Pergamon

European Journal of Cancer Vol. 30A, No. 8, pp. 1128–1133, 1994
Elsevier Science Ltd
Printed in Great Britain
0959–8049/94 $7.00+0.00

0959-8049(94)00211-8

Cancer Registration: a Feasibility Study in Northern Italy

S. Rodella, C. Picoco, F. Stanzial, M. Turazza, L. Fiore Donati and the Collaborative Group of Pathologists for Cancer Registration in Verona

A pilot study was carried out in the province of Verona, in the north-east of Italy, in order to assess the feasibility of establishing a population-based cancer registry in the area. The quality of routinely collected data, particularly histological diagnoses and hospital discharge codes, was evaluated for the year 1988. All the histologically confirmed incident cancers observed in the pathology departments of the study area were registered and compared to the expected cases. Moreover, computerised discharge codes were tested for accuracy in identifying hospitalised patients with cancer. Finally, the sensitivity of the two sources combined was measured (90.3%). This study could provide helpful information to cancer registries which intend to assess the quality of specific sources of data both for planning and periodical evaluation purposes.

Key words: cancer registries

Eur J Cancer, Vol. 30A, No. 8, pp. 1128–1133, 1994

INTRODUCTION

POPULATION CANCER registries collect information from different sources such as hospital records, histological diagnoses, death certificates and specific additional files [1]. Since the majority of patients with cancer are likely to undergo histological confirmation, clinical admission or both, the accuracy of these two sources can rightly be considered a crucial point in cancer registration.

The aim of this study was to assess the accuracy of histological diagnoses and discharge codes as main sources of data in cancer registration. The evaluation of completeness, not considered in this study, should be regarded as a subsequent step, to be achieved with well standardised methods, as already described by many authors [2, 3].

MATERIALS AND METHODS

The province of Verona, in the north-east of Italy, includes six local health units (LHU) and a population of 775 745 residents (1981 National Census). Eighteen public hospitals and seven private clinics are located in the area covered by a computerised Hospital Discharge Diagnosis Coding System (HDDCS). Five departments of pathology are active in the province. At the time the study was started, two of these had a computerised registration of diagnoses, using Systematised Nomenclature of Medicine (SNOMED) [4] and the International Classification of Diseases for Oncology codes (ICD-O) [5], respectively.

All histological diagnoses (HD) of malignant tumours performed between 1 January 1988 and 31 December 1988 in the pathology laboratories of the study area were collected using two different approaches. All records of four departments were systematically scrutinised by the authors in order to select all cancer diagnoses. Computer files extracted according to an ICD-O morphology code between 8000 and 9990 (fifth digit 1–9) were directly submitted by only one of the two computerised departments. All selected cases were individually checked against the total number of previous HDs in order to exclude prevalent cancers.

Incident cases were registered, and basic information concerning residence and identification of both patient and cancer were recorded. Primary sites of all cases were coded according to the International Classification of Diseases (ICD-IX) [6]. Topography and morphology codes were attributed to all, except computerised cases, according to the ICD-O. Malignant tumours were defined as those coded to the 140-208 section of the ICD-9 and for which the ICD-O fifth digit behaviour code was 3 or more, with the exception of bladder and brain tumours, which were all attributed to the ICD-IX codes 188 and 191, respectively, notwithstanding their morphology and behaviour, according to suggestions of the IACR [7] and criteria of other Italian registries [8]. Non-melanoma skin cancers were not registered.

Cytological and autopsy examinations were collected but not included in the present analysis. All coding procedures were carried out according to a previously defined protocol by two of the authors, and uncertain cases were jointly reviewed and classified. Incident cancers selected according to the described procedure (observed) were compared to the expected (both total and histologically-proven) incident cases according to the indirect standardisation method [9]. Expected numbers were

Correspondence to S. Rodella at the Istituto di Anatomia Patologica, Policlinico Borgo Roma, Via Le Grazie 8, Verona, Italy.
S. Rodella, C. Picoco, F. Stanzial, M. Turazza and L.F. Donati are at the Pathology Department; F. Stanzial is at the Biological Sciences Institute; and M. Turazza is at the Oncology Division, University of Verona, Verona, Italy. The members of the Collaborative Group of Pathologists for Cancer Registration in Verona are Enrico Buongiorno, Gianni Calabrese, Paola Castelli, Maurizio Lestani, Filippo Marino, Roberto Mencarelli and Giuseppe Pelosi.
Revised 7 Mar. 1994; accepted 17 May 1994.

obtained by 1978–1987 incidence rates and proportion of histological verification (HV) of the Lumbardy Cancer Registry (LCR) that covers a comparative area in northern Italy as regards population, cancer mortality and distribution of health facilities. Comparisons were performed for both sexes and all age categories.

A sample of 14583 clinical records, including all patients discharged from the 18 public hospitals in January 1988, was identified through the computerised HDDCS. The sample represented 9% of the 162043 admissions to the hospitals in 1988. A 3-month preliminary survey carried out on 322 patients discharged from private hospitals with a 140-239 code showed that 145 patients (45.0%) were not resident, 156 (48.4%) had previous or subsequent admissions to a public hospital or an HD in a pathology laboratory (or both), 18 (5.6%) were prevalent cases and only 3 (0.9%) could be regarded as incident cases. Based on these results, we decided to exclude private clinics from this study, assuming that only a small proportion of incident cancers would come from this source.

Clinical records were systematically reviewed by two authors for all patients included in the sample, and cancers in resident patients were registered and coded (registry code, RC) according to the ICD-IX. Computerised hospital discharge codes (HDC) concerning the same selected sample were collected from the HDDCS. Codes between 140 and 239 (malignant tumours, benign tumours, *in situ* carcinomas and tumours of uncertain nature or behaviour) were selected for resident patients. Codes above 208 were included to encompass all types of bladder and brain tumours, and to overcome a possible under-registration since, in our experience, many malignant tumours can be erroneously coded as benign or of uncertain nature or behaviour.

The HDCs were considered a "screening test" to detect cancer cases, and their accuracy, in terms of sensitivity, specificity and predictive value (PV), was measured versus RCs, which were considered the golden standard.

Finally, incident cases identified through clinical records were linked to HDs and HDCs. Incident cases were grouped in four categories, according to the results of the linkage:

(1) cases reported by both sources (both HD and 140-239 HDC available);
(2) cases reported only by histology (HD available but HDC not included in 140-239);
(3) cases reported only by HDDCS (HDC in 140-239 but HD not available);
(4) "missing" cases, detected through systematic examination of clinical records (neither HD nor a 140-239 HDC available).

The number of cases in categories 1–3 was regarded as a measure of sensitivity of the information system based on HDs and HDCs combined.

Confidence intervals were all calculated using CIA (confidence interval analysis) microcomputer program [10]. The major steps of the procedure described are illustrated in Figure 1.

RESULTS

Through HDs, 2637 incident cancers in resident patients were identified. The comparison to the expected number of total and histologically-proven incident cases is reported in Table 1, for both sexes and ICD-IX categories. Cancers of selected sites, particularly the oral cavity and pharynx (140-149), small intestine, colon and rectum (152-154), metastases (196-199), non-Hodgkin's lymphomas (200, 202, 204) and non-lymphatic leukaemias (205-208) were grouped together in an attempt to avoid small numbers and misclassification. Moreover, lymphatic leukaemias are grouped with non-Hodgkin's lymphomas in the most frequently used clinical and pathological classifications of these malignancies [9–11].

In order to more appropriately interpret the observed/expected (O/E) ratios, we also compared the 1980–1989 mortality for cancer for both the study and the reference area, on the basis of national statistics (ISTAT). No remarkable differences were found, except for a significantly lower mortality for gastric cancer in the Verona district (Table 2). This is an expected result, since LCR covers a high incidence area for this cancer in Italy [8].

Eight hundred and forty-seven cases of cancer (RCs included in 140-208) were identified in resident patients through a review of the 14583 clinical records, and of the 847, 257 were incident cancers (30%). Nine hundred and seventy-four HDCs in 140-239 were selected from the same sample. Since the HDC does not allow for the discrimination between incident and prevalent

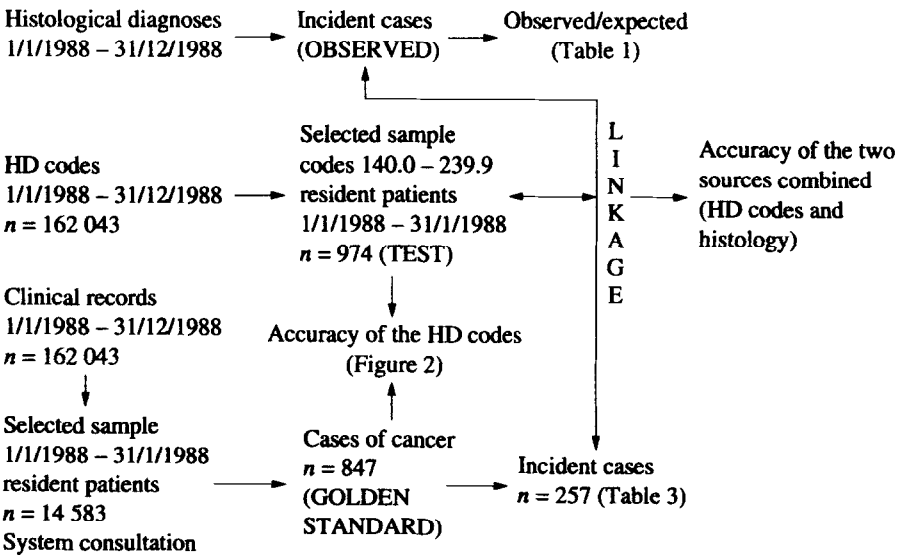

Figure 1. The major steps of the feasibility study for cancer registration in Verona.

Table 1. Comparison between observed and expected number of incident cases of cancer with histological verification in the Verona province, 1988, according to primary site, both sexes

ICD IX	Site	Expected (total *n*)	Expected (histological) E	Observed (histological) O	O/E	95% CI lower	95% CI upper
140–149	Oral cavity/pharynx	103.31	97.07	105	1.08	0.89	1.31
150	Oesophagus	46.00	33.96	44	1.30	0.94	1.74
151	Stomach	313.21	234.88	182	0.77‡	0.67	0.89
152–154	Small intestine, colon, rectum	394.64	325.92	361	1.11	0.99	1.23
155	Liver	67.99	26.73	30	1.12	0.76	1.60
156	Gall bladder	42.84	26.71	22	0.82	0.52	1.34
157	Pancreas	80.90	27.56	22	0.80	0.50	1.20
158, 159	Other digestive sites	14.80	5.45	10	1.83	0.88	3.32
161	Larynx	80.21	75.44	125	1.66‡	1.38	1.98
162	Bronchus, lung	471.40	223.49	269	1.20†	1.07	1.36
163	Pleura	5.32	2.94	5	1.70	0.75	4.45
160, 164–165	Other respiratory sites	9.86	8.43	2	0.24*	0.02	0.86
170	Bone	9.93	8.25	3	0.36	0.08	1.06
171	Connective tissue	15.72	15.16	31	2.04‡	1.39	2.90
172	Skin (melanoma)	41.05	40.30	60	1.49†	1.14	1.93
174	Breast, females	379.72	339.98	400	1.18†	1.06	1.29
175	Breast, males	2.39	2.12	1	0.47	0.01	2.63
180	Uterus, cervix	47.71	46.32	37	0.80	0.56	1.10
182	Uterus, corpus	83.92	82.95	62	0.75*	0.58	0.96
183	Ovary	62.40	52.28	49	0.94	0.69	1.24
179, 181, 184	Other female genitals	17.78	13.70	19	1.39	0.83	2.17
185	Prostate	127.03	90.97	149	1.64‡	1.39	1.93
186	Testis	15.11	14.90	9	0.60	0.28	1.15
187	Other male genitals	5.15	5.15	6	1.17	0.43	2.54
188	Bladder	187.90	168.02	160	0.95	0.81	1.11
189	Other urinary	78.92	60.56	107	1.77‡	1.45	2.14
191	Brain	46.86	24.27	23	0.95	0.60	1.42
190, 192	Eye, other nervous system	7.25	5.80	3	0.52	0.10	1.50
193	Thyroid	32.47	28.91	34	1.18	0.82	1.63
194	Other endocrine	3.84	2.94	1	0.34	0.01	1.89
195	Undefined sites	5.07	0.97	1	1.03	0.03	5.74
196–199	Primary site uncertain	74.84	28.84	112	3.88‡	3.21	4.69
201	Hodgkin's disease	30.45	29.68	29	0.98	0.65	1.41
200, 202, 204	Non-Hodgkin's lymphoma	125.14	111.22	148	1.33‡	1.13	1.57
203	Multiple myeloma	29.97	23.02	5	0.22‡	0.07	0.51
205–208	Non-lymphatic leukaemias	49.51	37.01	11	0.30‡	0.15	0.53
140–208	All sites	3110.61	2321.90	2637	1.14‡	1.09	1.18

*$P < 0.05$. †$P < 0.01$. ‡$P < 0.001$.

diseases, a 2 × 2 table could not be correctly completed solely for incident cancers and HDCs were considered a screening test for all cancer patients. Sensitivity (79.1%), specificity (97.4%) and positive and negative PVs (68.8 and 98.5%, respectively) were measured. The results of the comparison between HDCs and RCs are reported in Figure 2. The results of the linkage between the 257 incident cancers and all cases identified by the two sources combined are reported in Table 3. 159 cases (61.9%) were reported by both sources, 125 cases (78.6%) having a complete third digit level agreement between the HDC and ICD-IX code assigned to the HD. 52 cases (20.2%) were reported by HDCs only and 21 cases (8.2%) were reported by HDs only. Therefore, a total number of 180 incident cases (70%) had HV and 211 had a 140-239 HDC. Sensitivity of the HDCs with respect to incident cancers was 82.1% (CI 77.4–86.8%) and 170 of these cases (80%) had a complete third-digit-level agreement between HDC and RC.

In 25 patients (9.7%), an HD was lacking and the HDC was not included in the 140-239 section. These cases would have been overlooked had clinical records not been examined. Most of these patients were 70 years and older, and affected by malignancies of the digestive tract, lung and haematolymphopoietic system. As a final result, the sensitivity of the two sources combined was 90.3%.

DISCUSSION

While reliability of established cancer registries is usually based on the evaluation of completeness, quality assessment of specific sources of data can be a useful preliminary step when planning a cancer registry. Data from pathology laboratories and hospital discharges usually encompass the great majority of incident cases of cancer. This study assessed the accuracy of these two main sources in a geographical area of northern Italy where, given the existence of a considerable number of hospitals, the majority of resident patients with cancer are most likely to have their first diagnosis within their own area. The size of the

Table 2. Comparison between cancer mortality of the study area and reference area major sites, both sexes, 1980–1989, standardised rates per 100 000 (standard population: Italy 1981)

ICD IX	Site	Verona (study area)	LCR (reference area)	Rate ratio
140–149	Oral cavity/pharynx	7.18	6.31	1.14
150	Oesophagus	7.00	5.70	1.23
151	Stomach	19.32	30.71	0.63*
153–154	Colon and rectum	22.76	26.12	0.87
155	Liver	10.97	10.60	1.03
156	Gall bladder	4.32	3.56	1.21
157	Pancreas	10.30	9.84	1.05
161	Larynx	6.37	5.42	1.17
162	Bronchus, lung	50.18	53.87	0.93
172	Skin (melanoma)	2.10	2.27	0.93
174	Breast, females	36.34	37.66	0.96
179	Uterus, NOS	6.22	7.15	0.87
180	Uterus, cervix	1.42	1.29	1.10
182	Uterus, corpus	1.61	1.96	0.82
183	Ovary	8.85	9.21	0.96
185	Prostate	21.84	19.30	1.13
188	Bladder	6.93	8.34	0.83
189	Other urinary	5.57	4.93	1.13
191	Brain	4.98	5.34	0.93
196–199	Primary site uncertain	8.29	6.49	1.28
200 + 202 + 204	Non-Hodgkin's lymphoma	7.36	7.69	0.96
201	Hodgkin's disease	1.24	1.63	0.76
205–208	Non-lymphatic leukemias	4.78	5.77	0.83
140–239	All sites	239.30	256.55	0.93

LCR, Lumbardy Cancer Registry. *$P < 0.05$. NOS, not otherwise specified.

population and the total number of expected incident cancers per year in the study area are quite close to the median population size per registry (955 000) and to the median number of new cases observed (3500) in cancer registries of the European Community [12].

In our study, we focused more on the comprehensive accuracy in the identification of cancers than on precision in the definition of each case. For this reason, some ICO-IX categories were grouped together and results were not analysed according to topography and morphology, although this information was available. This is a reasonable approach in a feasibility study

		Clinical records: Cancer	Other	Total
Hospital	140-239	670	304	974
Codes	Other	177	11562	11739
	Total	847	11866	12713*

Sensitivity: 79.1%
Specificity: 97.4%
Positive predictive value: 68.8%
Negative predictive value: 98.5%

Figure 2. Accuracy of the screening test (hospital discharge codes) in identifying cases of cancer. *Number of clinical records of resident patients (total number of clinical records examined = 14 853).

primarily concerned with general objectives, such as potential for under-registration and general estimate of the workload required. It is implied that more refined criteria should be used when assessing the quality of registration in a well-established system.

In the Verona district, the observed histologically-proven incident cancers were compared to the number of expected cases, based on the known incidence rates and proportions of HV of the LCR. This method could be biased, namely because of the small size of the reference population (leading to inaccurate estimates of the expected numbers) and possible differences in incidence rates between the two areas. In order to minimise the first bias, we chose an incidence period as long as possible (10 years). Moreover, a substantial influence of the second factor can be reasonably excluded since mortality rates for different cancer sites seem not to differ significantly, with the exception of gastric cancer (Table 2).

The total number of observed cases with HV significantly exceeded the expected number (Table 1). When major sites were considered, the O/E ratio significantly exceeded 1 for larynx, lung, melanoma, breast, prostate, other urinary cancers and non-Hodgkin's lymphomas. Moreover, with the exception of lung cancer, the observed cases largely exceeded even the total number of expected cases. The O/E ratio was significantly lower than 1 for gastric cancer, multiple myeloma and non-lymphatic leukaemias. These results should be interpreted with some caution because of the potential for methodological biases. Firstly, erroneous inclusion of prevalent cases cannot be excluded. Secondly, the relatively remote time period used in the computation of the expected numbers can overlook temporal

Table 3. Results of the linkage between the 257 incident cases of cancer notified by the consultation of clinical records and all cases notified by the two sources of data (histological reports and hospital discharge codes)

ICD IX	Site	HDC +HV	HDC only*	HV only	Missing cases†	Total
140–149	Oral cavity/pharynx	7	1	0	0	8
150	Oesophagus	2	0	0	1	3
151	Stomach	13	4	1	1	19
152–154	Small intestine, colon, rectum	28	2	7	4	41
155	Liver	5	4	1	2	12
156	Gall bladder	2	1	0	0	3
157	Pancreas	3	1	0	3	7
158,159	Other digestive sites	1	0	0	0	1
161	Larynx	8	0	0	0	8
162	Bronchus, lung	26	13	1	4	44
163	Pleura	0	0	1	2	3
160, 164–165	Other respiratory sites	1	0	0	0	1
172	Skin (melanoma)	1	0	0	1	2
174	Breast, females	24	3	2	0	29
180	Uterus, cervix	1	0	0	1	2
182	Uterus, corpus	2	0	0	0	2
183	Ovary	2	2	0	0	4
179, 181, 184	Other female genital	1	0	0	0	1
185	Prostate	8	1	2	1	12
188	Bladder	8	1	2	0	11
189	Other urinary	6	4	0	0	10
191	Brain	1	1	0	0	2
190, 192	Eye, other nervous system	0	1	0	1	2
193	Thyroid	0	1	0	0	1
196–199	Primary site uncertain	1	6	1	1	9
200–202, 204	Lymphomas	8	0	3	0	11
203	Multiple myeloma	0	2	0	1	3
205–208	Non-lymphatic leukaemias	0	4	0	2	6
140–208	All sites	159	52	21	25	257

HDC, hospital discharge code; HV, histological verification. *The total number of HDCs concerning incident cases was 303, 92 of which were not malignant tumours. Therefore, positive predictive value = 69.6% (159 + 52/303). †These cases would have not been registered unless systematic consultation of clinical records had not been performed. The sensitivity of the two sources combined is: (159 + 52 + 21)/257 = 90.3%.

trends occurring in incidence rates for several sites of cancer [8], partly accounted for by some recent advances in diagnostic tools and a more diffuse attitude toward an earlier detection. As a consequence, a possible underestimate of the expected cases could partly exlain the excesses observed in our study. While observations concerning gastric cancer are consistent with mortality data and may reflect a true difference in incidence rates, the low number of multiple myelomas and non-lymphatic leukaemias represent an unexpected result. All these observations should be confirmed by further investigations. The remarkably high O/E ratio in the ICD-IX category 196-199 (metastases, primary site uncertain) and the high proportion (4.2%) of all categories including undefined primary sites (196-199, 159, 179) deserve an additional comment. This result reflects difficulties in attributing a primary site when histology is performed on metastatic tissues. More generally, histopathological reports may also lack precision in defining the site of origin in some specific organs, namely upper digestive tract and colon-rectum (pooled together in our study). Uncertainties concerning these items could be ultimately solved through examination of clinical records.

A specific quality control of the codes attributed by pathologists was carried out in the two computerised pathology laboratories on a selected sample of 120 cases, representing 2.2% of all HDs with some mention of cancer. In 14 cases, some incongruencies were found between the extensive diagnosis and the code. Four *in situ* bladder cancers were coded as invasive, 7 cases with a specific morphology (i.e. mucinous adenocarcinoma) were classified with a more general code (i.e. adenocarcinoma), 2 cases were attributed to a wrong topography and one oral cancer was erroneously coded as a skin cancer. Although these observations should be confirmed on a larger sample from all pathology laboratories they suggest that a certain lack of precision could derive from the use of computerised pathological data.

The accuracy of HDCs in identifying all cancers (Figure 2) was measured with respect to the examination of clinical records as reference sources for diagnosis. False negatives occurred in 177 out of 847 (21%). A comparable proportion of false negatives (21 + 25/257, 17.9%) was observed in the sample of incident cases (Table 3). Therefore, an important influence of the date of first diagnosis on the sensitivity of HDCs can be excluded in cancer patients.

While sensitivity of a source of data is a measure of completeness, the positive predictive value (PPV) is an indirect measure of the number of cases uselessly identified (false positives), representing the extra work to be carried out to avoid over-registration. The observed PPV of 68.8% means that approximately 30% of the 140-239 codes actually correspond to benign tumours or non-neoplastic pathologies. Restricting the test to

140-208 codes improved the PPV to 96.9% but reduced sensitivity to 66.3%. An analysis of the 257 incident cases gave comparable results, since 211 of these were coded to the 140-239 section (sensitivity = 82.1%) but only 178 were coded to the 140-208 section (sensitivity = 69.3%).

A possible interpretation of the negative predictive value (NPV) as an indirect measure of "missing" cases (false negatives) of cancer should also be mentioned. In Figure 2, a NPV of 98.5% means that a 1.5% proportion of "other" codes actually refer to cancer patients. Nonetheless, high NPVs may also be observed in the presence of a low sensitivity because of the high prevalence of non-neoplastic diseases among hospitalised patients. In our opinion, NPV can be a misleadingly optimistic measure of the number of "missing" cases.

Measuring validity of HDCs gives useful indications for the strategy to be adopted in defining cases solely reported by this source. Firstly, a large range of codes is preferable in order to ensure a higher sensitivity. Secondly, the clinical record should always be examined thoroughly in order to eliminate false positives.

The small sample size of the 257 "true" incident cases, in spite of the large number of clinical records at first examined, suggests some caution in drawing conclusions. Nevertheless, the number of cases identified, which corresponds to a 1-month survey, is very close to the mean number of cases expected monthly (3110/12 = 259), reasonably excluding an important under-registration. The observed proportion of HV (70%) is comparable to the proportion of 74.8 reported by the LCR in 1978–1987. The 159 cases with double notification can approximate the proportion to be possibly defined on the basis of computerised data. However, cases with severe disagreement between HD and HDC should always undergo further investigation even though, in our experience, histology is the most reliable source. We finally obtained a crude estimate of the expected number of clinical records to be reviewed each year by linking the 974 HDCs reported in Figure 2 with a computerised file, including all 140-239 HDCs of the previous 3 years (1985–1987), in order to eliminate prevalent cases. A residual number of 303 HDCs identified potential incident cases, and 144 of these, including 52 true positive HDCs only (Table 3) and 92 false positives (303 minus 211), represent the proportion of clinical records to be examined, which is at least 50% (144/303). Therefore, approximately 1700 clinical records are expected to be reviewed each year. This number could be partly reduced by extending collection of HDs to benign tumours.

The most important message is provided by the 90.3% sensitivity of the two sources combined, which implies a certain proportion of hospitalised incident cancers possibly overlooked in the routine registration. Cytological diagnoses, autopsies, HDCs from private clinics and mortality data are expected to increase the sensitivity in detection of hospitalised patients and contribute to the general completeness of cancer registration in the Verona district.

The primary purpose of the present study was to introduce some organised structure and opportunity for comparison into the usually challenging period which characterises the construction of a cancer registry. The reported results suggest the following conclusions:

—90.3% of hospitalised patients with incident cases of cancer in the study area could be identified from HDs and HDCs.
—Histological data, including both malignant and benign tumours, can rightly represent the main source in cancer registration.
—All cases with an HD also reporting a compatible HDC could be reasonably defined on the basis of computerised data.
—A thorough examination of the clinical records is recommended for cases only reported by HDCs, cases with severe disagreement between HD and HDC, cases with histology performed on metastatic tissues.
—Special and additional efforts should be directed towards identification of specific sites (namely gastrointestinal cancers, brain tumours and leukaemias).

This study can provide helpful information to cancer registries intending to evaluate the quality of routinely collected data, particularly computerised data, in planning new registration activities. Some of the methods described could also have a role in the periodical evaluation of well established cancer registries.

1. McLennan R, Muir C, Steinitz R, Winkler A. *Cancer Registration and its Techniques*. IARC Scientific Publications no. 21. Lyon, International Agency for Research on Cancer, 1978.
2. Goldberg J, Gelfand HM, Levy PS. Registry evaluation methods: a review and a case study. *Epidemiol Rev* 1980, **2**, 210–220.
3. Schouten LJ, Hoppener P, van den Brandt PA, Knottnerus JA, Jager JJ. Completeness of cancer registration in Limburg, the Netherlands. *Int. J Epidemiol* 1993, **22**, 369–376.
4. *Systematized Nomenclature of Pathology*. Chicago, IL, College of American Pathologists, 1965.
5. WHO. *International Classification of Diseases for Oncology*, first edition. Geneva, World Health Organization, 1976.
6. WHO. *International Classification of Diseases*, IX revision. Geneva, World Health Organization, 1975.
7. Muir C, Waterhouse J, Mack T, Powell J, Whelan S. *Cancer Incidence in Five Continents*, vol. V. IARC Scientific Publications no. 88, Lyon, International Agency for Research on Cancer, 1987.
8. Zanetti R, Crosignani P. *Cancer in Italy. Incidence Data from Cancer Registries* 1983–1987. Torino 1992.
9. Kleinbaum DG, Kupper LL, Morgenstern H. *Epidemiologic Research. Principles and Quantitative Methods*. New York, Van Nostrand Reinhold Company, 1982.
10. Gardner MJ, Gardner SB, Winter PD. Confidence interval analysis (CIA) microcomputer program. Statistics with confidence. *Br Med J* 1989.
11. Lukes RJ, Collins RD. Immunological characterization of human malignant lymphomas. *Cancer* 1974, **34**, 1488–1503.
12. National Cancer Institute sponsored study of classifications of non-Hodgkin's lymphomas: summary and description of a Working Formulation for clinical usage. *Cancer* 1982, **49**, 2112–2135.
13. Stansfeld AG, Diebold J, Noel H, *et al.* Updated Kiel classification for lymphomas. *Lancet* 1988, **1**, 292–293.
14. Coleman MP, Démaret E. Cancer registration in the European Community. *Int J Cancer* 1988, **42**, 339–345.

Acknowledgements—This study was partly supported by the Associazione Italiana per la Ricerca sul Cancro, Milano. We wish to thank Marisa Coato for her technical support and secretarial assistance, F. Cambiaso, L. Scolfaro, A. Serpelloni, L. Stevanella, M. Zamboni and the staff of the Data Analysis Units of the Verona province, who provided computerised data of clinical discharges. We are also indebted to Paolo Crosignani and Clotilde Vigano' (who kindly provided data from Lumbardy Cancer Registry) and Riccardo Capocaccia and Arduino Verdecchia, National Institute of Health (who made possible consultation of ISTAT data).

A special appreciation is expressed to Paolo Ricci (Local Health Unit 47, Mantova, Italy) for his valuable contribution in planning and discussing this project. We are particularly grateful to Franco Berrino and Andrea Micheli from the Lumbardy Cancer Registry, Benedetto Terracini from the Unit of Cancer Epidemiology, Department of Biomedical Sciences and Human Oncology (Turin, Italy) and Roberto Zanetti from the Piedmont Cancer Registry for their valuable suggestions and friendly encouragement.