A Comparison of Stage of Presentation for Pancreatic and Colorectal Cancer in Pennsylvania 2000-2005

MARK CHIRUMBOLE[1], NIRAJ GUSANI[2], ALICIA HOWARD[1], TIM LEONARD[1], PETER LEWIS[3] and JOSH MUSCAT[1]

Departments of [1]Public Health Science, [2]Surgical Oncology, and [3]Family and Community Medicine, Pennsylvania State University, Hershey, PA, U.S.A.

Abstract. *Background: The goal of this study was to examine how rurality, socioeconomic status (SES) and access to medical care are related to the stage at presentation of patients with colorectal (CRC) and pancreatic cancer (PC) in Pennsylvania. Materials and Methods: Incident CRC and PC cases were identified from the Pennsylvania Department of Health. Demographic, SES, and access variables were collected at the county level. Results: Increased urbanization, younger age, and male gender were shown to be significantly related to later stage at diagnosis for PC. Age and education level were significant predictors of the rate of PC, while age, education level, insurance status, rurality, and the ratio of oncologists to primary care physicians were significant predictors of the rate of CRC. Conclusion: Based on county-level data, urban residence, younger age, and male gender were shown to be predictors of later stage at diagnosis for PC. These findings should help guide further research into factors that may be important predictors of later stage of diagnosis.*

Pennsylvania (PA) is a large diverse state with a great degree of variation in population demographics, degree of rurality and access to medical care. While PA is the sixth most populous state in the United States (US) and encompasses two large urban centers (Philadelphia and Pittsburgh), 48 out of the 67 counties in PA are classified as rural by US Census population standards. PA also has the second highest percentage of population over the age of 65 years at 15.6% (1).

The American Cancer Society estimates that colorectal cancer (CRC) will be the third most common cancer in the US for both males and females in 2008 in terms of both incidence and mortality. Pancreatic cancer (PC) is estimated to have the tenth highest incidence of cancer for males and the fourth highest cancer mortality for both males and females (2). From 1995 to 2005 Pennsylvania was above the US average for CRC age-adjusted incidence and mortality rates for CRC (3).

Stage at presentation is an important factor in determining survival from CRC and PC. The Surveillance, Epidemiology and End Results (SEER) Program estimates that the five-year relative survival rate for CRC for all races, ages and sexes is 90.2% for localized stage at diagnosis, 69.2% for regional stage and 10.8% for distant stage. The outlook for PC is even more bleak, at 17.9% to 8.1% to 1.8% for localized, regional and distant stage at diagnosis respectively (4).

CRC and PC are different in several ways. CRC has established screening and early detection programs, while PC has no screening test currently available; thus CRC is easier to prevent and diagnose at an earlier and potentially more treatable stage. Surgical treatment may be less complicated for CRC than PC. CRC may be treated by a general surgeon, subspecialty trained colorectal surgeon, or a surgical oncologist whereas pancreatic cancer generally requires the care of a surgical oncologist. PC usually presents at a later stage, while CRC is more often diagnosed at early stage. CRC has relatively favorable survival when found at the localized or regional stage, while PC has dismal survival, especially at the distant stage. Thus knowing that there are clear differences in clinical presentation and available effective mechanisms for cancer screening between CRC and PC coupled with the burden of morbidity and mortality for these two types of cancer, the goal of our study was to examine how rurality, socioeconomic status (SES) and access to medical care are related to the stage at presentation for CRC and PC in PA and how these predictors may differ between CRC and PC.

Patients and Methods

CRC and PC cases were identified from the PA Department of Health, Bureau of Health Statistics Research web site (5). For each county in PA, all CRC and PC cases from 2000 to 2005 were included and were classified by stage: *in situ*, localized, regional,

Correspondence to: Mark Chirumbole, 1503 Louisa Lane, Mechanicsburg, PA 17050, U.S.A. Tel: +1 7176358886, e-mail: Mark.chirumbole@eds.com

Key Words: Colorectal cancer, pancreatic cancer, stage at diagnosis, rurality, access to care.

distant, and unstaged. Demographic and SES variables were collected at the county level from the 2000 Decennial Census, Summary File 3, provided by the US Department of Commerce, Bureau of the Census (1). These variables included age, gender, ethnicity, educational attainment, population below the poverty line and rural population. Physician data per county were obtained from the American Medical Association (AMA) Physician-Related Data Resources database which includes AMA and non-AMA members (6). Physician specialties are listed in the database according to the self declared specialties of licensed MDs and osteopathic doctors (DO). The Primary Care Physician (PCP) group included internal medicine, family practice, general practice, and general surgeons. The Oncology Physician group included gastroenterologists, abdominal surgeons, surgical oncologists, radiation oncologists, and medical oncologists. The ratio of these two categories, Oncology Physicians divided by PCPs, was one measure of physician density used in this study. The second measure of physician density used in this study was the total number of MDs and DOs from the AMA database per 100,000 population. PA insured adults by county were derived from the Small Area Health Insurance Estimates (SAHIE) program (7).

Two types of multivariate statistical models were run for both CRC and PC. The first type of model run for each type of cancer was a Poisson regression model to assess what variables are significantly associated with the incidence rate of CRC and PC. The next model was an analysis of covariance model (ANCOVA) in which the outcome variable was calculated as the ratio of non-local cases to cases. Local cases were those that were staged as *in situ* or local and non-local cases were those that were staged as regional or distant. Cases that were unstaged were not included in the ANCOVA model, but they were included in the Poisson model for the overall rate. All analyses were done using SAS Enterprise Guide, version 4.

Several counties did not have any local cases of PC, so their ratio of non-local to local cancer at the time of diagnosis was undefined, thus a grouping variable was sought to combine the low-density counties with counties having similar SES characteristics. One established grouping is the PA Workforce Investment Areas (WIA) that were authorized by the Federal Workforce Investment Act (8). The 67 PA counties are grouped into 23 WIA by the PA Workforce Investment Board. Each WIA operates an independent labor pool, linking employment services, training and employment with job seekers. These geographic designations are used with the state monthly release of civilian labor force data, a major economic indicator and other important employment statistics, establishing WIA as a valid method for grouping county level data. One WIA did not coincide with a county level grouping and was thus not used.

Descriptive statistics were computed and tests for normality were run to see which variables would be entered into the models as continuous and which would be entered as categorical. The continuous variables were determined to be age (percentage of population greater than 50 years old), gender (percentage male), percentage of population that is rural, percentage of population that is uninsured, percentage of population that is a college graduate, and percentage of population that is below the poverty line. The variables that were not normally distributed and thus categorized were race, number of physicians per 100,000 population, and the ratio of oncology physicians to PCPs. The race variable was split into three groups. WIA that had a population greater than 90% Caucasian were coded as zero, those with a population between 80% and 90% Caucasian were coded as one, and those less than 80% were coded as two. Physicians per 100,000 population was split into three groups, with greater than 850 coded as zero, 25 to 850 coded as one, and less than 25 coded as two. The oncology to PCP ratio was also split into three groups, with greater than 0.12 coded as zero, 0.065 to 0.12 coded as one and less than 0.065 coded as two.

Table I. *Poisson regression model results for pancreatic cancer.*

Variable	Risk ratio	95% confidence Interval	*p*-Value
Gender	0.7202	0.3433-1.5107	0.3851
Age	1.6159	1.3628-1.9159	0.0001***
Poverty	1.1438	0.9727-1.3449	0.1041
Education	1.0718	1.0079-1.1398	0.0271***
Insurance	0.9507	0.8593-1.0518	0.3269
Rurality	1.0037	0.9738-1.0345	0.8118
Race			
Group 0 *vs*. Group 1	0.9590	0.8761-1.0498	0.3645
Group 0 *vs*. Group 2	0.9154	0.6981-1.2004	0.5228
Group 1 *vs*. Group 2	0.9546	0.7370-1.2363	0.7245
MDs per 100k population			
Group 0 *vs*. Group 1	0.9865	0.9103-1.0690	0.7396
Group 0 *vs*. Group 2	0.9431	0.8445-1.0533	0.2988
Group 1 *vs*. Group 2	0.9561	0.8882-1.0291	0.2318
Oncologist/PCP ratio			
Group 0 *vs*. Group 1	0.9599	0.8774-1.0503	0.3729
Group 0 *vs*. Group 2	0.9102	0.8104-1.0224	0.1125
Group 1 *vs*. Group 2	0.9482	0.8742-1.0284	0.1991

Results

PC in Pennsylvania had an incidence of 10,414 cases for the time period of 2000 to 2005, including 821 local cases, 8,041 non-local cases and 1,552 unstaged cases. The overall non-local to local ratio was 9.79. CRC in Pennsylvania had an incidence of 56,767 cases for the same time period, including 24,649 local cases, 28,383 non-local cases and 3,735 unstaged cases. The overall non-local to local ratio was 1.15 (9).

Table I shows risk ratios for the Poisson regression model with rate of PC as the outcome variable. For the continuous variables, the risk ratio represents the increase in the PC rate for a 10% increase in the percentage of persons. WIA with a greater percentage of population over 50 years of age or a greater percentage of population that is a college graduate were significant predictors of an increase in the rate of PC.

Table II shows the results of the ANCOVA model with the ratio of non-local cases to local cases for PC as the outcome variable. This model explained 76.98% of the variation in the outcome variable. WIA with a greater percentage of population that is male or a greater percentage of population that is less than 50 years of age had a statistically significant increase in the non-local to local ratio. WIA with a greater percentage of population that is rural showed a statistically significant decrease in the non-local to local ratio.

Table II. *ANCOVA model results for pancreatic cancer.*

Variable	Beta coefficient	*p*-Value
Gender	226.83	0.0537***
Age	–56.69	0.0551***
Poverty	–5.10	0.8442
Education	–8.78	0.3671
Insurance	29.53	0.1111
Rurality	–11.30	0.0170***
MDs per 100k population		0.1710
Group 0	1.81	0.2879
Group 1	2.28	0.0765
Group 2	Reference	Reference
Oncologist/PCP ratio		0.2738
Group 0	–1.49	0.3028
Group 1	–1.70	0.1191
Group 2	Reference	Reference
Race		0.1514
Group 0	2.76	0.5510
Group 1	–0.20	0.9617
Group 2	Reference	Reference

Table III. *Poisson regression model results for colorectal cancer.*

Variable	Risk ratio	95% confidence interval	*p*-Value
Gender	1.0299	0.7552-1.4045	0.8523
Age	1.6598	1.5470-1.7808	0.0001***
Poverty	0.9905	0.9257-1.0599	0.7834
Education	0.9502	0.9253-0.9756	0.0002***
Insurance	0.9395	0.9005-0.9803	0.0040***
Rurality	0.9861	0.9736-0.9987	0.0313***
Race			
Group 0 *vs*. Group 1	1.0183	0.9795-1.0586	0.3610
Group 0 *vs*. Group 2	0.9589	0.8541-1.0766	0.4775
Group 1 *vs*. Group 2	0.9417	0.8435-1.0514	0.2855
MDs per 100k population			
Group 0 *vs*. Group 1	0.9953	0.9622-1.0296	0.7867
Group 0 *vs*. Group 2	0.9924	0.9465-1.0405	0.7531
Group 1 *vs*. Group 2	0.9971	0.9660-1.0291	0.8561
Oncologist/PCP ratio	Reference	Reference	
Group 0 *vs*. Group 1	1.0751	1.0353-1.1165	0.0002***
Group 0 *vs*. Group 2	1.0788	1.0275-1.1326	0.0023***
Group 1 *vs*. Group 2	1.0034	0.9695-1.0384	0.8478

Table III shows the analysis of parameter estimates for the Poisson regression model with rate of CRC as the outcome variable. For the continuous variables, the risk ratio represents the increase in the PC rate for a 10% increase in the percentage of persons. WIA with a greater percentage of population over 50 years of age were a significant predictor of an increased rate of PC. WIA with a greater percentage of population that is a college graduate, a greater percentage of population that is uninsured, or a greater percentage rurality showed a statistically significant decrease in the rate of PC. The oncologist to PCP ratio group was also statistically significant. The group with the highest oncologist to PCP ratio had a PC rate that was significantly higher than the middle group and the low group.

Table IV shows the results of the ANCOVA model with the ratio of non-local cases to local cases for CRC as the outcome variable. This model explained 21.40% of the variation in the outcome variable. None of the variables were significant at the $p<0.05$ level. This model revealed that the age variable was associated with the ratio outcome in the opposite direction to which it was associated with the rate outcome.

Table IV. *ANCOVA model results for colorectal cancer.*

Variable	Beta coefficient	*p*-Value
Gender	7.25	0.6220
Age	–0.19	0.9585
Poverty	–0.97	0.7886
Education	–0.09	0.9442
Insurance	–0.55	0.8190
Rurality	–0.32	0.5624
MD's per 100k population		0.4877
Group 0	0.28	0.2452
Group 1	0.14	0.4100
Group 2	Reference	Reference
Oncologist/PCP ratio		0.9995
Group 0	0.002	0.9919
Group 1	–0.003	0.9852
Group 2		
Race		0.7236
Group 0	–0.47	0.4647
Group 1	–0.47	0.4347
Group 2	Reference	Reference

Discussion

The most significant findings of this study are that increased urbanization, younger age and male gender were significantly related to later stage at diagnosis for PC. Age and education level were significant predictors of the rate of PC, while age, education level, insurance status, rurality and the ratio of oncologists to PCPs were significant predictors of the rate of CRC. None of the variables were significant at the $p<0.05$ level for the ratio of late stage to early stage for CRC.

Several previous studies have found that the incidence of PC is greater in males (10-17). It has also been reported that PC is diagnosed late in the natural history of disease given the few early indicators of illness and the lack of screening tests (18). While the incidence rate for PC in men was not significant in this case, men were more likely to be

diagnosed at later stages, so this suggests that men are also at greater risk for more advanced disease. Women, especially post-menopausal women, are more likely to receive health care than men (19), so it is possible that this increases their chances of being detected at an earlier stage. Furthermore, bloating is a symptom of both menopause and PC, and therefore may lead to increased diagnostic testing for PC.

Increasing age distribution is a significant predictor of CRC and PC. In addition to this, for PC, the increasing age composition also leads to a significant decrease in the ratio of late stage to early stage cancer. The finding that younger people are more likely to be diagnosed with later stages of PC suggests that there may be a genetic component that makes one susceptible to more aggressive forms of the disease.

Results from the literature are mixed on the impact of the urban/rural gradient on PC. Several studies have found no geographic or urban/rural differences (12, 20). However, Blot *et al*. found that rates for PC were higher in urban areas in the US (21) and Boyle *et al*. found the same in France (11). Other research showed that urban residents had significantly higher mortality rates than rural residents (22). Since late stage at diagnosis leads to higher mortality rates, our finding that increased urbanization is a significant predictor of increased late stage at diagnosis complements this previous finding. Increased urbanization was also a statistically significant predictor of the rate of CRC. This finding is consistent with several other studies that have found CRC rates to be higher in urban areas (23-25).

One of our measures of access to medical care (oncologist to PCP ratio) was a significant predictor of CRC rate. The areas in the highest oncologist to PCP category had a significantly higher rate of CRC than the lower two categories. The high rate of specialists in the upper category may increase general awareness and the use of better screening and diagnostic methods. Scheiman *et al*. showed evidence of this when they reported that gastrointestinal (GI) consultation increased the probability of CR cancer screening completion eightfold as compared to those who only consulted a PCP (26).

Education was a significant predictor of the rate of CRC and PC, albeit in opposite directions. An increased percentage of college graduates was a significant predictor of PC rate. This effect was the opposite of CRC in which a decreased percentage of college graduates was a significant predictor of CRC rate. For CRC, adults with less than a high school education are less likely to report use of screening services such as a fecal occult blood test (FOBT), flexible sigmoidoscopy (FSIG), or colonoscopy than all adults 50 years of age and older (27). Less utilization of available services among those with less education may be a factor contributing to higher CRC rates in this group. PC does not have such an array of screening services currently available, so perhaps this is why those with more education were not any less likely to develop PC.

One unexpected result was found in the study. As the percentage of uninsured individuals increased, the CRC rates decreased. It was expected that the rate of CRC would increase as the percentage of uninsured increased due to lack of access to CRC screening in the uninsured. In a univariate linear regression analysis, the rate of CRC did increase as the percentage of uninsured increased, as would be expected. However, this was a very weak relationship (R^2=0.0349) and it ceased to exist after adjusting for all of the other variables in the multivariate model.

Several limitations of this study should be noted. The ecological study design employed is useful for indicating where etiological relationships may exist, however it does not demonstrate that causal pathways exist. The data collected are at the group (*i.e*. county) level, so collecting individual level data should be the next step to investigate these relationships. Personal factors such as family history, smoking status, and diet have also been indicated as causative factors for PC and should be looked at in more detail (28, 29).

Collecting data at the county level may affect the analysis in other ways also. County of residence reflected the place of residence at the time of diagnosis, however it may not be indicative of a person's lifetime experience with regards to the variables studied and it also may not reflect where an individual has access to medical care. For example, one may live in an urban environment for most of their life, but retire to a rural setting and then subsequently be diagnosed with cancer. In addition, as previously mentioned, the 67 PA counties were grouped into 22 WIA due to lack of data in some counties. Doing this facilitated data analysis, but it also decreased the sample size. A smaller sample size results in fewer degrees of freedom after adjusting for all of the covariates in the statistical models and thus less ability to detect significant associations.

Acknowledgements

This study was supported by research grants K07 CA104231 from the National Cancer Institute, and PA-DOH 4100038714, PA-DOH 4100038715 from the Pennsylvania Department of Health.

References

1 U.S. Census Bureau. Census 2000 Summary File 3 – Pennsylvania. Washington, DC: U.S. Department of Commerce, 2002.

2 Jemal A, Siegel R, Ward E, Murray T, Xu J and Thun MJ: Cancer statistics, 2007. Cancer J Clinicians *57(1)*: 43-66, 2007.

3 Pennsylvania Department of Health. Cancer Facts and Figures Pennsylvania, 2008. Available at: www.dsf.health.state.pa.us/health/lib/health/FactsnFigures/CFF_2008.pdf Accessed March, 2009.

4 Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov) SEER*Stat Database: Incidence - SEER 17 Regs Limited-Use + Hurricane Katrina Impacted Louisiana Cases, Nov 2007 Sub (1973-2005 varying), National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch, released April 2008, based on the November 2007 submission. Accessed March, 2009.
5 Pennsylvania Department of Health. Epidemiologic Query and Mapping System. Available at: http://app2.health.state.pa.us/epiqms/Asp/ChooseDataset.asp. Accessed April, 2008.
6 American Medical Association. Available at: http://www.ama-assn.org/cgi-bin/sserver/datalist.cgi. Accessed April, 2008.
7 U.S. Census Bureau. Data Integration Division, Small Area Estimates Branch. Available at: http://www.census.gov//did/www/sahie/methods/source.html. Accessed April, 2008.
8 Pennsylvania Partners. Workforce Investment Areas. Available at: http://www.papartners.org/pawia.htm. Accessed July, 2008.
9 Pennsylvania Department of Health. Bureau of Health Statistics Research. Colorectal and pancreatic cancer incidence in Pennsylvania by county 2000-2005. Available at: http://app2.health.state.pa.us/epiqms/. Accessed April, 2008.
10 Nasca PC, Burnett WS, Greenwald P, Brennan K, Wolfgang P and Carlton K: Population density as an indicator of urban-rural differences in cancer incidence, upstate New York, 1968-1972. Am J Epidemiol *112(3)*: 362-375, 1980.
11 Boyle P, Hsieh CC, Maisonneuve P, La Vecchia C, Macfarlane GJ, Walker AM and Trichopoulos D: Epidemiology of pancreas cancer. Int J Pancreatology *5(4)*: 327-346, 1989.
12 Levin DL, Connelly RR and Devesa SS: Demographic characteristics of cancer of the pancreas: mortality, incidence, and survival. Cancer *47(6 Suppl)*: 1456-1468, 1981.
13 Faivre J, Bedenne L, Arveux P and Klepping C: Descriptive epidemiology of cancer of the pancreas. Bull Cancer *77(1)*: 39-46, 1990.
14 Gold EB: Epidemiology of and risk factors for pancreatic cancer. Surg Clin North America *75(5)*: 819-843, 1995.
15 Ghadirian P, Lynch HT and Krewski D: Epidemiology of pancreatic cancer: an overview. Cancer Det Prev *27(2)*: 87-93, 2003.
16 ShaibYH, Davila JA and El-Serag HB: The epidemiology of pancreatic cancer in the United States: changes below the surface. Aliment Pharm Ther *24(1)*: 87-94, 2006.
17 Mack TM and Paganini-Hill A: Epidemiology of pancreas cancer in Los Angeles. Cancer *47(6 Suppl)*: 1474-1484, 1981.
18 Michaud DS: Epidemiology of pancreatic cancer. Minerva Chirurgica *59(2)*: 99-111, 2004.
19 Owens GM: Gender differences in health care expenditures, resource utilization, and quality of health care. J Manag Care Pharm *14(3)(suppl S)*: S2-S6, 2008.
20 Ahlgren JD: Epidemiology and risk factors in pancreatic cancer. Sem Oncol *23(2)*: 241-250, 1996.
21 Blot WJ, Fraumeni JF Jr and Stone BJ: Geographic correlates of pancreas cancer in the United States. Cancer *42(1)*: 373-80, 1978.
22 Chen KX, Wang PP, Zhang SW, Li LD, Lu FZ and Hao XS: Regional variations in mortality rates of pancreatic cancer in China: results from 1990-1992. World J Gastroenterol *9(11)*: 2557-2560, 2003.
23 Valerianova Z, Gill C, Duffy SW and Danon SE: Trends in Incidence of Various Cancers in Bulgaria, 1981-1990. Int J Epidemiology *23(6)*: 1117-1126, 1994.
24 Richmond RE, Rickabaugh J, Huffman J and Epperly N: Colorectal Cancer Mortality and Incidence in Campbell County, Kentucky. South Med J *80(8)*: 953-957, 1987.
25 Coughlin SS, Richards TB, Thompson T, Miller BA, Van Eenwyk J, Goodman MT and Sherman RL. Rural/nonrural differences in colorectal cancer incidence in the United States, 1998-2001. Cancer *107(5 Suppl)*: 1181-1188, 2006.
26 Menees SB, Scheiman J, Carlos R, Mulder A and Frederick AM: Gastroenterologists utilize the referral for EGD to enhance colon cancer screening more effectively than primary care physicians. Aliment Pharmacol Ther *23(7)*: 953-962, 2006.
27 Centers for Disease Control and Prevention. National Center for Chronic Disease Prevention and Health Promotion, Behavioral Risk Factor Surveillance System. CD-ROM (1996-1997, 1999) and Public Use Data Tape (2001, 2002, 2004, 2006).
28 Lowenfels AB and Maisonneuve P: Epidemiologic and etiologic factors of pancreatic cancer. Hematol – Oncol Clin North Am *16(1)*: 1-16, 2002.
29 Hart A, Kennedy H and Harvey I: Pancreatic cancer: A review of the evidence of causation. Clinical Gastroenterol Hepatol *6*: 275-282, 2008.

Received March 16, 2009
Revised May 26, 2009
Accepted May 28, 2009