Ireland. *Cancer Causes Control,* **11,** 899-905.

Schouten LJ, Jager JJ, Van den Brandt PA (1993). Quality of cancer registry data: a comparison of data provided by clinicians with those of registration personnel. *Br J Cancer,* **68,** 974-7.

Schouten LJ, Langendijk JA, Jager JJ, et al (1997). Validity of the stage of lung cancer in records of the Maastricht cancer registry, The Netherlands. *Lung Cancer,* **17,** 115-22.

Posner KL, Sampson PD, Caplan RA, et al (1990). Measuring interrater reliability among multiple raters: an example of methods for nominal data. *Stat Med*, **9,** 1103-15.

Skeet RG (1991). Quality and quality control. In: Jensen OM, Parkin DM, editors. Cancer Registration: Principles and Methods. *IARC*, 101-7.

Thoburn KK, German RR, Lewis M, et al (2007). Case completeness and data accuracy in the Centres for Disease Control and Prevention's National Program of Cancer Registries. *Cancer,* **109,** 1607-16.