Accuracy of Cancer Death Certificates and Its Effect on Cancer Mortality Statistics

CONSTANCE PERCY, MS, EDWARD STANEK, III, MS, AND LYNN GLOECKLER, MS

Abstract: A study to determine the accuracy of cancer mortality data was done using cancer deaths occurring during 1970 and 1971 in eight of the nine areas included in the Third National Cancer Survey (TNCS). Death certificates with an underlying cause of death of cancer were compared to the hospital diagnosis for 48,826 resident cases of single primary cancers.

The underlying cause of death as coded on the death certificate was found to be accurate for about 65 per cent of the cancer deaths in this study. Misclassification problems occurred for colorectal cancer, the second leading cause of death from cancer. Colon cancer was overreported and rectal cancer was underreported on death certificates. Other misclassification problems were found for cancers of the uterus, brain, and buccal cavity including most of its sub-sites. Physicians tended to report a non-specific site of cancer on the death certificate rather than the specific site identified by the hospital diagnosis. (*Am J Public Health* 1981; 71:242-250.)

Cancer mortality statistics have been part of many epidemiological investigations including a variety of etiological studies of cancer. Deaths from cancer have been used in geographic studies, studies of time trends, correlation studies, and therapy evaluation. They have also been used to identify cases for retrospective evaluation of possible etiological factors. Because of the varied uses of mortality data, it is important that they be reliable and accurate.

In the past, a number of authors have pointed out the inaccuracies of cancer death certificates[1-3] by comparing the specified underlying cause of death to autopsy diagnosis and to more specific hospital and pathologic information. Most of these studies have been of limited scope and dealt only with a small series of cases. To date, the only large-scale studies on the accuracy of cancer death certification are Dorn and Horn's 1941 study[4] based on the First National Cancer Survey; Dorn and Cutler's 1958 study[5] based on the Second National Cancer Survey; the Pan American Health Association study in 1967[6]; and studies by the Atomic Bomb Casualty Commission[7, 8] in Japan. No national assessment of the accuracy of cancer mortality data in the United States has been undertaken in the last 20 years. The study to be reported compares the underlying cause of death from cancer with the hospital cancer diagnosis of persons in specific areas of the United States. In order to be included in the analyses to be presented, a patient must have had a medical record which indicated that a diagnosis of one and only one cancer had been made and must have died with cancer coded as the underlying cause of death on the death certificate. Furthermore, the patient must have been a resident of the areas involved in the study.

From the Biometry Branch, DCCP, NCI, NIH. Address reprint requests to Constance Percy, Biometry Branch, National Cancer Institute, National Institutes of Health, Bethesda, MD 20205. This paper, submitted to the Journal October 26, 1979, was revised and accepted for publication October 28, 1980.

Editor's Note: See related editorial p. 231, this issue.

Materials and Methods

The data on deaths used in this study were drawn from the Third National Cancer Survey (TNCS), which was conducted in two states and seven metropolitan areas during the years 1969-1971. The survey reviewed hospital records, including autopsy and surgical pathology reports, for all active cancers, both incident and prevalent, in all area hospitals. Over 90 per cent of the TNCS cases were microscopically proven. This data set differs from that reported in Monograph 41[9] which dealt mainly with incident cases. Prevalent cases were included in this study in addition to the subset of the incident cases reported in this monograph.*

Death certificates with any mention of cancer were routinely collected from vital statistics sections of state health departments during the study period. However, not all copies of certificates received by TNCS included an International Classification of Diseases Adapted (ICDA-8)[10] code for the underlying cause of death. Sometimes copies of the certificates were requested by TNCS staff before the coding was done in the state health department or copies were sent without the underlying cause code appearing on the copies of the certificate. It would have been possible for the authors to code the underlying cause of death directly from the death certificate when the ICDA-8 code was missing. However, the coding rules for ICD-8[11] are complicated and codes are usually assigned by a trained nostologist. To prevent possible bias, no coding of the death certificates was attempted by the authors. As a result, deaths whose certificates lacked the ICDA-8 codes for cause of death had to be excluded from the analysis.

*Prevalent cases were defined in the TNCS study as having active cancer during 1969-1971 but having been first diagnosed previous to 1969.

In 1970, TNCS staff began abstracting from the death certificates the underlying cause of death coded according to the ICDA-8.[10] Death certificates must be obtained from the state health departments because the National Center for Health Statistics (NCHS) receives certificates without any personal identification. NCHS recodes all death certificates sent from the states but both the state and NCHS use the same classification and coding schemes, so the results of this study should be applicable to both U.S. mortality data (NCHS) and state mortality data.

If a hospital abstract was not on file for a death certificate with an underlying cause of cancer, an abstract of the medical record was obtained. Only 2.1 per cent of the cases in TNCS were "death certificate only" cases, i.e., the death certificate was the only evidence that the person had cancer. These cases were not included in this study since there was no corresponding hospital record. When an autopsy protocol indicated cancer, the death certificate was obtained and the autopsy diagnosis was considered as the hospital diagnosis.

Hospital diagnoses of cancer were abstracted and coded for site and histologic type according to the Manual of Tumor Nomenclature and Coding (MOTNAC).[12] To facilitate comparison, a computer program was written converting the MOTNAC codes to ICDA-8 categories.[13] Since this conversion consists of converting a detailed code involving both site and histologic type to the less specific categories of ICDA-8 which is based mainly on site, it is highly reliable for the site involved. In this study, malignant neoplasms are defined as ICDA-8 categories 140-207.

For a death to be included in this study, the copy of the death certificate received at the National Cancer Institute (NCI) must have an ICDA-8 cancer category coded for the underlying cause of death and it must have a hospital diagnosis of cancer. When a person had two or more cancers, it was difficult to determine which cancer caused the death. Therefore, patients with multiple primaries were excluded.

There were 48,826 deaths among residents during 1970 and 1971 in the TNCS areas for persons with only a single primary cancer and an underlying cause of death category of cancer coded on the death certificate. Excluded from these 48,826 deaths were those from the Minneapolis-St. Paul area which did not start recording the underlying cause of death categories until 1971. Also excluded were less than 100 deaths from malignant carcinoids (category 258) that were not considered to be malignant neoplasms by ICDA-8. Of these 48,826 patients, 60 per cent were incident (diagnosed in 1970 or 1971) and 40 per cent prevalent (diagnosed before 1970) cases.

To assess the representativeness of the cancer deaths in this study, a comparison of these 48,826 cancer deaths was made to those deaths reported by the NCHS for the survey areas. The 48,826 cancer death certificates from TNCS comprised 80.6 per cent of those reported as cancer deaths in the TNCS areas by NCHS in 1970 and 1971. In 1970, 70 per cent of the NCHS deaths were in our study and in 1971, 90 per cent were included. The discrepancy results from excluding patients with multiple primaries, death certificate only cases, and certificates received without an ICDA-8 category for the underlying cause of death. Overall, the most common cause of exclusion was the lack of an ICDA-8 cause of death category coded on the certificate. After the first year of the study, a special effort was made to get this item from the health departments (hence the difference between the 1970 and 1971 percentages). Cross classifications of hospital diagnoses with underlying cause of death by site were examined for the individual years, 1970 and 1971, and found to be very similar. Therefore, it was decided to combine the data for these years. The survey deaths were cross-classified by sex, race (White, non-White), and age (0-49, 50-69, 70+), and subsequently compared to NCHS deaths (Table 1). Although there were slight differences in percentages, none of the comparisons suggested serious biases in the data.

The accuracy of the death certificate code was assessed by comparing the primary cancer site reported on the hospital diagnosis with the cancer site coded as the underlying cause of death on the death certificate (Table 2). This was done at a 3-digit level of the ICDA-8 classification[10] for most sites. Most mortality data are reported by similar groupings. The basis of the analysis is a cross-tabulation of the hospital diagnosis (vertical axis) with the underlying cause of death (horizontal axis). The diagonal of the cross-tabulation represents cases whose hospital diagnosis and underlying cause of

TABLE 1—Number of Cancer Deaths Reported to the National Center for Health Statistics (NCHS) for TNCS Areas* and Percentage of Such Deaths Included in this Study by Age, Sex, and Race (1970 + 1971)

Race and Sex	No. of NCHS Cancer Deaths by Age				% of NCHS Cancer Deaths Included in this Study by Age			
	0-49	50-69	70+	All Ages	0-49	50-69	70+	All Ages
White	6,552	23,983	23,022	53,557	77.6	79.5	81.8	80.3
Males	3,142	13,400	12,462	29,004	79.9	80.0	80.1	80.0
Females	3,410	10,583	10,560	24,553	75.5	78.7	84.0	80.6
Non-white	1,269	3,696	2,069	7,034	80.5	80.4	89.0	82.9
Males	614	2,202	1,241	4,057	80.8	81.6	87.6	83.3
Females	655	1,494	828	2,977	80.2	78.6	91.1	82.4
Total	7,821	27,679	25,091	60,591	78.1	79.6	82.5	80.6

*Minneapolis-St. Paul area excluded.

TABLE 2—Comparison of Hospital Diagnoses with Underlying Cause of Death by Cancer Sites (ICDA-8: 140-207) for Cancer Deaths from the Third National Cancer Survey, 1970–71

HOSPITAL DIAGNOSIS		UNDERLYING CAUSE OF DEATH FROM DEATH CERTIFICATE																													
PRIMARY SITE	ICDA-8 CATEGORY	140-207*	140-149	150	151	153	154	155.0, 197.8	155.1, 156	157	152, 158-9	161	162	160, 163	172	174	180	182	183	181, 184	185	186, 187	188	189	191, 192	193, 194	201	200, 202	203	204-207	170-1, 173, 190	195-199†
ALL SITES		48826	1187	997	2321	5131	1367	687	709	2531	290	433	10178	173	486	4583	869	768	1499	138	2579	171	1179	930	1171	205	536	1470	688	2140	647	2763
BUC. CAV.	140-149	1397	1098	20	2	2	0	1	0	0	1	100	22	9	1	1	1	0	0	1	1	0	1	0	3	0	1	3	0	0	46	83
ESOPHAGUS	150	921	5	858	24	0	0	0	1	2	0	3	13	0	0	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	11
STOMACH	151	2365	0	85	2109	26	2	8	1	17	12	0	10	0	0	5	1	0	5	0	3	1	0	0	1	1	1	4	0	0	5	68
COLON	153	4546	0	3	19	4062	138	25	4	15	22	0	25	1	0	11	6	2	14	0	3	0	6	3	3	3	1	1	0	0	2	177
RECTUM	154	2098	0	0	5	791	1180	8	3	2	4	0	9	0	3	1	4	3	1	0	5	0	1	1	0	0	0	0	0	0	8	69
LIVER	155.0, 197.8	462	0	0	4	6	0	377	10	8	1	0	8	0	0	0	0	2	1	0	1	0	0	0	0	0	0	1	0	0	0	43
GALLBL.	155.1, 156	801	0	0		6	0	67	627	55	4	0	3	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	1	28
PANCREAS	157	2496	0	1	20	20	0	20	21	2252	11	1	24	0	0	2	0	0	5	0	3	0	1	1	1	0	0	2	0	0	1	110
OTH.DIG.	152, 158-9	355	0	1	22	40	4	6	7	17	146	0	4	0	0	2	0	1	3	0	0	0	0	2	3	2	0	5	0	0	38	52
LARYNX	161	395	40	4	0	1	0	0	0	0	0	317	16	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	14
LUNG +	162	10059	6	13	11	11	1	16	1	17	3	6	9560	38	3	19	5	1	3	0	6	1	3	9	21	10	0	12	0	4	13	266
OTH.RESP.	160, 163	169	10	1	1	0	0	0	0	0	2	0	15	89	4	0	0	1	0	0	0	2	0	0	4	2	0	1	0	0	13	24
MELANOMA	172	502	1	0	1	1	2	1	0	0	0	0	11	0	440	0	0	0	0	0	0	0	0	1	13	0	0	3	4	0	7	17
BREAST	174	4734	1	1	4	9	1	8	2	5	1	0	42	6	0	4498	1	3	8	0	2	0	0	0	8	4	0	2	0	0	13	115
CERVIX	180	995	0	0	2	7	1	0	0	1	1	0	10	2	0	3	786	111	16	2	0	0	5	1	1	0	0	0	0	0	4	42
CORPUS	182	674	0	0	2	5	0	1	0	0	2	0	8	0	0	3	25	549	34	9	0	0	0	0	0	0	0	0	1	0	5	30
OVARY +	183	1497	0	0	4	26	1	2	2	5	6	0	6	0	0	4	4	21	1322	2	0	0	1	0	0	1	0	2	0	0	3	85
OTH.FEM.	181, 184	232	0	0	1	3	3	0	0	1	0	0	2	0	0	0	17	38	10	117	0	0	3	1	1	0	0	0	0	0	18	17
PROSTATE	185	2621	1	1	3	11	2	1	1	8	2	0	21	1	0	0	0	0	1	0	2483	1	27	1	1	0	0	1	1	0	6	47
OTH.MALE	186, 187	186	0	0	0	0	0	0	1	0	1	0	7	0	0	0	0	0	0	1	1	156	0	3	2	0	0	1	0	0	1	12
BLADDER	188	1211	0	0	0	4	2	1	1	1	1	0	9	1	0	1	5	1	2	0	34	3	1103	18	1	0	0	0	0	1	2	20
KIDNEY	189	984	1	0	1	4	0	2	0	0	3	0	25	0	0	2	0	1	1	1	5	1	17	865	4	0	0	3	0	0	6	42
BRAIN +	191, 192	1074	0	0	1	0	0	0	0	0	0	0	4	0	1	1	0	0	0	0	0	0	0	1	1044	0	0	2	0	0	9	11
THYROID +	193, 194	206	0	2	0	0	0	0	1	2	1	0	6	0	1	1	0	0	0	0	0	1	0	2	5	169	0	1	0	0	0	14
HODGKINS	201	572	1	0	1	0	0	0	0	0	1	1	2	1	1	0	0	0	0	0	0	0	0	0	0	1	496	61	0	2	2	2
N-H.LYMPH.	200, 202	1562	5	0	28	10	4	2	1	9	13	0	11	3	0	4	1	1	0	0	6	2	1	3	5	3	34	1300	3	57	6	50
MULT.MYELOMA	203	699	0	0	0	0	0	0	0	1	1	0	2	0	2	0	0	0	0	0	0	0	0	0	2	0	1	3	675	6	2	4
LEUKEMIA	204-207	2152	0	0	0	1	0	2	0	2	0	0	1	1	0	2	0	0	1	0	0	0	1	0	0	0	0	35	1	2069	28	8
OTHER SPEC. SITES		541	3	0	2	6	17	2	0	1	4	0	15	3	26	6	2	12	7	2	0	1	0	0	8	1	0	9	0	0	343	71
UNKNOWN	195-199†	2320	15	7	47	79	9	137	25	110	47	5	287	18	3	15	11	21	64	3	26	1	9	18	39	6	2	18	3	1	63	1231

*Excludes 197.8

†See ICDA-8 (ref. 10) for complete description of categories.

death were identical. This agreement was measured by two different rates, detection and confirmation:

- The *detection rate* for a specific site was defined as the number of cases diagnosed as cancer of that site in the hospital and having cancer of the same site on the death certificate divided by the total number of persons diagnosed with that specific site of cancer in the hospital and dying of cancer. It is, therefore, the proportion of hospital diagnoses with cancer of a certain site in which the cause of death reflects the same hospital diagnosis.
- The *confirmation rate* is the same numerator divided by the number of persons who died with this particular site as the underlying cause of death and had had previous diagnosis of cancer. It is therefore the proportion of cancer deaths in which the specified underlying cause is confirmed by the hospital diagnosis.

Selected sites of cancer were classified into four groups according to the relative and absolute values of their detection and confirmation rates as follows:

Group 1: *Both high detection and confirmation rates at about the same level (over 80 per cent).* This means that there was good agreement between the primary site diagnosed in the hospital and that recorded on the death certificate. Mortality rates for these sites should be fairly accurate. The majority of the deaths were found in this group.

Group 2: *Both low detection and confirmation rates at about the same level (under 80 per cent).* In this group, there was considerable disagreement between hospital diagnoses and the corresponding underlying causes reported on the death certificates. Nevertheless, since both the detection and confirmation rates were equally low, the number diagnosed in the hospital and the number of deaths reported on the death certificates were about the same (columns 1 and 2 of Table 3). Such rates were observed usually for sites of low frequency. For cancer sites in this group the proportionate mortality would remain the same.

Group 3: *Detection rate higher than confirmation rate.* The sites in this group are characterized by a greater number of deaths reported on death certificates for the specific site than reported by the hospital diagnosis; this means that these sites will be overreported in the subsequent mortality statistics. This occurs for a site such as bone where metastasis from other primary sites frequently occurs. Unless the certifier specifies bone as a metastatic site, it will be coded as the primary site.

Group 4: *Confirmation rate higher than detection rate.* In this group more cases of a site were diagnosed in the hospital than were actually reported on the death certificates. This resulted in an underreporting on death certificates.

Results

Frequencies and detection and confirmation rates are presented in the Tables for the 48,826 cancer deaths. The detection and confirmation rates are given with their standard errors for each site to aid researchers in evaluating the accuracy of cancer mortality statistics for these sites.

As mentioned above, Table 2 cross-classifies the hospital diagnoses of the 48,826 cancer deaths with the underlying cause of death on the certificate. Overall, when the 30-site groupings shown in this Table were used, 86.7 per cent had the same site reported on the hospital diagnosis as that indicated as the underlying cause of death on the certificate. However, if the number of groups (3-digit categories) is increased to 49 as in Table 3, the overall agreement decreases to 82.7 per cent. Naturally, if 4-digit categories are used (Tables 5-7), agreement decreases even more.

Table 3 presents the data of Table 2 in a different format and includes the detection and confirmation rates and their standard errors. The figures and rates are given for each 3-digit primary site and a few combinations of sites.

Table 4 identifies the principal sites which fell into the four groups described earlier. In this study 65 per cent of the total cancer deaths belong in Group 1. A similar percentage (64) of United States cancer deaths in 1970 were attributed to cancers of the sites included in Group 1. Extremely high detection and confirmation rates of over 93 per cent were found for cancers of the lung and bronchus, breast, prostate, and multiple myeloma.

Connective tissue tumors are a good example of Group 2. Although there was much misclassification between the hospital diagnosis and the underlying cause of death, about the same number (259 cases) were reported in the hospital as were found for the underlying cause on the death certificate (252 cases). However, as seen in Table 3, column 3, only 142 connective tissue cancers were diagnosed and confirmed on death certificates, a detection rate of 54.8 per cent and a confirmation rate of 56.3 per cent. Very few cancer deaths are in Group 2.

The category "malignant neoplasm of the bone" is a good example of Group 3 (detection rate higher than confirmation rate). Table 3 shows that 160 bone cancers were diagnosed in the hospital but many more (252) were stated as the underlying cause of death. Since only 125 of these were confirmed in the hospital, the confirmation rate was only 49.6 per cent but the detection rate was 78.1 per cent. An examination of those deaths coded to primary cancer of bone on the death certificate showed diagnoses in the hospital of primary cancers of many other sites. No doubt bone was actually a secondary site of these cases and thus they were misclassified. Therefore mortality figures for bone cancer are considerably overreported in vital statistics data. Colon cancer, one of the principal sites of cancer, falls into Group 3.

Group 4 sites have higher confirmation rates than detection rates. For example, more buccal cavity cancers were diagnosed in the hospital (1,397 cases), than were reported on death certificates (1,187 cases). Since 1,098 deaths were confirmed by the hospital diagnosis, the confirmation rate (92.5 per cent) was much higher than the detection rate (78.6 per cent). Malignancies of the rectum, a frequent site of cancer, fall into this group as well.

For many sites, a specific diagnosis was made in the hospital but only a non-specific site was stated on the death certificate. Table 5 shows a detailed comparison of cancer of the cervix uteri, corpus uteri and uterus, NOS (not otherwise specified). Many more deaths were reported for cancer

TABLE 3—Number of Cases Diagnosed in Hospital and Number with Underlying Cause of Death by Site and Number with Same Site on both, Cancer Deaths from Third National Cancer Survey, 1970–1971 Detection and Confirmation Rates

ICDA-8 Category	Primary Site	No. diagnosed in hospital with this site	No. with this site as the cause of death	No. with same site on both	Detection Rate %	Detection Rate Standard error	Confirmation Rate %	Confirmation Rate Standard error
140-207	All Sites	48,826	48,826	40,379	82.7	0.17	82.7	0.17
140-149	Buccal cavity	1,397	1,187	1,098	78.6	1.10	92.5	0.76
140	Lip	32	14	14	43.8	8.77	100.0	
141	Tongue	324	267	219	67.6	2.60	82.0	2.35
142	Salivary gland	90	67	56	62.2	5.11	83.6	4.53
143	Gum	54	23	14	25.9	5.96	60.9	10.18
144	Floor of mouth	135	79	51	37.8	4.17	64.6	5.38
145	Mouth, NOS	176	177	83	47.2	3.76	46.9	3.75
146	Oropharynx	220	149	115	52.3	3.37	77.2	3.44
147	Nasopharynx	123	104	82	66.7	4.25	78.8	4.01
148	Hypopharynx	174	111	74	42.5	3.75	66.7	4.47
149	Pharynx, NOS	69	196	45	65.2	5.73	23.0	3.00
150	Esophagus	921	997	858	93.2	0.83	86.1	1.10
151	Stomach	2,365	2,321	2,109	89.2	0.64	90.9	0.60
152	Small intestine	113	109	75	66.4	4.44	68.8	4.44
153-154	Colon & Rectum	6,644	6,498	6,171	92.9	0.32	95.0	0.27
153	Colon	4,546	5,131	4,062	89.4	0.46	79.2	0.57
154	Rectum	2,098	1,367	1,180	56.2	1.08	86.3	0.93
155	Liver & Intrahep. bile duct	536	347	266	49.6	2.16	76.7	2.27
156	Gallbladder	727	674	583	80.2	1.48	86.5	1.32
157	Pancreas	2,496	2,531	2,252	90.2	0.59	89.0	0.62
158	Retroperitoneum	108	102	45	41.7	4.74	44.1	4.92
159	Peritoneum	134	79	20	14.9	3.08	25.3	4.89
160	Nose, ear & sinuses	93	71	53	57.0	5.13	74.6	5.16
161	Larynx	395	433	317	80.3	2.00	73.2	2.13
162	Trachea, lung & bronchus	10,059	10,178	9,560	95.0	0.22	93.9	0.24
163	Pleura, med. & other resp.	76	102	36	47.4	5.73	35.3	4.73
170	Bone	160	252	125	78.1	3.27	49.6	3.15
171	Connective tissue	259	252	142	54.8	3.09	56.3	3.12
172	Melanoma of skin	502	486	440	87.6	1.47	90.5	1.33
173	Other skin	51	67	18	35.3	6.69	26.9	5.42
174	Breast	4,734	4,583	4,498	95.0	0.32	98.1	0.20
180	Cervix	995	869	786	79.0	1.29	90.4	1.00
182	Corpus & uterus, NOS	674	768	549	81.5	1.50	71.5	1.63
183	Ovary, f. tube, etc.	1,497	1,499	1,322	88.3	0.83	88.2	0.83
181+184	Other female genital	232	138	117	50.4	3.30	84.8	3.10
185	Prostate	2,621	2,579	2,483	94.7	0.44	96.3	0.73
186	Testis	136	125	113	83.1	3.21	90.4	2.64
187	Other male genital	50	46	41	82.0	5.43	89.1	4.59
188	Bladder	1,211	1,179	1,103	91.1	0.82	93.6	0.72
189	Kidney	984	930	865	87.9	1.04	93.0	0.84
190	Eye	71	43	35	49.3	5.93	81.4	5.93
191-192	Brain & other nerv.	1,074	1,171	1,044	97.2	0.50	89.2	0.90
191	Brain	965	919	803	83.2	1.20	87.4	1.10
192	Other nervous system	109	252	79	72.5	4.28	31.3	2.92
193-194	Thyroid & other endo.	206	205	169	82.0	2.70	82.4	2.70
193	Thyroid	149	142	130	87.2	2.73	91.5	2.33
194	Other endocrine	57	63	39	68.4	6.16	61.9	6.12
195-199	Ill defined & unknown	2,320	2,763	1,231	53.1	1.00	44.6	0.90
200+202	Non-Hodgkin's lymphoma	1,562	1,470	1,300	83.2	0.90	88.4	0.80
201	Hodgkin's disease	572	536	496	86.7	1.42	92.5	1.14
203	Multiple myeloma	699	688	675	96.6	0.69	98.1	0.52
204-207	Leukemias	2,152	2,140	2,069	96.1	0.40	96.7	0.40
204	Lymphocytic	743	688	594	79.9	1.47	86.3	1.31
205	Myeloid	1,107	914	843	76.2	1.28	92.2	0.89
206	Monocytic	98	104	56	57.1	5.00	53.8	4.89
207	Other & unspecified	204	434	149	73.0	3.11	34.3	2.28

of the uterus, NOS (category 182.9) on the death certificates than were so diagnosed in the hospital. Of the 470 cases having cancer of the uterus, NOS on the death certificate, 103 were diagnosed in the hospital as cervical cancer and 184 as corpus cancer. In contrast, only 89 were diagnosed cancer of the uterus, NOS in the hospital, a confirmation rate of only 18.9 per cent.

Table 6 shows a comparison of the 4-digit ICDA-8 categories for cancer of the colon and rectum. The non-specific site of cancer of the colon, NOS (category 153.8) was found

TABLE 4—Groupings of Certain Cancer Sites According to the Absolute and Relative Values of the Detection and Confirmation Rates* (Cancer Deaths from Third National Cancer Survey, 1970–1971)

Group 1	Group 2	Group 3	Group 4
Both rates high	Both rates low	Detection rate higher than Confirmation Rate (overreporting on death certificates)	Confirmation Rate higher than Detection Rate (underreporting on death certificates)
Stomach (151)	Mouth, NOS (145)	Colon (153)	Buccal cavity (140–149)
Pancreas (157)	Small intestine (152)	Larynx (161)	Rectum (154)
Bronchus and Lung (162)	Connective tissue (171)	Bone (170)	Cervix (180)
Melanoma of skin (172)		Uterus, NOS (182.9)	Corpus (182.0)
Breast (174)		Pharynx, NOS (149)	Eye (190)
Ovary (183.0)		Ill defined and	Myeloid leukemia (205)
Prostate (185)		Unknown sites (195–199)	Transverse colon (153.1)
Bladder (188)			Sigmoid colon (153.3)
Thyroid (193)			
Multiple myeloma (203)			

*See Table 3 for rates.
NOTE: ICDA-8 (ref. 10) categories given in parentheses.

as the underlying cause of death on 3,299 certificates, but only 554 cases were classified in this category from the hospital diagnosis, a detection rate of 76.0 per cent but a confirmation rate of only 12.8 per cent (Group 3). Of these 3,299 colon cancers (category 153.8), 468 were diagnosed in the hospital as cancer of the transverse colon, 749 as sigmoid colon, 266 as recto-sigmoid, etc. As can be seen from this Table, a specific site was diagnosed in the hospital, but the physician did not record this specific site as the underlying cause on the death certificate. When all the sites of the colon (category 153) are combined, the detection rate is 89 per cent and the confirmation rate is 79 per cent, putting this site in Group 3. Cancer of the rectum, on the other hand, falls into Group 4 because the confirmation rate (86 per cent) was much higher than the detection rate (56 per cent). Thus cancer of the colon was overreported on death certificates and cancer of the rectum was underreported. If cancer of the colon and rectum are combined, the detection rate is 93 per cent and the confirmation rate is 95 per cent, placing the combination in Group 1.

ICDA-8 has several categories for coding cancers of the liver: 155.0—primary liver cancers, such as hepatocellular carcinoma; 197.8—malignant neoplasms of the liver, not specified primary or secondary; and 197.7—liver cancer specified as secondary. Table 7 gives the comparison and rates for the various liver categories. Of the 536 cases diagnosed as primary liver cancer in the hospital, only 347 cases had an underlying cause of death category of 155.0 (primary malignant neoplasm of liver) on the death certificate, a detection rate of 49.6 per cent. However, in addition there were 375 cases with a cause of death category of 197.8 (malignant neoplasm of liver, unspecified). By adding the deaths from category 155.0 and category 197.8 together, the detection rate rose to 79.1 per cent as shown in Table 7. If only the category 155.0 were used, there would be gross underreporting of malignant neoplasms of liver on death certificates. Therefore, a more accurate total of liver malignancy deaths was obtained by combining ICDA-8 categories 155.0 and 197.8 together.

Myeloid leukemia and lymphocytic leukemia were underreported on death certificates (Table 3). Again the diagnosis was specific in the hospital but the physician just wrote "leukemia" (category 207) on the death certificate. Over twice as many cases were in this unspecified category on the

TABLE 5—Compairson of Hospital Diagnosis with Underlying Cause of Death: Cancer Deaths from the Third National Cancer Survey, 1970–1971, Detection and Confirmation Rates for Malignant Neoplasms of Specific Sites of the Uterus

Hospital Diagnosis		Underlying Cause of Death			Total no. diagnosed in hosp. of this site*	Detection Rate		Confirmation Rate	
ICDA-8 Category	Malignant Neoplasm of	Cervix 180	Corpus 182.0	Uterus NOS† 182.9		Per Cent	Standard Error	Per Cent	Standard Error
180	Cervix Uteri	786	8	103	995	79.0	1.29	90.4	1.00
182.0	Corpus Uteri	23	265	184	547	48.4	2.14	89.5	1.78
182.9	Uterus, NOS†	2	10	89	127	70.1	4.06	18.9	1.81
Total no. with this underlying cause*		869	296	470					

*Figures for other sites are included in this total but not shown in the Table
†Not Otherwise Specified.

TABLE 6—Comparison of Hospital Diagnosis with Underlying Cause of Death: Cancer Deaths from the Third National Cancer Survey, 1970–1971, Detection and Confirmation Rates for Malignant Neoplasms of Colon and Rectum

ICDA-8 Category	Malignant Neoplasm of	Underlying Cause of Death: 153.0	153.1	153.2	153.3	153.8	153.9	154.0	154.1	154.2	Total no.* diagnosed in hospital of this site	Detection Rate: Per Cent	Detection Rate: Standard Error	Confirmation Rate: Per Cent	Confirmation Rate: Standard Error
153.0	Cecum, ascending colon & appendix	551	8	4	13	610	25	1	8	-	1,311	42.0	1.36	88.6	1.28
153.1	Transverse colon	18	144	16	8	468	22	2	2	-	723	19.9	1.49	87.8	2.56
153.2	Descending colon	4	2	51	23	260	5	2	8	-	394	12.9	1.69	38.3	3.75
153.3	Sigmoid colon	13	1	33	473	749	47	37	62	-	1,519	31.1	1.19	71.1	1.76
153.8	Colon, NOS†	13	3	8	13	421	19	1	15	-	554	76.0	1.81	12.8	0.58
153.9	Intestine, NOS†	-	-	-	2	10	25	-	-	-	45	55.6	7.41	10.1	1.91
154.0	Recto-sigmoid	2	-	11	80	266	18	86	127	-	627	13.7	1.37	51.2	3.86
154.1	Rectum	2	1	7	39	334	25	36	906	5	1,435	63.1	1.27	76.2	1.24
154.2	Anus, etc.	-	-	-	1	4	1	-	18	2				50.0	12.65
Total no. with this underlying cause*		622	164	133	665	3,299	248	168	1,189	10					

*Figures for other sites are included in total but not shown in the table.
†Not Otherwise Specified.

death certificate as were diagnosed in the hospital.

A final observation concerns deaths and diagnoses in ICDA-8 categories 195–199, cancers of ill-defined and unknown sites. The detection rate was 53.1 per cent and the confirmation rate was 44.6 per cent. As seen in the last row of Table 2, many of the cases diagnosed as "unknown" in the hospital were coded to a specific site on the death certificate. Large frequencies were found for common metastatic sites such as lung, liver, and bone. Conversely, when we look in the other direction (at the column) for this category in Table 2, we note that over one-half of the cases with unknown or ill-defined site on the death certificate had a hospital diagnosis of cancer of a specific site.

Discussion and Summary

Variability and biases in cancer mortality data can distort, limit, or inhibit the value of this important resource in epidemiologic studies. The data in this study have been used to pinpoint some of these problems.

The reasons the absolute numbers of cancer deaths in TNCS did not correspond to those reported by NCHS have been previously presented. To ensure that this discrepancy in numbers did not bias this study, a comparison was made of the site distribution of cancers (ICDA-8 categories 140–207) reported by NCHS[14] for the total U.S. in 1970–71 to the TNCS cancer deaths in 1970–71 used in this study. Table 8

TABLE 7—Comparison of Hospital Diagnosis with Underlying Cause of Death: Cancer Deaths from the Third National Cancer Survey, 1970–1971, Detection and Confirmation Rates for Malignant Neoplasms of the Liver

ICDA-8 Category	Hospital Diagnosis	Underlying Cause of Death: Malignant Neoplasm of: Primary Liver 155.0 + 197.8	Liver, NOS‡	Total Liver 155.0 + 197.8	Secondary Liver 197.7	Unknown Site 195–199 **	Total no. diagnosed in hospital of this site* 195–199	Detection Rates: Per Cent	Detection Rates: Standard Error	Per Cent	Standard Error	Confirmation Rate: Per Cent	Confirmation Rate: Standard Error
155.0	Liver, primary	229	148	377	14	29	462	49.6	2.33	81.6	1.80	72.7 39.5 52.2	2.51 2.52 1.86
Total no. with this underlying cause*		315	375	722	213								

*Figures for other sites are included in this total but not shown in the Table.
**Excluding sites 197.7 and 197.8.
‡Not Otherwise Specified.

TABLE 8—Comparison of Percentages of U.S. Cancer Deaths and TNCS Cancer Deaths by Death Certificate and Hospital Diagnosis for Selected Sites of Cancer, 1970-1971

Site				
ICDA Category	Term	% of total cancer deaths in U.S. (14)	% of TNCS Cancer deaths by cause of death on certificates	% of TNCS Cancer deaths by hospital diagnosis
140-149	Buccal cavity	2.3	2.4	2.9
150	Esophagus	1.8	2.0	1.9
151	Stomach	4.8	4.8	4.8
153.1	Transverse colon	0.4	0.3	1.5
153.2	Descending colon	0.3	0.3	0.8
153.3	Sigmoid colon	1.5	1.4	3.1
153.8	Colon, NOS	6.7	6.8	1.1
154	Rectum	3.1	2.8	4.3
160-163	Lung and other respiratory	21.4	22.1	21.8
174	Breast	9.0	9.3	9.4
180	Cervix	2.0	1.8	2.0
183	Ovary	3.0	3.1	3.1
185	Prostate	5.3	5.3	5.4

displays these data for selected sites comprising over 60 per cent of the cancer deaths, showing that there is very good agreement between the distribution of U.S. cancer deaths (column 1) and TNCS cancer deaths (column 2) by site.

What would happen to cancer mortality statistics if the death certificates were reviewed and the underlying causes changed to correspond to the site of cancer as reported in the hospital? The absolute value of the mortality rate would not change much for any particular site since the number of deaths is small in comparison to the total population of the areas. However, the percentage change is dependent on the site. Sites which fall into Groups 3 and 4 would have substantial percentage changes in the mortality rate. This would be important in any time trend analysis. Similarly, the proportionate cancer mortality ratio due to a particular site of cancer would show little change (Table 8, columns 2 and 3) but the percentage change based on the number of deaths of a particular site would be large for sites in Group 3 and 4. For example, if the hospital diagnosis was used instead of the cause recorded on the death certificate, the number of transverse colon deaths would increase from 164 to 723 (see Table 6), an increase of 340.9 per cent (723-164)/164. Similar calculations for descending colon, sigmoid colon, colon, NOS, rectum, and buccal cavity yield changes of 196.2 per cent, 128.4 per cent, − 81.9 per cent, 53.5 per cent, and 17.7 per cent, respectively. This illustrates the fact that mortality trends may show a per cent change due to an artifact caused by physicians changing their method of recording the diagnoses when they fill out the death certificate.

Death certificates are frequently used as a means of case ascertainment for retrospective studies. Cases identified for sites with low confirmation rates would be subject to possible bias, as many cases identified by the death certificates for a particular site would not have that site as the hospital diagnosis. For example, if bone cancer cases were selected from the death certificates, over one-half would not be bone cancers according to the hospital record. Another bias is that if the detection rate is low, cases pulled from the death certificates for a particular site will not represent all the cases diagnosed of that site. For example, over one-half the deaths due to malignant neoplasms of eye (category 190) would be missed if only death certificates were used because the certificate often says only "melanoma" with no mention of eye; these certificates are coded to category 172 (melanoma of skin) and not to eye.

There are a few additional factors besides the site of the neoplasm that may affect the accuracy of death certificates, such as the presence of an autopsy, evidence of microscopic confirmation, age at death, sex, race and geographic area. Although mechanisms exist for amending the original death certificates after autopsies are completed, this is rarely done. As a result, the performance of autopsies does not routinely result in improvements in death certificate accuracy. There was a slight variation in overall accuracy from area to area but no substantial differences were observed except for a few sites of low frequency. The overall agreement of the other factors mentioned, i.e, age, sex, and race, did not cause a difference of more than one or two per cent from the 86.7 per cent overall agreement in Table 2. The place of death may influence the accuracy of the recorded cause of death. In this study 73 per cent of the deaths occurred in the hospital, 12 per cent in nursing homes, and 15 per cent at home.

Certifiers are not always familiar with the indexing of ICDA-8 or the guidelines and instructions[15] for filling out death certificates and do not realize how the order of entry of the terms which they record ultimately determines the selection of the underlying cause of death by the coders.

Even though, in some cases, the NCHS had expanded the rules[16] of ICD-8, it was still difficult for coders to apply the rules in a uniform and consistent manner. The ICD-9,[17] which went into effect in January 1979, has further expanded the rules and improved the index. Word choice is critical for determining the correct code. For example, when malignant brain tumors such as astrocytoma or glioma were diagnosed in the hospital, the physician frequently signed out the death certificate as "brain tumor." This cause of death is coded to category 238.1, neoplasm of unspecified nature of the brain and is therefore not recorded as a cancer death. Differences

in understanding between certifiers and coders of the meaning of a term may introduce systematic biases to the cause of death assignments. The term "metastasic" is often difficult to interpret on death certificates. For instance "metastatic lung cancer" can mean metastasis from a primary lung cancer or metastasis to the lung.[18]

The combination of the detection and confirmation rates for each site should help determine the reliability of the frequency of each particular site. Among the ten leading sites of cancer deaths,[19] seven sites—lung, breast, prostate, pancreas, urinary bladder, ovary, and leukemia—fell into Group 1. Thus, with both a high detection and confirmation rate, the mortality rates for these sites can be considered reliable. Furthermore these sites represent 65 per cent of all cancer deaths in this study.

However, statistics on the remaining one-third of the cancer sites are not as accurate:

Cancer of the colon and cancer of the rectum, together the second leading cause of cancer deaths (18 per cent), are different entities with separate characteristics. Colon cancer is overreported on death certificates and rectal cancer is underreported. This crossover forces some statisticians to combine mortality figures of colon and rectal cancer in order to get a more accurate count of colorectal cancer. Unfortunately, this obliterates the individual characteristics of each disease.

Cancer of the uterus, the fifth leading cause of cancer deaths among females, has the same type of problem as colorectal cancer. Cancers of the corpus uteri and cervix uteri are not the same disease and have different etiologies. They should be identified specifically on the death certificate and not be called simply "uterus".

Buccal cavity cancer, the tenth leading cause of cancer deaths, and its components—cancers of lip, tongue, salivary gland, gum, etc.—are underreported on death certificates.

If physicians understand how their reporting affects the classification of the underlying cause of death, they are likely to be more accurate in completing the death certificate. State vital statistics departments can and must play a role in monitoring and encouraging accurate certification. Ensuring that autopsy diagnoses are reflected on death certificates would be a step in the right direction.

The quality of cancer death certificate data will continue to be an important issue for cancer researchers, not simply to acknowledge but also to address. When using cancer mortality data in studies of a specific cancer site, attention must be paid to the limitations as well as to the potentials of this data source.

REFERENCES

1. Steer A, Land CE, Moriyama IM, *et al*: Accuracy of diagnosis of cancer among autopsy cases: JNIH-ABCC Population for Hiroshima and Nagasaki. Gann 1976; 67:625-632.
2. Moriyama IM, Baum WS, Haenszel WM, *et al*: Inquiry into diagnostic evidence supporting medical certifications of death. Am J Public Health 1958; 48:1376-1387.
3. James G, Patton RE and Heslin AS: Accuracy of cause-of-death statements on death certificates. Pub Health Rep 1955; 70:39-51.
4. Dorn HF and Horn JL: The reliability of certificates of deaths from cancer. Am J of Hygiene 1941; 34:12-23.
5. Dorn HF and Cutler SJ: Morbidity from cancer in the United States. Comparison of death certificates and case reports. Pub Health Mono 1958; 56:117-124, 206-207.
6. Puffer RR and Griffith GW: Patterns of urban mortality. Pan American Health Organization, WHO, 1967, pg. 90.
7. Steer A, Moriyama IM and Shimizu K: The autopsy program and the life span study, Jan. 1951-Dec. 1970 ABCC-JNIH Pathology Studies, Hiroshima and Nagasaki. Report 3, Technical Report 16-73.
8. Yamamoto T, Moriyama IM, Asano M and Guralnick L: The autopsy program and the life span study, Jan. 1961-Dec. 1975. Radiation Effects Research Foundation, Report 4, Technical Report RERF TR 18-78.
9. Cutler SJ and Young JL Jr (eds): Third National Cancer Survey: Incidence Data, NCI Monograph 41, DHEW Publication No. (NIH) 75-787, 1975.
10. U.S. Dept. of Health, Education, and Welfare. International Classification of Diseases, Adapted for use in the United States, Eighth Revision, PHS Publication No. 1693, 1967.
11. World Health Organization: International Classification of Diseases, 1965 revision: Geneva: WHO, 1967.
12. Percy C, Berg JW and Thomas LB (eds): Manual of Tumor Nomenclature and Coding, 1968 Edition. New York: American Cancer Society, 1968.
13. American Cancer Society: Conversion Table No. 3, Conversion of Neoplasms by site and histology from Manual of Tumor Nomenclature and Coding, 1968 Edition to 8th Revision of the International Classification of Diseases Adapted. ACS No. 0803, 1973.
14. Vital Statistics of United States, II-Mortality, Part A. National Center for Health Statistics, US DHEW, 1970, 1971.
15. U.S. Dept. of Health, Education, and Welfare: Physicians' handbook on medical certification: death, birth, fetal death. Pub. No. (PHS) 78-1108.
16. Vital statistics: Instructions for classifying the underlying cause of death, 1976-77. National Center for Health Statistics, Part 2a, DHEW, 1975.
17. World Health Organization: International Classification of Diseases, 1975 revision: Geneva: WHO, 1977.
18. Percy C and Dolman A: Comparison of the coding of death certificates related to cancer in seven countries. Pub Health Rep 1978; 93:335-350.
19. American Cancer Society: Cancer Facts and Figures, New York, ACS, 1980.

ACKNOWLEDGMENTS

The authors wish to thank Valerie Van Holten for the extensive programming of the MOTNAC/ICDA-8 conversion. Without this program, this study would not have been possible. Also the authors thank William McKay of the National Cancer Institute for supplying the necessary figures for Table 1 that he adapted from U.S. mortality tapes supplied by NCHS. The authors appreciate the comments of Dr. Susan Devesa, and the technical assistance of Carol Kalb and other members of the Demographic Analysis Section of the Biometry Branch of the National Cancer Institute.

An earlier version of this paper was presented at the American Public Health Association Annual Meeting on November 7, 1979 in New York City.