Cancer of the Breast, Colon, Ovary, and Testis in the United States: Rates 1970–78 from a Hospital Reporting System

JANE B. PORTER, MS, ALEXANDER M. WALKER, MD, DRPH, AND HERSHEL JICK, MD

Abstract: We have explored the use of data from the Commission on Professional and Hospital Activities-Professional Activity Study (CPHA-PAS) for ascertaining information on cancer incidence, with regional breakdowns. Extirpative surgical procedures were linked with discharge diagnoses to provide estimates of numbers of incident cases. We calculated incidence rates for four cancers: breast, colon, ovary, and testis. CPHA-PAS inferred rates corresponded closely to those of other reporting systems for breast cancer in most age groups, and for colonic and testicular cancer in some age groups. Ovarian cancer rates were consistently underestimated. We conclude that a cancer incidence reporting scheme based on hospital discharge data can work for certain cancers, and be very inexpensive and efficient. It must, however, be used with care. (*Am J Public Health* 1984; 74:585–588.)

Introduction

The expense of effective cancer registration schemes has led to consideration of alternative methods for monitoring cancer incidence in large populations.[1] One technique, evolved for the study of endometrial cancer rates,[2,3] uses a national data base of hospital discharges, the research sample of the Commission on Professional and Hospital Activities-Professional Activity Study (CPHA-PAS). For endometrial cancer, the procedure involved identifying hospital discharges in which there was both a diagnosis of endometrial cancer and a report of hysterectomy having been performed during hospitalization. These were taken to be incident cases of endometrial cancer and served as the basis for calculation of cancer rates very similar to those obtained from more conventional sources. We have now attempted to use this procedure for four more cancers—those of the breast, colon, ovary, and testis.

Materials and Methods

Each year CPHA-PAS records information on a substantial fraction of hospital discharges in the United States. In cooperation with IMS America Ltd., CPHA-PAS draws from its records a sample corresponding to about 1 per cent of all discharges in the country, constructed so as to correspond to all United States hospital discharges in its distribution by census region (Northeast, South, North Central, West), hospital size (1 to 99, 100 to 199, 200 to 399, and 400 plus beds), and hospital ownership (governmental, nongovernmental). National data on hospital discharges used in constructing the probability sample were obtained from each year's "Annual Survey of Hospitals," a series of Master Facilities Inventory Tapes developed by the National Center for Health Statistics.[4]

Using computer files we were able to identify the numbers of patients 25 years and older in the following disease categories: 8,318 females with a diagnosis of cancer of the breast, according to the International Classification of Diseases, Hospital Adaptation (HICDA), eighth revision (codes 174.0–174.2) coupled with mastectomy (codes 86.0–86.9); 2,392 males and 3,068 females with the diagnosis of cancer of the colon (codes 153.0–153.4) with the performance of a local excision or destruction of a lesion or tissue of the large intestine, or colectomy (codes 45.2, 45.5, 45.6); 1,043 females with the diagnosis of cancer of the ovary (183.0) with excision or destruction of a lesion or tissue of the ovary, oophorectomy, or pelvic exenteration or evisceration (codes 65.1–5, 68.7); and 259 males with the diagnosis of cancer of the testis (code 186.0) with orchiectomy (codes 62.2–62.3). These were considered to be incident cases and were each cross-tabulated by age in five-year intervals, and by year of diagnosis (1970–1978). Projected numbers of new cases were obtained by dividing the number of cases in each age-year category of the sample by the sampling fraction for the corresponding year. These served as numerators for the cancer rates. Approximate age-specific populations at risk were obtained for 1970 and 1980 from the United States Census for those years; populations for the years in between were obtained by linear interpolation. Age- and year-specific rates were calculated by dividing projected numbers of cases by estimated populations at risk in each category. When specific rates were combined over several years, crude summary rates were calculated; when rates were combined across age categories, they were standardized to the 1970 United States male or female population. Coefficients of variation of rates (standard error divided by the point estimate) were calculated on the assumption that observed numerator counts had an underlying Poisson variation. The coefficient of variation, a measure of relative variation, expresses the standard deviation as a percentage of the point estimate. Age categories for which the coefficients of variation were greater than 20 per cent are not presented; the methods suggested here are not appropriate to these categories. CPHA-PAS estimates were compared with those calculated from data from the Third National Cancer Survey (TNCS) (1969–1971)[5] and from the Surveillance Epidemiology and End Results Program (SEER) (1973–1977).[6] Again, a standardization procedure was used when rates for different age categories were combined.

Results

Breast Cancer

Age-specific rates, as inferred from CPHA-PAS data, for breast cancer in the United States as a whole for three-year intervals 1970–1978 are presented in Table 1. Rates for women with ages 65–74 rose somewhat between the first and second intervals but remained essentially unchanged through the third; whereas, for women 75 years and over, rates rose and then fell in the third interval. Women younger

From the Boston Collaborative Drug Surveillance Program, Boston University Medical Center. Address reprint requests to Jane B. Porter, Boston Collaborative Drug Surveillance Program, BUMC, 400-1 Totten Pond Road, Waltham, MA 02154. This paper, submitted to the Journal March 15, 1983, was revised and accepted for publication December 28, 1983.

TABLE 1—United States Breast Cancer Rates*

Age (years)	Years 1970–72	1973–75	1976–78
25–34	16.0	14.8	14.2
35–44	76.1	67.6	69.7
45–54	158.0	167.5	155.4
55–64	188.1	196.7	200.0
65–74	201.0	242.5	239.0
75+	213.8	251.5	230.3

*Incident cases per 100,000 woman-years at risk. From CPHA-PAS (see text).

than 45 had relatively unchanging rates. Rates for women in the perimenopausal years, ages 45–54, rose and fell, but the degree of change for that age group is of the same order as the inherent variability in the data (see below) and may represent chance variation.

The question arises as to whether cases are adequately ascertained. Use of mastectomy as an indicator of incident cases underestimates the number, by the proportion of breast cancers not undergoing mastectomy. The estimate by contrast is inflated by the proportion of cases in which there is a repeat mastectomy, either by extension of an initial partial mastectomy, or by removal of the remaining breast. Figures for 1973–1976 from the Michigan Cancer Foundation Cancer Registry show that in women with breast cancer, mastectomy was performed in 94.5 per cent of women under age 50, 91.5 per cent of women age 50–69, and 85.5 per cent of women aged 70 or above, with little overall variation during the four years of observation, and no secular trend in the rates.[7] Cumulative rates for repeat mastectomy can be inferred from reports of cancers in the breast after an initial mastectomy. The cumulative incidence of such "asynchronous" cancers has been reported by Urban, *et al,* to be 5.8 per cent[8] and by Lewison to be 6.3 per cent.[9] Overall, the combined effects of non-operation for advanced cases and repeat surgery for recurrent cases will result in estimates with very little bias for women under the age of 70, and an underestimate of almost 10 per cent for older women. Recent trends toward the use of "lumpectomy" and local irradiation in the treatment of early breast cancers should not introduce any notable undercounting, because the surgical procedure would still be counted as a partial mastectomy.

The reliability of breast cancer rates estimated from CPHA-PAS is further dependent on the degree of chance variation in the data. Chance variation will be largest in population groups with smallest numbers of observed cases. Coefficients of variations were in the range from 3.7 per cent for ages 55–64 (1976–1978) to 11.4 per cent for ages 25–34 (1970–1972).

TABLE 2—Breast Cancer Rates* from the Third National Cancer Survey (TNCS) (1969–1971), Surveillance, Epidemiology, and End Results (SEER) (1973–1977), and as Estimated from CPHA-PAS Data

Age (years)	TNCS 1969–1971	CPHA-PAS 1970–1972	SEER 1973–1977	CPHA-PAS 1973–1977
25–34	15	16	17	15
35–44	79	76	83	72
45–54	162	158	184	164
55–64	204	188	239	198
65–74	240	201	291	241
75+	299	214	350	230

*Cases per 100,000 women per year.

Comparisons of breast cancer rates calculated from the CPHA-PAS data with the results of more formal cancer surveillance networks show a fairly high degree of concordance for young women (Table 2). For the oldest women, there is the anticipated underestimate due to lower mastectomy rates. In most age groups CPHA-PAS rates are lower than those of the TNCS or SEER. The breast cancer rates for all age groups increased between the TNCS and SEER. The CPHA-PAS data reflected the increase between the two time intervals for women 45 years and older.

Colon Cancer

Age-specific rates for colon cancer were estimable from CPHA-PAS data for men and women age 45 and older (Table 3). In persons less than 55 years of age, females had higher rates of colon cancer than males. The rates for both males and females showed an increase from the first to the third time interval in almost all age groups. In males the important increases occurred in those age 65 and over. In females, rates increased from the second to the third interval for all age groups. Again, the largest overall increase was in those 75 years and older.

Figures from SEER[10] show that in the period 1970 to 1973, 86 per cent of Whites and 80 per cent of Blacks were treated with some type of surgery. There were no large differences for the sexes for either race.* In another national series, in which more than 75 per cent of the cases were diagnosed before 1971, 83 per cent of cases underwent resection.[11] Therefore, there is some underestimation in using colectomy to indicate incident cases. For men, coefficients of variation ranged from 5.7 per cent for ages 65–74 (1976–1978) to 12.1 per cent for ages 45–54 (1973–1975). For

*This SEER series consisted of all newly diagnosed cancer patients in the state of Connecticut, approximately one-fifth of the patients in California, and all patients at two teaching hospitals—Charity Hospital in New Orleans, and the University of Iowa Hospital from 1950 to 1973.

TABLE 3—United States Colon Cancer Rates*

Age (years)	Males 1970–1972	1973–1975	1976–1978	Females 1970–1972	1973–1975	1976–1978
45–54	19.4	17.6	19.8	26.3	23.8	23.6
55–64	59.4	53.9	65.7	56.8	53.0	60.3
65–74	114.0	138.1	149.9	118.4	115.9	138.6
75+	191.3	202.9	234.5	168.7	178.3	212.6

*Incident cases per 100,000 person-years at risk. From CPHA-PAS (see text).

TABLE 4—Colon Cancer Rates* by Sex from the Third National Cancer Survey (TNCS) (1969–1971), Surveillance, Epidemiology and End Results (SEER) (1973–1977), and as Estimated from CPHA-PAS Data

Age (years)	TNCS 1969–1971	CPHA-PAS 1970–1972	SEER 1973–1977	CPHA-PAS 1973–1977
Males				
45–54	25.6	19.4	25.5	18.2
55–64	74.7	59.4	80.0	59.7
65–74	173.1	114.0	192.5	141.4
75+	300.0	191.3	362.2	213.9
Females				
45–54	28.2	26.3	27.0	23.3
55–64	68.9	56.8	72.5	57.1
65–74	144.8	118.4	157.8	124.1
75+	259.6	168.7	286.9	184.9

*Cases per 100,000 persons per year.

women, the range was 5.0 per cent for ages 75 and over (1976–1978) to 10.4 per cent for ages 45–54 (1976–1978).

In Table 4, colon cancer rates calculated from CPHA-PAS data are compared with rates calculated from other national cancer surveys.[5,6] The CPHA-PAS estimates are closest to the TNCS and SEER figures for middle-aged men and women. For older populations, colectomy ceases to be a reliable correlate of new disease. Nevertheless the CPHA-PAS data show an increase in rates for men and women 45 years and older with time as do those of the TNCS and SEER.

Ovarian Cancer

Age-specific rates for ovarian cancer were estimable for women aged 35 years and older (Table 5). In all age groups, the rates increased betwen the first and third intervals. The increases were most significant in those aged 65 and older. Rates for women in the perimenopausal years, ages 45–54, showed the least change over the whole time period.

Use of oophorectomy as an indicator of incident cases results in an underestimation of incident cases. Figures from SEER[10] show that in the period, 1970–1973, 71 per cent of White women and 64 per cent of Black women diagnosed with ovarian cancer were treated with surgery alone or in combination with another treatment. The recommended treatment for ovarian cancer is a maximal surgical effort—total hysterectomy, bilateral salpingo-oophorectomy, and omentectomy.[12–15] Even when one ovary appears to be normal, both are usually removed since the likelihood of malignancy spreading to the retained ovary is quite high (12 per cent to 18 per cent).[13] It is usually only in young women of childbearing age that an ovary is retained. Therefore, although

TABLE 5—United States Ovarian Cancer Rates*

Age (years)	Years 1970–1972	1973–1975	1976–1978
35–44	8.3	8.0	10.3
45–54	22.6	24.8	25.1
55–64	23.5	22.9	30.1
65–74	20.2	25.9	31.7
75+	17.5	18.0	26.3

*Incident cases per 100,000 woman-years at risk. From CPHA-PAS (see text).

TABLE 6—Ovarian Cancer Rates* from the Third National Cancer Survey (TNCS) (1969–1971), Surveillance, Epidemiology, and End Results (SEER) (1973–1977), and as Estimated from CPHA-PAS Data

Age (years)	TNCS 1969–1971	CPHA-PAS 1970–1972	SEER 1973–1977	CPHA-PAS 1973–1977
35–44	11.2	8.3	11.5	8.4
45–54	29.5	22.6	27.3	24.2
55–64	39.6	23.5	41.5	25.5
65–74	49.6	20.2	47.6	27.8
75+	46.9	17.5	49.2	20.8

*Cases per 100,000 women per year.

our rates underestimate incident cases, it seems unlikely that they would be inflated materially by repeated procedure.

Coefficients of variation ranged from 9.7 per cent for ages 55–64 (1976–1978) to 18.0 per cent for ages 75 and over (1970–1972, 1973–1975). Rates calculated from TNCS and SEER[5,6] are compared with rates calculated from CPHA-PAS data in Table 6. The CPHA-PAS rates underestimate those of the TNCS and SEER, moreso with increasing age. The rates have remained relatively stable from the TNCS to SEER, but those of CPHA-PAS increased slightly over the two time periods.

Testicular Cancer

Age-specific rates for testicular cancer in the United States could be estimated only for men aged 25–34 years (Table 7). The rates were relatively stable over the three time intervals.

Use of orchiectomy as an indicator of incident cases is probably a relatively good approximation to the number of testicular cancers, since figures from SEER indicate that from 1950 to 1973 about 95 per cent of White patients with testicular cancer underwent surgery.

The smallest age-year count of observed cases was 30 in 1970–1972, and the largest count was 46 in 1976–1978. These counts yield coefficients of variation of approximately 18.3 per cent and 14.7 per cent (Table 7).

Testicular cancer rates calculated from national cancer surveillance systems[5,6] are in close agreement with rates calculated from CPHA-PAS data (Table 7).

Discussion

CPHA-PAS is a resource which has several advantages. It is capable of providing up-to-date cancer incidence rates quickly—within about six months of hospitalization. In

TABLE 7—United States Testicular Cancer Rates, Ages 25–34*

	Source			
	CPHA-PAS		TNCS	SEER
Years	Rate	Coefficient of Variation	1969–1971	1973–1977
		%		
1970–1972	6.4	18.3	6.4	
1973–1975	8.5	15.2	—	8.6
1976–1978	8.7	14.7	—	

*Incident cases per 100,000 man-years at risk. From CPHA-PAS (see text).

addition, these data can be obtained very inexpensively. The primary problem encountered in using CPHA-PAS data is the identification of incident cases of cancers. For some cancers, this is more difficult than for others.

Usefulness of the case ascertainment procedure rests on the apparent completeness of ascertainment and on the degree to which the population in which the cases are ascertained is representative of the target population. As noted previously, the research file was chosen so as to correspond to discharges nationally and regionally according to hospital characteristics, eliminating biases deriving from relationships betwen utilization of CPHA-PAS and hospital size, location, or ownership. Factors resulting in biased ascertainment of cases would have to fulfill three criteria. First, they would have to be hospital characteristics. Secondly, they would have to affect a hospital's propensity to utilize CPHA-PAS services, independently of any association with hospital size, location, or ownership. Finally, they would have to affect the proportion of these cancer cases among all discharges, again independently of hospital size, location, and ownership. There are no readily apparent characteristics which meet all these criteria, and it is probable that the net effect of such biases as do occur is not large. In addition, for regional, age, or secular comparisons to be affected, one would have to posit biases which operate differentially across places, ages, or time.

CPHA-PAS data are of some value for the study of temporal fluctuations in the incidence of certain cancers. This resource is most useful for studying cancers which are relatively common and have been regularly treated by surgical procedures. Estimates tend to be closest to those of established registries in the middle age groups. At either extreme of age, there tends to be an underestimation of numbers of cases in the CPHA-PAS data. Among older patients, attempts at surgical therapy appear to be less common for most cancers than among middle-aged patients. In the youngest age groups, the numbers of cases are usually too small to permit reliable estimates of disease incidence.

For breast cancer, the CPHA-PAS estimates were quite accurate over a wide age range. For colon and testicular cancer, useful estimates were restricted to a few age groups. The procedures used in this paper provide low estimates of ovarian cancer rates. In addition, an attempt was made to estimate pancreatic cancer rates, but calculated rates appeared to be grossly underestimated.

Fuller use of this data resource will have to await the availability of detailed age- and sex-specific figures for surgery rates among patients afflicted with various cancers. These surgery rates could then be used to create effective "correction factors" to reverse the patterns of cancer rate underestimation which could result from uncritical use of the CPHA-PAS data.

REFERENCES

1. Greenberg ER, Colton T, Bagne C, Fayen E: Alternative method for determining cancer incidence and survival in the United States. Phase II final report to HEW Contract NCI NO1CP71047. Bethesda, MD: NCI, 1979.
2. Walker AM, Jick H: Cancer of the corpus uteri: increasing incidence in the United States, 1970–1975. Am J Epidemiol 1979; 110:47–51.
3. Walker AM, Jick H: Declining rates of endometrial cancer. Obstet Gynecol 1980; 56:733–736.
4. The Hospital Record Study: A joint study by the Commission on Professional and Hospital Activities-Professional Activity Study (CPHA-PAS), Ann Arbor, MI and IMS America Ltd, Ambler, PA, 1976.
5. Cutler SJ, Young JL (eds): Third National Cancer Survey: Incidence Data. Bethesda, MD: National Cancer Institute, Monograph 41, 1975.
6. Surveillance, Epidemiology, and End Results (SEER): Incidence and Mortality Data, 1973–1977. US Department of Health and Human Services, National Cancer Institute Monograph #57, NIH Pub. No. 81-2330, 1981.
7. Albert S, Belle S, Swanson GM: Recent trends in the treatment of primary breast cancer. Cancer 1978; 41:2399–2404.
8. Urban JA, Papachistou D, Taylor J: Bilateral breast cancer, biopsy of the opposite breast. Cancer 1977; 40:1968–1973.
9. Lewison EF, Neto AS: Bilateral breast cancer at the Johns Hopkins Hospital, a discussion of the dilemma of contralateral breast cancer. Cancer 1971; 28:1297–1301.
10. Report from the Cancer Survival, Epidemiology and End Results (SEER) Program: Cancer Patient Survival, Report Number 5. US Department of Health, Education, and Welfare, NIH Pub. No. 77-992, 1976.
11. Evans JT, Vana J, Aronoff BL, *et al:* Management and survival of carcinoma of the colon: results of a national survey by the American College of Surgeons. Ann Surg 1978; 188:716–720.
12. Munnell EW: The changing prognosis and treatment in cancer of the ovary. Am J Obstet Gynecol 1968; 100:790–805.
13. Clark DG, Hilaris BS, Ochoa M: Treatment of cancer of the ovary. Clinics in Obstet Gynaecol 1976; 3:159–179.
14. Girtanner R: Ovarian cancer: diagnosis and treatment. Compr Ther 1980; 6:30–38.
15. Gusberg SB, Frick HC: Corscaden's Gynecologic Cancer, 5th Ed. Baltimore: Williams and Wilkins, 1978.

ACKNOWLEDGMENT

The Boston Collaborative Drug Surveillance Program is supported in part by the Food and Drug Administration (Cooperative Agreement 5 U01 FD01223).

Name Change to Family Health International Adds New Emphasis

In 1982, the International Fertility Research Program, widely known as IFRP, changed its name to Family Health International. The new name reflects an expanded effort to improve the health of women and their families.

Family Health International continues IFRP's tradition of support for research and training on all aspects of contraceptive use and safety and adds a new emphasis on work to improve health care. Malcolm Potts is President of Family Health International. Headquarters remain in Research Triangle Park, North Carolina 27709, USA.