ARTÍCULO ORIGINAL

Data quality at population-based cancer registries in Colombia

Ricardo Cendales[1], Constanza Pardo[1], Claudia Uribe[2], Guillermo López[3],
María Clara Yepez[4], Luis Eduardo Bravo[5]

[1] Grupo de Vigilancia Epidemiológica del Cáncer, Instituto Nacional de Cancerología, Bogotá, D.C, Colombia
[2] Registro Poblacional (sic.) de Cáncer del Área Metropolitana de Bucaramanga, Universidad Autónoma de
  Bucaramanga, Bucaramanga, Colombia
[3] Registro Poblacional (sic.) de Cáncer de Manizales, Universidad de Caldas, Manizales, Colombia
[4] Registro Poblacional (sic.) de Cáncer de Pasto, Universidad de Nariño, Pasto, Colombia
[5] Registro Poblacional (sic.) de Cáncer de Cali, Universidad del Valle, Cali, Colombia

Introduction. Population-based cancer registries provide vital information for planning, prevention and cancer management. Information generated by the registries must be comprehensive, valid and comparable. Because of their importance, regular quality assessments are recommended.

Objective. The quality of cancer incidence data were assessed at four population-based cancer registries in Colombia for cancer incidence estimations.

Material and methods. Data collected at population-based cancer registries of Bucaramanga, Cali, Pasto, and Manizales were included. Completeness was assessed by the use of graphs in illustrating the mortality incidence ratios and their relation to the survival. Validity was evaluated by the description of morphologically verified cases, cases identified from death certificates only, and the internal consistency of the data.

Results. There was a global under-coverage of cancer registration at Bucaramanga and Manizales, whereas a more specific under-coverage for certain localizations was observed in Cali and Pasto. Validity analyses established that death certificates were little used as a source of information, and some inconsistencies appeared among the data associated with the most valid basis of diagnosis and morphology.

Conclusions. In Colombia, the data quality at population-based cancer registries can be further improved by considering the use of additional sources of information, such as death certificates, the use of specialized software for data capture, and automatic validation of internal consistency. Mortality certification must be improved in areas where a population-based cancer registry is operating.

Key words: vital statistics, neoplasms, quality control, data collection, reproducibility of results, Colombia.

Calidad de los datos en los registros de cáncer de base de población en Colombia

Introducción. La información de los registros de población de cáncer resulta de vital importancia en la planeación, prevención y manejo del cáncer. La información generada por los registros debe ser exhaustiva, válida y comparable, por lo que se recomienda hacer evaluaciones periódicas de calidad.

Objetivo. Evaluar algunos aspectos relacionados con la exhaustividad y la validez de la información recolectada por los registros de población de cáncer en Colombia para estimar las cifras de cáncer en el país.

Materiales y métodos. Se evaluó la información de los registros de población de Bucaramanga, Cali, Pasto y Manizales. La exhaustividad se analizó mediante las razones de mortalidad, incidencia y su relación con la supervivencia. La validez se estableció con la descripción de los casos verificados morfológicamente y los registrados en certificados de defunción. Finalmente, se describieron las inconsistencias generadas.

Resultados. El análisis de exhaustividad mostró un subregistro general en los registros de Bucaramanga y Manizales, y un subregistro para algunas localizaciones en todos los registros. El análisis de validez indicó un bajo uso de los certificados de defunción como fuente de información y

porcentajes importantes de inconsistencias entre la base de diagnóstico y la histología.

Conclusiones. La calidad de la información de los registros de población de cáncer en Colombia puede ser mejorada al considerar fuentes adicionales de información, el uso de *software* especializado en captura y validación, y fortalecimiento en el registro de la mortalidad en las áreas de cobertura.

Palabras clave: estadísticas vitales, neoplasias, control de calidad, recolección de datos, reproducibilidad de resultados, Colombia.

The control of malnutrition and communicable diseases, the increasingly ageing population, changes in lifestyle and exposure to occupational and environmental risk factors have increased the burden of cancer in developing countries. The World Health Organization has promoted the development of national programs aimed toward cancer control (1). Population-based cancer registries have major relevance for this purpose. They are valuable sources of information for estimating cancer burden, planning cancer care facilities, targeting and monitoring interventions, and setting priorities for cancer control and prevention (2).

The Cali Cancer Registry (CCR) is the most important source of descriptive epidemiology in Colombia and South America (3). It has been in continuous operation since 1962. Although the CCR includes only a small fraction of the incident cancer cases in Colombia (4), the information therein has been used as the main input for estimating cancer incidence at the national level (5). Other regions in Colombia such as Antioquia, Barranquilla, Bucaramanga, Manizales, Cartagena, Cesar, Huila, and Pasto have replicated the efforts of CCR and have provided preliminary results (6-10).

Information generated by cancer registries must be comparable and reproducible in order to accomplish validity. At the initial stages of establishing a cancer registry, efforts must be made to achieve collection of complete and high quality data (11). Periodic quality evaluations must be performed to guarantee the reliability of data produced by cancer registries, particularly in developing countries with a scarcity of resources, facilities and training. The technical report published by the International Agency for Research on Cancer (IARC), describes methodologies for judging comparability and quality

control in cancer registration (12). An updated methodological approach has been published recently in a series of two methods articles (13,14) and one applied research article (15). Herein, an updated methodological proposal of four dimensions are described: comparability, completeness, validity and timeliness.

Comparability is essential for the interpretation of collected data. A critical aspect is the standardization of codification and collection practices. Completeness consists in evaluating how many of the incident cases are effectively recorded by the population-based cancer registries. Validity refers to the proportion of the identified cancer cases with a given characteristic that actually have that attribute. Timeliness of reporting of a cancer registry is also a key aspect in cancer registries (13).

Completeness can be evaluated through qualitative, semi-quantitative and quantitative methods. One of the semi-quantitative methods is generated through the analysis of the mortality: incidence ratio (M:I); this ratio compares cancer deaths with incident cancer cases, specifically for each localization in the same time period. If quality of both sources is good, under some specific circumstances, the M:I ratio will be approximately 1 — a given five-year survival probability in the area of the registry or in an area with similar conditions. Distortions in this ratio will suggest completion issues that can be easily identified from simple graphic analyses (14).

Validity can be evaluated through four groups of methods: reabstracting and recoding, diagnostic criteria methods (histological verification, and death certificate only), missing information analyses, and internal consistency methods. Reabstracting and recoding is used to evaluate the concordance between collectors and between the collectors and the original data sources. Diagnostic criteria methods are performed through the percentage of cases morphologically verified and categorized by sex and site, as well as the proportion of cases identified by death certificates only. Missing information analyses describe the percentage of ill-defined sites, age unknown, and unknown basis of

Corresponding author:
Ricardo Cendales, Grupo de Vigilancia Epidemiológica del Cáncer, Instituto Nacional de Cancerología, Calle 1 Nº 9-85, Bogotá, D.C., Colombia
Phone: (571) 334 1111
acardocen@yahoo.com

diagnosis (13). Internal consistency evaluation is established by the description of errors identified through a standardized validation program (16).

In Colombia, two studies of cancer registries have been performed. The first one evaluated twelve institution-based cancer registries and considered six domains: human, technological, logistical and political resources, quality of the information and information diffusion (17). The second study considered eight population-based cancer registries that evaluated coverage of information sources, information validity, quality assurance, comparability and sustainability (Pontificia Universidad Javeriana. Memorias del Primer Congreso Nacional de Salud Pública. Evaluación de los registros poblacionales de cáncer en Colombia. Bogotá: Facultad de Ciencias Económicas y Administrativas; 2006). However, no study has evaluated the quality of information collected in the databases of these registries.

The aim of the current study is to provide an evaluation of some aspects related to validity and completeness of the information collected by four population-based cancer registries in Colombia [Bucaramanga Cancer Registry (BCR), Cali Cancer Registry (CCR), Manizales Cancer Registry (MCR), and Pasto Cancer Registry (PCR)]. These four registries are among the seven currently active in Colombia, and they have produced updated cancer information for a period of at least five years. The information generated herein will be used in taking decisions for the inclusion of the information provided by these four population-based cancer registries, in the validation and estimation processes used in generating cancer incidence estimations in Colombia.

Materials and methods

Sources of information

Official databases of the BCR (2000-2005), the MCR (2002-2006), and the PCR (1998-2002) were obtained. Incidence data and consistency indicators for CCR (1998-2002) were not processed, but obtained from published data (5). Only confirmed cases residing in the target area of the cancer registry were included in the incidence analyses. The IARCcrgTools software was used to create equivalences between the International Classification of Diseases for Oncology, third edition (ICD-O 3) and the International Classification of Diseases, tenth edition (ICD-10). Conversion errors and warnings were described for each registry with the IARCcrgTools software (18). Cases with

conversion errors were excluded from the analyses; cases with conversion warnings were reviewed but not excluded.

Official national mortality databases of the national department for vital statistics (DANE in Spanish for *Departamento Administrativo Nacional de Estadística*) were obtained for the respective time periods in each cancer registry. A complete analysis of mortality quality was performed as a secondary objective and independently published. The study concluded that mortality certification in Colombia and particularly in the population-based cancer registries areas has high quality standards for overall death and cancer death certification (19). Cancer deaths not reported by physicians were excluded from the analyses. The city of habitual residence was not imputed by the occurrence city, because imputing this information leads to an overestimation of deaths for two reasons: (1) because most cancer deaths occur in the main cities or capital cities of each district and (2) because cancer patients are usually referred to specialized health services located at capital cities. Ill-defined natural deaths and ill-defined cancer deaths were not redistributed. Deaths from uterine cancer not otherwise specified were proportionally redistributed among deaths from uterine cancer in specified sites, for each registry and age group, using methods derived of probabilistic sampling (20). The traditional coverage correction for mortality of 79% estimated for Colombia (21) was not applied to the mortality data, because of a presumed very high mortality coverage in the registry areas.

M:I ratio was calculated by making use of the traditional groups for cancer localizations (22). Only deaths of people with habitual residence in the area of registries were considered in the M:I ratio calculation. The analyses were limited to cancer localizations with more than twelve cases registered in the period because lower numbers are associated with greatly fluctuating values that can make the estimates and data analyses difficult (23).

The five-year accumulated probability of survival was based on data from the Cancer Registry of Norway (24), since no local estimations of survival are available, and established specifically for each localization, considering all stages, for cases diagnosed in 1999-2003.

Statistical methods

Databases were consolidated in a single file with all the critical information required for the analysis

(identification number, birth date, incidence date, sex, topographical code, morphological code, tumor behavior, histological differentiation, and basis of diagnosis). Frequencies and percentages were employed to describe categorical variables. Completeness graphical analyses were performed through bivariance scatter plots. Data were analyzed using the SPSS software version 15.0.

Aspects described in completeness evaluation

M:I ratio was calculated for each registry, sex and cancer localization. A bivariate scatter plot was used for graphical analysis, depicting the M:I ratio in the Y axis and the 1 minus 5 year cumulated probability of survival in the X axis. Localizations with M:I ratios higher than 1.6 were excluded from graphical analyses. A qualitative analysis of the graphical information was then performed.

Aspects described in validity evaluation

Percentage of cases morphologically verified, and death certification only cases were described specifically for each registry according to sex and localization. For each registry, the percentage of cases with unknown age or unknown or invalid basis of diagnosis were also described. Inconsistencies between age and birth date, age/site/histology, site/histology, sex/site, sex/histology, behavior/site, behavior/histology, grade/histology and basis of diagnosis/histology were identified using the IARCcrgTools software (18).

Results

Incidence data analysis

A total of 7,590 incident cases of the BCR, 2,709 of the PCR, and 2,785 of the MCR were analyzed. Access was unavailable to the CCR databases for analysis of the 16,660 cases reported in this registry; however, the CCR quality indicators were published previously by IARC (5). ICD-O 3 to ICD-10 conversions demonstrated one case that was unable to be converted in the BCR because of an invalid combination of sex, localization or histology; 45 cases were not converted in the MCR because they did not have valid ICD-O 3 codes; the PCR did not demonstrate conversion inconsistencies. Two conversion warnings were described for the MCR due to an improbable combination of behavior/site/histology; no other conversion warnings appeared for the BCR or the PCR. After the conversion, 31 cases of the BCR and 1 case of the MCR that were originally classified as malignant, were reclassified as benign or of uncertain behavior; these 32 cases were

also excluded. No cases were excluded because of lack of information on habitual residence, age or sex, except in the CCR where 707 cases were excluded because of an unknown age.

Several cancer localizations were excluded from the analysis because fewer than twelve cases were recorded in the period. These included the following codes, according to the ICD-10: small intestine (C17), other and ill-defined digestive organs (C26), nasal cavity, middle ear and accessory sinuses (C30-C31), thymus, heart, mediastinum, pleura and other and ill-defined sites within respiratory system and intrathoracic organs (C37-C39), bones, joints and articular cartilage (C40-C41), mesothelioma (C45), Kaposi sarcoma (C46), peripheral nerves and autonomic nervous system (C47); connective, subcutaneous and other soft tissues (C49), eye and adnexa (C69), adrenal gland and other endocrine glands and related structures (C74-C75), malignant inmunoproliferative diseases (C88) and, multiple myeloma and plasma cell tumors (C90). Specifically for males breast (C50), penis (C60) and, other and unspecified males genital organs (C63) were excluded; for female sex, larynx (C32) and placenta (C58) were excluded.

Mortality data analysis

The total deaths analyzed were as follows: 5,355 in Bucaramanga and its metropolitan area (2000-2005), 2,046 in Manizales (2002-2006), 1,346 in Pasto (1998-2002), and 9,105 in Cali (1998-2002). In Bucaramanga and its metropolitan area, 5 deaths were excluded from the analyses because they were not certified by a physician, 32 more deaths were excluded because they were of unknown age. In Cali, 11 deaths were excluded because they were not certified by a physician and 114 more were excluded because they were of unknown age. In Manizales, only 2 deaths were excluded because they were of unknown age. In Pasto, 31 deaths were excluded because they were not certified by a physician, 30 more deaths were excluded because they were of unknown age.

Completeness analysis

The BCR demonstrated high M:I ratios for all localizations; however, in males unusual values were observed for pancreas, liver and gallbladder, and in females, unusual values were observed for liver, lung, gallbladder, stomach, esophagus and kidney (Table 1, figure 1a). The CCR demonstrated coherent values both for men and women with the exception of liver that showed a high M:I ratio (Table

1, figure 1b). The MCR showed high M:I ratios in general for all sites; with extreme values in men for liver, pancreas, esophagus, lung and leukemia; in women for liver, pancreas, esophagus, gallbladder, lung, central nervous system, and leukemia (Table 1, figure 1c). The PCR showed consistent values both in men and women for all sites, with the exception of several unusual values for lung, pancreas and liver in both sexes (Table 1, figure 1d).

Validity analysis

The analysis of cases established only from death certification revealed a low use of death certifications as a source of information in the BCR and the MCR. In the three registries (BCR, MCR, and PCR), a low percentage of morphologically verified cases were present for certain localizations such as liver, gallbladder, pancreas, lung, kidney, multiple myeloma and leukemia. The remaining localizations had, in general, a higher use of morphological verification (Table 2).

The validation software found one error in the combination of sex and localization in the BCR; in the MCR an error was found related to date of birth in relation to the date of incidence, 37 errors in the code of the histological type, 9 errors in the code of topographic localization, 32 errors in the age at

diagnosis, and one error in the combination of sex and localization. No errors were found in the PCR. This software found also other inconsistencies. The most frequent was related to basis of diagnosis and histology, followed by localization and histology. The registry with the lowest inconsistencies was the BCR. The MCR and the PCR exhibited similar values for inconsistencies (Table 3).

Discussion

This is the first data quality evaluation based on the information provided by population-based cancer registries in Colombia. The methodology was in accordance with the recently published recommendations by Bray and Parkin (13,14). The current study was not as comprehensive, however, since it did not consider all the components described for data quality evaluation at population-based cancer registries but rather was focused on validity and completeness. The present time was appropriate for evaluation since most registries had just produced their first consolidated results, and thereby provided an opportunity for improving processes established for collecting and registering cancer incident cases.

The lack of local estimates of overall survival for most cancer sites became a major limitation of

Table 1. Mortality: Incidence ratios according to sex, site, and cancer registry in Colombia

Sites	BCR[a] 2000-2005		CCR[b] 1998-2002		MCR[c] 2002-2006		PCR[d] 1998-2002	
	Males	Females	Males	Females	Males	Females	Males	Females
Lip, oral cavity and pharynx	0.73	0.45	0.48	0.35	0.73	0.45	0.31	0.18
Esophagus	1.18	1.37	0.87	0.76	1.18	1.37	0.74	0.58
Stomach	1.18	1.55	0.74	0.75	1.18	1.55	0.61	0.65
Colon, rectum, anus	0.72	0.72	0.57	0.59	0.72	0.72	0.52	0.65
Liver	2.98	2.63	1.83	1.70	2.98	2.63	1.08	1.12
Gallbladder	1.86	1.80	0.65	0.70	1.86	1.80	0.59	0.62
Pancreas	4.46	4.09	1.11	0.99	4.46	4.09	1.12	1.67
Larynx	0.83	0.90	0.5	0.70	0.83	0.9	1.50	NAe
Lung, trachea and bronchus	2.17	2.19	0.88	0.87	2.17	2.19	1.33	1.52
Melanoma of skin	0.45	0.37	0.19	0.15	0.45	0.37	0.36	0.10
Breast		0.36		0.34		0.36		0.31
Cervix uteri		0.56		0.44		0.56		0.41
Corpus uteri		0.32		0.30		0.32		0.03
Ovary		0.85		0.47		0.85		0.55
Prostate	0.43		0.35		0.43		0.42	
Testis	0.30		0.22		0.30		0.09	
Kidney	0.95	1.26	0.47	0.39	0.95	1.26	0.54	0.75
Bladder	0.51	0.51	0.38	0.34	0.51	0.51	0.26	0.41
Brain and central nervous system	0.93	0.90	0.69	0.70	0.93	0.90	0.87	0.79
Thyroid	0.22	0.09	0.29	0.07	0.22	0.09	0.27	0.13
Non-Hodgkin lymphoma	0.69	0.60	0.37	0.38	0.69	0.60	0.35	0.36
Hodgkin lymphoma	0.62	0.54	0.29	0.38	0.62	0.54	0.60	0.63
Leukemia	0.93	1.06	0.75	0.72	0.93	1.06	0.61	0.81

[a]BCR: Bucaramanga Cancer Registry, [b]CCR: Cali Cancer Registry, [c]MCR: Manizales Cancer Registry, [d]PCR: Pasto Cancer Registry, [e]NA: No data available because there were no registered cases.
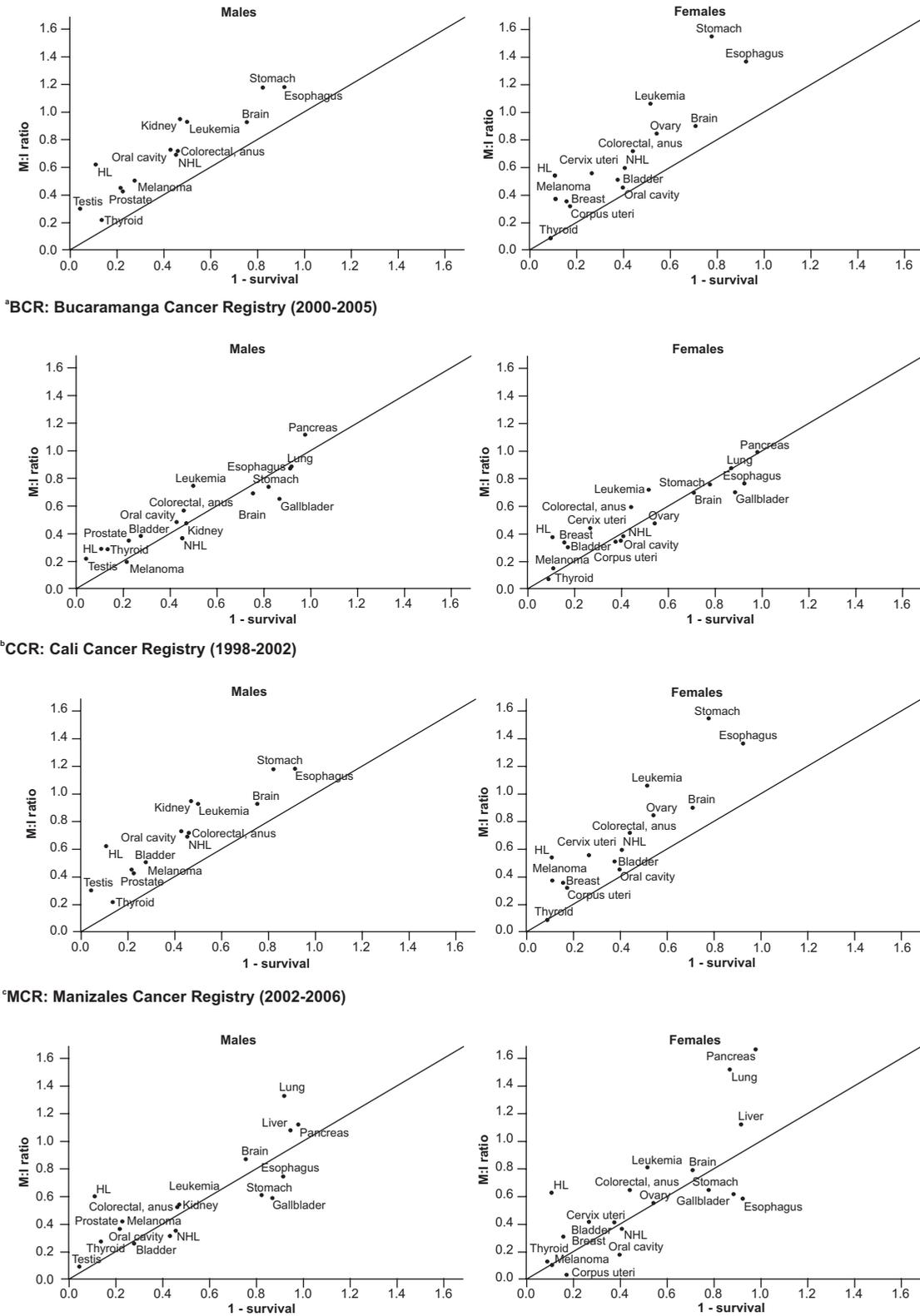
ᵃBCR: Bucaramanga Cancer Registry (2000-2005)

ᵇCCR: Cali Cancer Registry (1998-2002)

ᶜMCR: Manizales Cancer Registry (2002-2006)

ᵈPCR: Pasto Cancer Registry (1998-2002)
*Localizations with M:I ratios higher than 1.6 were excluded from graphical analyses.

Figure 1. Mortality: Incidence ratios of ᵃBCR, ᵇCCR, ᶜMCR, AND ᵈPCR versus 1-survival

Table 2. Indices of data quality, according to sex, site, and cancer registry in Colombia.

Sites	BCR[a] 2000-2005				CCR[b] 1998-2002 *				MCR[c] 2002-2006				PCR[d] 1998-2002			
	Males		Females		Males		Females		Males		Females		Males		Females	
	MV[e]	DCO[f]	MV	DCO	MV	DCO	MV	DCO	MV	DCO	MV	DCO	MV	DCO	MV	DCO
Lip, oral cavity and pharynx	97.3	0.0	98.7	0.0	86.7	3.9	92.0	0.6	100.0	0.0	96.7	0.0	75.0	18.8	82.4	5.9
Esophagus	97.2	0.0	100.0	0.0	79.5	9.8	79.5	10.3	85.7	0.0	94.7	0.0	80.0	17.1	83.3	16.7
Stomach	97.6	0.0	95.4	0.0	76.7	10.1	71.5	15.4	97.2	0.0	93.2	0.0	81.0	13.9	70.9	20.6
Colon, rectum, anus	98.7	0.0	98.1	0.0	84.1	6.6	83.0	7.1	95.2	0.0	96.2	0.0	88.6	6.8	75.4	10.8
Liver	97.7	0.0	94.7	0.0	66.7	8.6	68.5	8.9	66.7	16.7	52.9	0.0	53.8	30.8	28.0	48.0
Gallbladder	95.2	0.0	100.0	0.0					69.2	7.7	70.0	0.0	76.5	11.8	61.7	25.0
Pancreas	76.9	0.0	91.3	0.0	36.9	14.9	35.7	21.4	50.0	0.0	46.2	0.0	41.2	23.5	23.8	38.1
Larynx	100.0	0.0	100.0	0.0	85.1	3.3	68.1	8.5	100.0	0.0	100.0	0.0	100.0	0.0	NA	NA[g]
Lung, trachea and bronchus	79.7	0.0	78.7	0.0	55.3	12.7	56.8	13.6	73.3	1.0	82.3	0.0	45.7	23.9	64.0	32.0
Melanoma of skin	97.6	0.0	100.0	0.0	96.8	0.0	99.3	0.0	95.7	0.0	100.0	0.0	100.0	0.0	1000	0.0
Breast			98.2	0.0			92.9	2.2			91.1	0.0			89.3	4.2
Cervix uteri			99.5	0.0			93.8	2.1			94.1	0.0			88.9	8.2
Corpus uteri			100.0	0.0			87.8[h]	6.3			94.1	0.0			86.5	8.1
Ovary			92.7	0.0			72.8	7.7			90.2	0.0			82.8	6.9
Prostate	98.4	0.0			85.1	5.1			95.5	0.0			71.9	15.7		
Testis	97.7	0.0			93.0	0.9			96.4	0.0			100.0	0.0		
Kidney	92.5	0.0	94.7	0.0	80.6	2.2	79.5	6.0	88.9	0.0	86.4	0.0	69.2	7.7	75.0	8.3
Bladder	92.8	0.0	95.3	0.0	88.7	2.0	83.5	3.7	95.1	0.0	93.8	0.0	82.9	8.6	82.4	11.8
Brain and central nervous system	92.0	0.0	90.1	0.0	78.8	2.8	76.1	4.8	80.6	0.0	78.6	7.1	73.9	4.3	58.3	33.3
Thyroid	92.7	0.0	95.0	0.0	89.4	2.4	97.8	0.0	91.7	0.0	81.8	0.0	100.0	0.0	90.1	2.8
Non-Hodgkin lymphoma	97.2	0.0	95.4	0.0	90.5[j]	0.4	93.0	0.6	95.9	0.0	93.9	0.0	87.7	5.3	87.3	3.6
Hodgkin lymphoma	93.1	0.0	91.7	0.0					100.0	0.0	100.0	0.0	86.7	0.0	100.0	0.0
Leukemia	68.5	0.0	70.3	0.0	91.1	0.3	90.0	1.7	71.4	0.0	70.0	0.0	17.1	7.3	40.5	16.2
Total	**94.5**	**0.0**	**94.9**	**0.0**	**79.2**	**6.3**	**82.9**	**6.0**	**91.5**	**0.3**	**90.2**	**0.1**	**77.7**	**11.1**	**81.2**	**10.8**

[a]BCR: Bucaramanga Cancer Registry; [b]CCR: Cali Cancer Registry, *Source: Cancer Incidence in Five Continents IX; [c]MCR: Manizales Cancer Registry; [d]PCR: Pasto Cancer Registry; [e]MV: Morphologically verified; [f]DCO: Death certificate only; [g]NA: No data available because there were no registered cases; [h]Corpus uteri includes C54-C55 in CCR; [j]Non-Hodgkin Lymphoma includes Non-Hodgkin lymphoma and Hodgkin lymphoma in CCR.

Table 3. Percentage of consistencies according to cancer registry in Colombia.

Consistency	BCR[a] (2000-2005)	MCR[b] (2002-2006)	PCR[c] (1998-2002)
Age/site/histology	99.7	99.5	99.9
Basis of diagnosis/histology	99.6	97.6	92.0
Site/histology	99.8	99.1	99.3
Behavior/histology	100.0	99.9	100.0

[a]BCR: Bucaramanga Cancer Registry, [b]MCR: Manizales Cancer Registry, [c]PCR: Pasto Cancer Registry

this study, since Norway cancer survival data (24) cannot be directly applied to the Colombian populations. The graphical behavior of the M:I ratio in relation to 1−survival showed that all localizations are very close to the diagonal line in CCR. This indicated that Norway survival was an appropriate surrogate for local survival. However, the data were increasingly more distant in the PCR, BCR and MCR. This was particularly true for the female data, suggesting differences in the recording of the cancer incidence for females. The high M:I ratio put in evidence a lack of identification of cancer cases from death certificates. These very high M:I ratios may also be associated with misclassifications in the habitual residence of the deceased, since most cancer cases occurring in small, distant towns are referred to district capitals. Therefore, deaths occurring in district capitals may be certified as people habitually residing in the district capital when this is not in fact the case. Consequently, a formal definition of habitual residence must be implemented for death certification in areas where population-based cancer registries are established. Physicians in charge of death certification in these areas must be duly trained in the proper filling of death certificates (25).

Because all of the registries, including CCR, demonstrated high M:I ratios for liver, pancreas

and lung, we hypothesized a systematic error in death certification. It is possible that some of the metastatic liver and lung tumors could have been erroneously registered as primary tumors of liver and lung in death certificates. Other reasons for the high M:I ratios observed for esophagus, gallbladder and leukemia may be systematic errors in incidence data collection where relevant data sources are omitted such as endoscopic centers, general hospitals, or hematological clinics. Strategies such as establishing personal contact with general surgeons, internists or hematologists working in the area with the purpose of reporting cases that probably will not be registered by the traditional data sources have been successfully implemented (26). Electronic media for case report will also enhance the opportunity and the completeness of the population-based cancer registries in Colombia.

The analysis of the diagnostic confirmation method demonstrated that only CCR and PCR use death registries as a source of information. This finding confirmed the hypothesis of a lowered recording of cancer incidence in highly lethal cancers such as liver, gallbladder, pancreas, lung, multiple myeloma and leukemia. Population-based cancer registries in Colombia are aware of this problem (7); in fact, the MCR started to include the information from death certificates immediately after the conclusion of the current study. Exploration of alternatives is recommended that are aimed to facilitate the direct or indirect access to mortality databases. Other countries with similar difficulties have achieved unrestricted access to mortality databases and have successfully accomplished the inclusion of mortality information in cancer registries (26,27). In this regard, making use of the national mortality databases (not local mortality databases) is to be commended, because a sizeable percentage of cases are recorded in a city different from the city of residence.

The percentage of cases with morphologically verified cases and with known age was satisfactory in all registries, although it signalled the very low percentage of cases with unknown age and the high percentage of morphologically verified cases. This may be due to an incomplete search of cases, probably based solely in pathology centers; this may also be the explanation for the high M:I ratios. Validity analysis showed a low but persistent percentage of inconsistencies between the basis of diagnosis and the histological type; it demonstrated the lack of software capable of fully validating the data input. The experience of CCR and the

institutional-based cancer registry of the Instituto Nacional de Cancerología may help to develop the activities of the younger registries. The PCR demonstrated better validity indicators, reflecting the assistance provided by the CCR.

Future studies must verify that a completely depurated database is obtained from registries, because some validation processes that are routinely applied to the information collected by the registries cannot be applied to the databases. This situation will create an underestimation of the data quality.

The extensive effort required to collect cancer incidence information in developing countries must continue, but quality of the collected data must be evaluated periodically and systematically in order to achieve high quality standards that permit the unrestricted use of the information provided by young registries. If this is not accomplished, the information will continue to be collected at an enormous expense, but without a clear impact on cancer information due to low quality standards.

References

1. **World Health Organization.** National Cancer control programmes: Policies and managerial guidelines. Second edition. Geneva: World Health Organization; 2002.

2. **Jensen OM, Storm H.** Purposes and uses of cancer registration. In: Jensen OM, Parkin DM, Maclennan R, Muir CS, Skeet RG, editors. Cancer registration: Principles and methods. Lyon: International Agency for Research on Cancer; 1991. p. 7-21.

3. **Universidad del Valle.** Registro Poblacional (sic.) de Cáncer de Cali. Cali: Facultad de Salud, Universidad del

Valle. Fecha de consulta: 14 de julio de 2010. Disponible en: http://rpcc.univalle.edu.co/in/

4. **Parkin DM.** The evolution of population-based cancer registry. Nat Rev Cancer. 2006;6:603-12. http://dx.doi.org/10.1038/nrc1948

5. **Curado MP, Edwards B, Shin HR, Storm H, Ferlay J, Heanue M, *et al.*** Cancer incidence in five continents, Vol. IX. Lyon: International Agency for Research on Cancer; 2009.

6. **Brome M, Galeano L, Santamaría I, Salas C, Álvarez A.** Registro poblacional de cáncer de Antioquia. Revista Epidemiológica de Antioquia. 2006;28:7-27.

7. **Uribe CJ, Meza EE.** Cancer incidence in Bucaramanga Metropolitan Area, 2000-2004. First five years report from population-based cancer registry of Metropolitan Area of Bucaramanga. MedUNAB. 2007;10:147-72.

8. **López G, Vanegas M, Escobar M.** Registro poblacional (sic.) de cáncer Manizales − Caldas, resultados 2006. Quinquenio 2002-2006. Manizales: Universidad de Caldas; 2009.

9. **Secretaría de Salud Departamental del Huila.** Registro poblacional (sic.) de cáncer del Huila. Boletín epidemiológico No. 1. Neiva: Registro poblacional (sic.) de cáncer de Huila; 2007.

10. **Bolaños H, Hidalgo A, Yépez MC.** Incidencia de cáncer en el municipio de Pasto, 1998-2002. San Juan de Pasto: Universidad de Nariño; 2007.

11. **Valsecchi MG, Steliarova-Foucher E.** Cancer registration in developing countries: luxury or necessity? Lancet Oncol. 2008;9:159-67. http://dx.doi.org/10.1016/S1470-2045 (08) 70028-7

12. **Parkin DM, Chen VW, Ferlay J, Galceran J, Storm HH, Whelan SL.** Comparability and quality control in cancer registration. Lyon: International Agency for Research on Cancer and International Association of Cancer Registries; 1994.

13. **Bray F, Parkin DM.** Evaluation of data quality in the cancer registry: principles and methods. Part I. Comparability, validity and timeliness. Eur J Cancer. 2009;45:747-55. http://dx.doi.org/10.1016/j.ejca.2008.11.032

14. **Parkin DM, Bray F.** Evaluation of data quality in the cancer registry: Principles and methods Part II. Completeness. Eur J Cancer. 2009;45:756-64. http://dx.doi.org/10.1016/j.ejca.2008.11.033

15. **Larsen IK, Småstuen M, Johannesen TB, Langmark F, Parkin DM, Bray F, *et al.*** Data quality at the Cancer Registry of Norway: An overview of comparability, completeness, validity and timeliness. Eur J Cancer. 2009;45:1218-31. http://dx.doi.org/10.1016/j.ejca.2008.10.037

16. **Ferlay J, Burkhard C, Whelan S, Parkin DM.** Check and conversion programs for cancer registries (IARC/IACR tools for cancer registries). Lyon: International Agency for Research on Cancer; 2005.

17. **Cuervo LG, Roca S, Rodríguez MN, Stein J, Izquierdo J, Trujillo A, *et al.*** Evaluation of institutional cancer registries in Colombia. Rev Panam Salud Pública. 1999;6:202-6. http://dx.doi.org/10.1590/S1020-49891999000800008

18. **International Association of Cancer Registries.** Software IARCcrgTools Version 2.05. Built Jun 9 2008. Fecha de consulta: 30 de junio de 2010. Disponible en: http://www.iacr.com.fr/iacr_iarccrgtools.htm

19. **Särndal C, Swensson B, Wretman J.** Basic ideas in estimation from probability simples. In: Särndal C, Swensson B, Wretman J, editors. Model assisted survey sampling. New York: Springer-Verlak; 1997. p. 24-58.

20. **Mathers CD, Fat DM, Inoue M, Rao C, López AD.** Counting the dead and what they died from: An assessment of the global status of cause of death data. Bull World Health Organ. 2005;83:171-7. http://dx.doi.org/10.1590/S0042-96862005000300009

21. **Cendales R, Pardo C.** Colombian death certificate quality, 2002-2006. Rev Salud Pública (Bogotá). 2011;13:229-38. http://dx.doi.org/10.1590/S0124-00642011000200005

22. **Parkin DM, Whelan SL, Ferlay J, Raymond L, Young J.** Cancer incidence in five continents. Vol. VII. Lyon: International Agency for Research on Cancer; 1997.

23. **Devesa SS, Grauman DJ, Blot WJ, Pennello G, Hoover RN, Fraumeni JF Jr.** Atlas of cancer mortality in the United States, 1950-94. Washington, D.C: US Government Print Office; 1999.

24. **Cancer Registry of Norway.** Cancer in Norway 2008 - Cancer incidence, mortality, survival and prevalence in Norway. Oslo: Cancer Registry of Norway, 2009. Fecha de consulta: 12 de julio de 2010. Disponible en: http://www.kreftregisteret.no/Global/Publikasjoner%20og%20rapporter/CiN2008part1_web.pdf

25. **Departamento Administrativo Nacional de Estadística.** Metodología estadísticas vitales. Colección documentos - Actualización 2009 Num. 82. Fecha de consulta: 16 de marzo de 2010. Disponible en: http://www.dane.gov.co/files/investigaciones/fichas/Estadisticas_vitales.pdf

26. **Navarro C.** The National Death Index: A largely expected advance in the access to mortality data. Gac Sanit. 2006;20:421-3.

27. **Morales DL, McClellan AJ, Jacobs JP.** Empowering a database with national long-term data about mortality: The use of national death registries. Cardiol Young. 2008;1(Suppl.2):S188-95. http://dx.doi.org/10.1017/S1047951108002916