VARIATION IN COLORECTAL CANCER INCIDENCE AMONG STATES AND METROPOLITAN AREAS IN THE UNITED STATES, 1991-1995

John P. Fulton, Xiao-Cheng Wu, Susan E. Carozza, Robert Greenlee, Lihua Liu, Brooke Steele, Vivien W. Chen

INTRODUCTION

Colorectal cancer is among the five most common malignant neoplasms diagnosed in the United States, accounting for an estimated 130,200 new cases and 56,000 deaths annually.[1] One of the prominent risks for colorectal cancer is diet, although the nutrients that cause the disease remain the subjects of research and debate.[2] Red meat may increase the risk of colorectal cancer, while diets rich in folate may have a protective effect, although other dietary components have been shown to be related to colorectal cancer risk, as well.[3] Research has demonstrated the salutary effects of screening for colorectal cancer with the fecal occult blood test, sigmoidoscopy, colonoscopy or the barium enema.[4-10] Colorectal cancer screening has the added benefit of preventing some malignancies by facilitating the removal of pre-cancerous polyps from the colon.

Colorectal cancer incidence in the United States has been monitored since 1973, with the advent of the Surveillance, Epidemiology, and End Results (SEER) program sponsored by the National Cancer Institute. The SEER system currently collects population-based incidence data from geographical areas in which about 14 percent of the U.S. population resides. Among the primary goals of the SEER system are producing accurate estimates of the incidence of cancer in the U.S. as a whole, and measuring racial and ethnic differences in cancer incidence.[11] Between 1973 and 1985, colorectal cancer incidence increased in SEER areas. Beginning in 1986, however, colorectal cancer incidence began to decrease, and it has decreased steadily since then, accompanied by a shift in the distribution of stage of disease at diagnosis toward less advanced disease.[12,13] Although SEER data may be used to explore geographic differentials in cancer incidence rates, this is not their strength, as their coverage encompasses only five states (CT, HI, IA, NM, UT), and nine additional metropolitan areas, groups of rural counties, and special populations.

The present study uses data available from population-based central cancer registries and other sources to measure gross geographic variation in age-adjusted colorectal cancer incidence across states and metropolitan areas in the United States, 1991-1995, and, using selected essential variables, to explore potential determinants of geographic variation, suggesting directions for future surveillance and analysis. Cancer incidence data were assembled by the North American Association of Central Cancer Registries (NAACCR) for its annual publication, *Cancer in North America* (CINA), including data from 11 of the 14 SEER areas (five states, six metropolitan areas), and data from another 25 states across the United States. NAACCR also publishes data from Canada, but these data were not used, because assessing the comparability of U.S. and Canadian colorectal cancer incidence data was beyond the scope of the present study.

METHODS

Data on newly diagnosed cases of invasive colorectal cancer (from NAACCR) and on population (from the U.S. Bureau of the Census) were used to compute age-adjusted colorectal cancer incidence rates (all races combined) by sex and year (1991-1995) for each of the 36 geographic areas. Average annual age-adjusted colorectal cancer incidence rates were computed similarly for the five-year period, 1991-1995. The 1970 population of the United States was used as the standard population in computing age-adjusted rates.

The trend in age-adjusted colorectal cancer incidence rates, 1991-1995, by sex, was computed for each of the 36 geographic areas using a measure pioneered by the SEER program, the estimated annual percent change, or EAPC.[11] The EAPC is computed by regressing the natural logarithm of the variable of interest (age-adjusted incidence), against a measure of time (calendar year), then using the slope of the line generated, ***m***, to estimate the average annual percent change in the variable of interest:

$$EAPC = 100*(e^{m}-1)$$

The principal advantage of the EAPC over "percent change" between two points in time is that EAPC is less sensitive to random variations in the variable of interest, if more than two points in time are available for analysis.

Data from the U.S. census bureau and CDC's Behavioral Risk Factor Surveillance System (BRFSS) were used to create six additional variables to be tested as potential determinants of age-adjusted colorectal cancer incidence rates and EAPC for each of the 36 geographic areas. These variables included:

1) SEER status (proxy for completeness of case reporting)

2) Inclusion by NAACCR in combined U.S. rates (proxy for completeness of case reporting)[14]

    - Explanation: Geographic variation in the completeness of cancer case reporting may contribute to the geographic variation in cancer incidence rates. SEER registries and central cancer registries whose data pass rigorous quality tests to be used by NAACCR in combined U.S. rates have more complete cancer case reporting than other central cancer registries. Areas with more complete case reporting generally have higher incidence rates than areas with less complete case reporting.

3) Percent white in the 1993 U.S. population, by sex (racial composition)

    - Explanation: Geographic variation in racial composition may contribute to geographic variation in cancer incidence rates if the incidence rates in question vary substantially by race. (In fact, colorectal cancer incidence rates do not have substantial differentials by race in the United States. The indicator for racial composition was included in the analysis to be thorough.)

4) Percent of the population ages 50 and older who were ages 65 and older in 1993, by sex (age structure)

    - Explanation: Even though age-adjustment is intended to remove the effects of age from age-adjusted cancer incidence rates, it is a crude process. Residual age effects may contribute to geographic variation in age-adjusted cancer incidence rates if the age structures of the populations at risk vary substantially across geographic areas.

5) Percent of the 1990 population who lived in urban areas as defined by the U.S. census (urbanization)

    - Explanation: Certain cancers have higher rates in urban areas than non-urban areas.[15] Urbanization may contribute to geographic variation in cancer incidence rates if the geographic areas in question vary substantially by level of urbanization.

6) Percent of the population ages 50 and older in 1993 who had received a proctoscopic exam in the past 5 years (screening behavior).

- Explanation: Screening may lower cancer incidence rates by facilitating the removal of pre-cancerous lesions, or may raise cancer incidence rates, at least temporarily, by identifying cancerous lesions earlier than they would have been identified without screening. Therefore, geographic variation in screening rates may contribute to geographic variation in cancer incidence rates.

Because BRFSS data are available for states only, state-level data were allocated to metropolitan areas as gross estimates. The latter probably had the effect of attenuating observed relationships between variables, because of the error thus introduced.

The geographic distributions of incidence and EAPC were described by computing ranges, medians, means, and standard deviations, and by constructing histograms. The geographic distributions of the six potential determinants of incidence and EAPC were evaluated to assure that each had sufficient variance to be used as a variable in bivariate analysis. All six were found to have moderate to substantial variance, without extreme outliers. Bivariate Pearson correlation coefficients (R) were computed for all pairs of variables. The 0.05 probability level was used to evaluate statistical significance.

RESULTS

Geographic Variation in Incidence

The geographic variation in the average annual age-adjusted incidence of colorectal cancer for the period 1991-1995 was substantial, ranging from 40.4 to 65.7 for males, and from 27.8 to 44.7 for females (Table 1). (All rates are expressed as cases per 100,000 population at risk.) The geographic distribution of age-adjusted incidence was approximately normal (Figures 1 and 2). Incidence rates were higher for males than females; means and medians were approximately equal within genders (cf. 52.9 with 53.6 for males; 37.2 with 37.1 for females).

Geographic Variation in EAPC

The EAPC for the period 1991-1995 varied substantially across geographic regions, ranging from -6.8 percent to 5.5 percent for males, and -7.7 percent to 2.5 percent for females (Table 1). The mean EAPC for all geographic areas combined reflected an approximately two percent annual decrease in the rate for both men and women. The geographic distribution of EAPC was approximately normal (Figures 3 and 4). As with incidence rates, means and medians were approximately equal within genders.

Bivariate Analysis: Incidence

Incidence was moderately correlated with the age structure of the population at risk (males, 0.34; females, 0.51) (Table 2). Incidence was higher in older populations. This residual age effect may explain 12-26 percent of the variance in age-adjusted incidence rates (R^2). Incidence was not very sensitive to data quality, racial composition, or urbanization. Correlations between incidence and these variables ranged between 0.01 and 0.15 (disregarding sign). The correlation between incidence and screening behavior was suggestive of a relationship (males, 0.29; females, 0.25), but was not statistically significant at the 0.05 probability level.

Bivariate Analysis: EAPC

APC was moderately correlated with screening behavior and urbanization. Greater decreases in incidence were observed in populations with higher screening rates and populations that were more urban. These relationships were statistically significant at the 0.05 probability level for males, but not for females. Screening behavior and urbanization were themselves moderately correlated. EAPC also was moderately

correlated with one of the two measures of completeness of cancer case reporting ("included in combined rates"). This apparent relationship may be spurious, however, because the measure of completeness of case reporting was itself moderately correlated with screening behavior and urbanization.

CONCLUSIONS

The geographical variation in colorectal cancer incidence observed across states and metropolitan areas in the United States was quite substantial, suggesting the possibility of large future declines in areas with a high colorectal cancer burden.

The correlations between age-adjusted incidence rates and an indicator of age structure *suggest* the existence of residual, uncontrolled age effects in the rates. This phenomenon is understood, but seldom measured when age-adjusted rates are compared. Residual age effects may be large when certain conditions prevail (e.g., large differences in age-specific rates *within* age groups, especially final age groups, like "85+," and populations at risk with very irregular age distributions), and should be assessed. Of course, it is also possible that the relationship observed may be spurious, caused by a third variable related to each of the other two. These alternative explanations may be evaluated by recomputing new age-adjusted rates, based on narrower age groups, and then recomputing the correlation between rates age structure. If the new correlation coefficient is substantially smaller than the first, the magnitude of the first was probably caused by residual age effects. If the new correlation coefficient is about the same as the original, its magnitude may be caused by a third variable related to each of the other two.

Geographic variation in EAPC ran the gamut from positive to negative, although the central tendency was clearly negative. The downward trend in colorectal cancer incidence (EAPC) was moderately correlated with level of screening. Colorectal cancer screening, like screening with the Pap test, can actually *reduce* cancer incidence by identifying and facilitating the removal of *precancerous* lesions. The level of screening appears to be more intense in more urbanized areas than less urbanized areas, suggesting greater access to screening services in the former. Hopefully, this geographic differential will diminish over time, as colorectal cancer screening becomes adopted in more primary care practices, especially in less urban areas.

A once-strong urban-rural differential in colorectal cancer *mortality* began declining in the United States over 20 years ago.[15] Correlations between colorectal cancer incidence in 1991-1995 and urbanization were quite low, despite sufficient variance in the latter variable to be strongly correlated with other basic variables, suggesting that perhaps the decline in the rural-urban differential has run its course. The negative correlations between urbanization and EAPC may explain why correlations between urbanization and colorectal cancer were so low. Apparently, colorectal cancer incidence declined faster in more urban than less urban areas in 1991-1995. This differential in the rate of decline may have reduced the original geographic differential in colorectal cancer incidence. If this is indeed true, and were this differential in the rate of decline to persist, it might create a new differential, this time with incidence rates higher in *less urban* areas. This, of course, is quite speculative. If geographic differentials in screening lessen, as suggested earlier, the low correlation between urbanization and colorectal cancer may persist indefinitely. These variables should be closely monitored, as their trends and relationships have important implications for the deployment of new screening assets.

The present exploratory study analyzes *ecological* relationships, and as such is merely suggestive, not conclusive. Its purpose was to measure gross geographic variation in age-adjusted colorectal cancer incidence across states and metropolitan areas in the United States, 1991-1995, and, using selected essential variables, to explore potential determinants of geographic variation, suggesting directions for future surveillance and analysis. Clearly, the relationships among urbanization, screening, and incidence

deserve closer attention. Attempting to replicate the present analysis using county data might be a good place to start.

REFERENCES

1. American Cancer Society (ACS). Cancer Facts & Figures - 2000. Atlanta, Georgia, 2000.

2. Trock B, Lanza E, Greenwald P. Dietary fiber, vegetables and colon cancer: critical review and meta-analyses of the epidemiologic evidence. *J Natl Cancer Inst* 1990;82:650-661.

3. Colditz G, DeJong W, Hunter D, Trichopoulos D, Willett W. Harvard report on cancer prevention. Volume 1. *Cancer Causes and Control* 1996;7(Supplement 1):S1-S59.

4. Mandel JS, Bond JH, Church TR, et al. Reducing mortality from colorectal cancer by screening for fecal occult blood. *N Engl J Med* 1993;328:1365-1371.

5. Newcomb PA, Norfleet RG, Storer BE, et al. Screening sigmoidoscopy and colorectal cancer mortality. *J Natl Cancer Inst* 1992;84:1572-1575.

6. Selby JV, Friedman GD, Quesenberry CP, et al. A case-control study of screening sigmoidoscopy and mortality from colorectal cancer. *N Engl J Med* 1992;326:653-657.

7. Gilbertsen VA. Proctosigmoidoscopy and polypectomy in reducing the incidence of rectal cancer. *Cancer* 1974;34(Supply):936-939.

8. Winawer SJ, Zauber AG, Ho MH, et al. Prevention of colorectal cancer by colonoscopic polypectomy. *N Engl J Med* 1993;329:1977-1981.

9. Thiss-Evensen E, Hoff GS, Sauar J, Langmark F, Majak BM, Vatn MH. Population-based surveillance by colonoscopy: effect on the incidence of colorectal cancer. Telemark Polyp Study I. *Scand J Gastroenterol* 1999;34:414-420.

10. Winawer SJ, Fletcher RH, Miller L, *et al.* Colorectal cancer screening: clinical guidelines and rationale. *Gastroenterology* 1997;112:594-642.

11. Ries LAG, Kosary CL, Hankey BF, Miller BA, Clegg L, Edward BK (eds). *SEER Cancer Statistics Review, 1973-1996.* Bethesda, MD: National Cancer Institute, 1999.

12. Chu KC, tarone RE, Chow WH, Alexander GA. Colorectal cancer trends by race and anatomic subsites, 1975 to 1991. *Arch Fam Med* 1995;4:849-856.

13. Troisi RJ, Freedman AN, Devesa SS. Incidence of colorectal carcinoma in the U.S.: an update of trends by gender, race, age, subsite, and stage, 1975-1994. *Cancer* 1999;85(8):1670-1676.

14. Chen VW, Wu XC, Andrews PA (eds). 1999 Cancer in North America 1991-1995. Volume One: Incidence. Sacramento, CA: North American Association of Central Cancer Registries, April, 1999.

15. Greenberg MR. Urbanization and cancer mortality. The United States experience, 1950-1975. New York: Oxford University Press, 1983.

Table 1. Annual age-adjusted* incidence** for the years 1991, 1992, 1993, 1994, and 1995, average annual age-adjusted incidence for the period 1991-1995, and estimated annual percent change (EAPC) in age-adjusted incidence of colo-rectal cancer, by sex, all races, 36 NAACCR registries.

Males								Females							
	Age-Adjusted Incidence								Age-Adjusted Incidence						
REGISTRY	1991	1992	1993	1994	1995	1991-1995	EAPC	REGISTRY	1991	1992	1993	1994	1995	1991-1995	EAPC
AZ	39.0	48.0	40.8	39.4	47.0	42.9	1.8	AZ	25.8	31.4	29.0	29.2	30.4	29.2	2.5
CA	53.5	51.9	49.6	49.6	47.5	50.3	-2.8	CA	36.7	35.8	36.0	33.7	32.9	35.0	-2.8
CALA	53.9	52.4	49.7	50.2	48.7	50.9	-2.5	CALA	36.7	34.4	34.9	33.7	32.5	34.4	-2.6
CASF	57.2	57.1	53.5	51.8	48.0	53.4	-4.4	CASF	37.5	37.4	38.3	35.7	33.7	36.5	-2.5
CASJ	54.3	50.2	46.9	48.3	47.1	49.3	-3.2	CASJ	37.0	38.6	38.7	33.3	36.2	36.7	-1.9
CO	51.5	48.6	46.6	43.2	43.4	46.5	-4.5	CO	36.1	35.2	32.8	30.1	33.3	33.4	-3.1
CT	63.5	57.9	61.8	61.7	54.1	59.8	-2.5	CT	40.8	41.6	40.4	41.7	41.8	41.2	0.5
DE	76.7	61.3	48.8	61.8	59.9	61.5	-4.7	DE	55.3	42.5	34.7	37.8	39.3	41.8	-7.7
FL	59.2	58.8	57.0	54.7	54.5	56.8	-2.4	FL	41.6	41.2	39.4	39.5	38.3	40.0	-2.0
GAAT	57.0	56.0	56.4	48.7	49.4	53.4	-4.2	GAAT	39.6	42.5	38.9	44.3	36.4	40.3	-1.3
HI	62.5	52.7	57.0	52.1	56.4	55.9	-2.1	HI	36.8	38.2	39.0	32.4	35.5	36.3	-2.3
IA	58.5	61.7	61.6	55.5	58.7	59.2	-1.0	IA	44.5	43.9	45.5	41.7	41.7	43.5	-1.8
ID	45.2	45.3	45.9	43.7	41.2	44.2	-2.2	ID	33.9	32.6	29.7	28.4	30.8	31.0	-3.2
IL	58.7	58.6	56.2	55.4	54.7	56.7	-2.0	IL	40.0	38.6	39.5	41.7	39.7	39.9	0.6
KY	53.0	53.7	51.3	53.8	56.4	53.7	1.3	KY	39.8	38.4	38.0	39.2	39.4	39.0	0.0
LA	54.0	58.1	53.2	51.8	52.6	53.9	-1.7	LA	40.2	37.0	42.6	39.7	37.2	38.8	-1.6
MA	67.4	61.7	62.3	61.1	56.0	61.6	-3.7	MA	43.4	40.7	40.7	39.6	38.7	40.6	-2.6
MI	62.5	58.7	55.3	52.2	50.5	55.7	-5.3	MI	42.0	41.0	38.0	37.5	35.1	38.7	-4.4
MIDE	65.8	61.5	59.9	57.4	53.8	59.6	-4.6	MIDE	42.4	39.9	39.9	38.4	40.3	40.1	-1.4
MN	56.3	57.6	52.3	52.0	52.6	54.1	-2.3	MN	39.6	37.9	37.3	37.1	36.1	37.6	-2.1
MT	53.7	46.5	47.6	46.6	43.4	47.4	-4.2	MT	35.4	31.8	35.0	34.2	35.1	34.3	0.5
NC	48.6	46.1	46.0	44.8	46.1	46.3	-1.3	NC	34.2	34.1	33.7	31.9	33.4	33.5	-1.1
NE	56.0	52.4	50.3	53.7	52.0	52.9	-1.2	NE	37.7	34.4	35.6	36.8	36.5	36.2	0.2
NH	56.1	58.9	49.9	59.1	56.9	56.0	0.3	NH	40.8	41.3	42.9	46.4	40.2	42.3	0.9
NJ	71.0	68.9	65.2	61.9	61.8	65.7	-3.8	NJ	47.2	45.2	44.8	43.3	43.0	44.7	-2.3
NM	41.9	44.5	36.4	41.8	39.7	40.8	-1.7	NM	30.4	30.1	24.3	31.4	26.8	28.5	-2.1
NV	54.8	51.8	47.1	34.9	47.0	46.7	-6.8	NV	40.2	35.9	30.2	25.7	33.4	32.8	-6.8
NY	58.7	56.2	54.9	56.3	54.0	56.0	-1.7	NY	40.1	39.2	39.8	41.2	39.3	39.9	0.1
PA	63.9	65.5	61.5	60.4	59.4	62.1	-2.3	PA	44.7	42.8	42.7	42.6	40.6	42.7	-2.0
RI	68.1	63.8	61.2	67.5	61.5	64.4	-1.5	RI	45.5	41.1	47.6	43.2	37.5	43.0	-3.3
TN	47.5	47.4	45.7	46.8	47.3	46.9	-0.2	TN	33.2	35.0	35.0	35.2	34.0	34.5	0.5
UT	44.7	45.0	37.6	41.2	38.7	41.3	-3.7	UT	28.7	29.0	27.4	26.4	27.7	27.8	-1.7
VA	47.6	46.5	46.4	43.1	42.7	45.2	-2.9	VA	33.1	32.1	34.0	32.9	32.0	32.8	-0.4
WASE	54.6	53.4	49.0	51.0	46.7	50.9	-3.5	WASE	38.5	37.6	36.1	34.5	35.8	36.5	-2.3
WI	61.0	63.5	61.7	59.4	58.1	60.6	-1.6	WI	41.0	44.1	40.6	40.8	41.8	41.7	-0.4
WY	32.6	42.0	42.7	41.0	43.1	40.4	5.5	WY	35.4	34.2	31.2	36.8	35.3	34.5	0.7
RANGE								RANGE							
High	76.7	68.9	65.2	67.5	61.8	65.7	5.5	High	55.3	45.2	47.6	46.4	43.0	44.7	2.5
Median	56.1	54.9	50.8	51.9	51.3	53.6	-2.4	Median	39.6	38.1	38.0	36.9	36.2	37.1	-1.9
Low	32.6	42.0	36.4	34.9	38.7	40.4	-6.8	Low	25.8	29.0	24.3	25.7	26.8	27.8	-7.7
MEAN	55.8	54.6	51.9	51.5	50.9	52.9	-2.3	MEAN	35.8	37.7	37.1	36.6	36.2	37.2	-1.6
STD DEV	9.0	6.8	7.3	7.6	6.4	6.9	2.2	STD DEV	5.4	4.2	5.2	5.2	4.0	4.4	2.0

* Age-adjusted to the 1970 population of the U.S.

** Incidence expressed as cases per 100,000.

Figure 1. Average annual age-adjusted colorectal cancer incidence, males, all races, 1991-1995, 36 NAACCR registries.

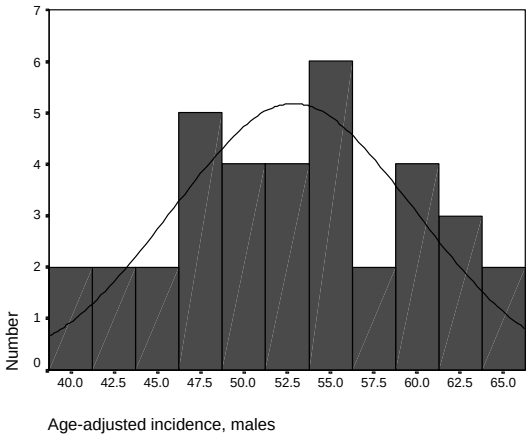

Figure 2. Average annual age-adjusted colorectal cancer incidence, females, all races, 1991-1995, 36 NAACCR registries.

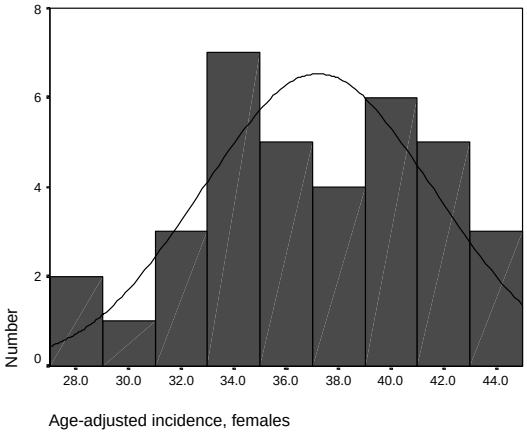

Figure 3. Estimated annual % change in age-adjusted colorectal cancer incidence, males, all races, 1991-1995, 36 NAACCR registries.

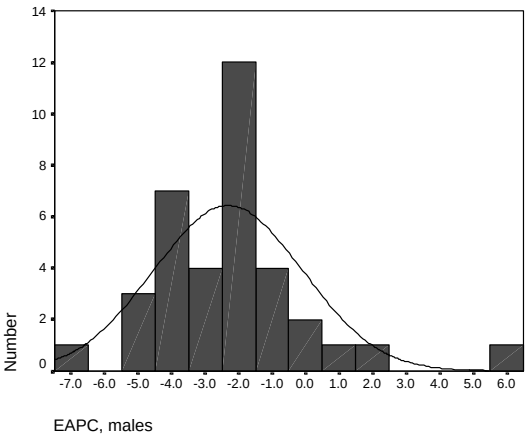

Figure 4. Estimated annual % change in age-adjusted colorectal cancer incidence, females, all races, 1991-1995, 36 NAACCR registries.

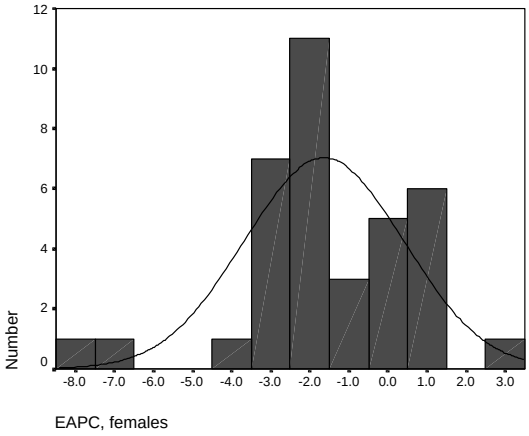

Table 2. Correlation matrix, average annual age-adjusted* incidence,** and estimated annual percent change (EAPC) in age-adjusted incidence of colo-rectal cancer, by sex, all races, 1991-1995, with key potential determinants of incidence and EAPC.

Males

Variable number and description ↓ // Variable number →	1	2	3	4	5	6	7	8
1. Age-adjusted incidence, 1991-1995	1.00	-0.24	0.15	-0.06	-0.12	§ 0.34	0.13	0.29
2. EAPC	-0.24	1.00	-0.09	-0.22	0.22	0.13	§-0.43	§-0.47
3. Included in combined rates	0.15	-0.09	1.00	§ 0.44	-0.15	-0.12	§ 0.35	§ 0.42
4. Part of SEER system	-0.06	-0.22	§ 0.44	1.00	§-0.35	-0.26	§ 0.53	§ 0.36
5. % pop who are white	-0.12	0.22	-0.15	§-0.35	1.00	0.13	§-0.40	-0.26
6. % pop 50+ who are 65+, 1993	§ 0.34	0.13	-0.12	-0.26	0.13	1.00	-0.08	-0.10
7. % pop in urban places, 1990	0.13	§-0.43	§ 0.35	§ 0.53	§-0.40	-0.08	1.00	§ 0.41
8. % pop 50+ procto exam, past 5 years, 1993	0.29	§-0.47	§ 0.42	§ 0.36	-0.26	-0.10	§ 0.41	1.00

Females

Variable number and description ↓ // Variable number →	1	2	3	4	5	6	7	8
1. Age-adjusted incidence, 1991-1995	1.00	-0.08	0.15	-0.11	-0.03	§ 0.51	0.01	0.25
2. EAPC	-0.07	1.00	§-0.50	-0.20	0.13	0.03	-0.32	-0.32
3. Included in combined rates	0.15	§-0.50	1.00	§ 0.44	-0.15	-0.15	§ 0.35	§ 0.42
4. Part of SEER system	-0.11	-0.20	§ 0.44	1.00	§-0.35	-0.30	§ 0.53	§ 0.36
5. % pop who are white	-0.03	0.13	-0.15	-0.35	1.00	0.35	-0.40	-0.26
6. % pop 50+ who are 65+, 1993	§ 0.51	0.03	-0.15	-0.30	§ 0.35	1.00	-0.19	-0.10
7. % pop in urban places, 1990	0.01	-0.32	§ 0.35	§ 0.53	§-0.40	-0.19	1.00	§ 0.41
8. % pop 50+ procto exam, past 5 years, 1993	0.25	-0.32	§ 0.42	§ 0.36	-0.26	-0.10	§ 0.41	1.00

* Age-adjusted to the 1970 population of the U.S.
** Incidence expressed as cases per 100,000.
§ $P \leq .05$