

Contribution of Socioeconomic Status to Black/White Differences in Cancer Incidence

WILLIAM P. MCWHORTER, MD, MPH, ARTHUR G. SCHATZKIN, MD, DRPH, JOHN W. HORM, MSc,
AND CHARLES C. BROWN, PhD

Blacks and Whites have very different cancer incidence rates for many sites, but this may largely be due to the racial differences in socioeconomic status (SES). The authors tested this hypothesis by determining the effect of adjustment for SES on the black/white incidence ratios for 12 cancer sites. Race-specific census tract-level SES variables (median family income, percent below poverty level, and years of education) were obtained from the 1980 US census and applied to approximately 20,000 black and 88,000 white cancer cases from the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) program for the years 1978 to 1982. For each cancer site (with each sex considered separately), Poisson regression was used to produce age-adjusted black/white incidence ratios, with and without adjustment for SES. The SES variable with the strongest adjusting power was percent below poverty level. For many sites (breast, *in situ* and invasive cervix, esophagus, male lung, pancreas, stomach) poverty accounted for much or all of the racial differences. For several sites (bladder, multiple myeloma, prostate, uterine corpus), large racial differences persisted after adjustment for poverty, and these findings suggest directions for investigating the etiology of these cancers.

Cancer 63:982-987, 1989.

BLACKS AND WHITES are known to have very different cancer incidence rates for a number of anatomic sites. For many sites (*e.g.*, cervix, esophagus, larynx, male lung, multiple myeloma, pancreas, prostate, stomach), black rates are higher than those of whites. For a smaller number of sites (*e.g.*, bladder, breast, ovary, rectum, uterine corpus), black rates are lower than those of whites. This has been summarized recently using data from the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) program.¹

The causes of these racial differences have not been clearly shown, but several lines of evidence support the belief that a large factor is socioeconomic status (SES), a broad term which usually includes some measure of income, education, or occupation: (1) as a group, blacks have considerably lower measures of SES than Whites²; (2) exposures to some known carcinogens, *e.g.*, smoking, vary with SES³⁻⁵; and (3) cancer rates are known to vary with SES for a number of sites.⁶⁻⁹

From the Division of Cancer Prevention and Control, National Cancer Institute, Bethesda, Maryland.

The authors thank Bob Smucker of IMS, Inc., for his assistance; and Dr. Steven Piantadosi, formerly of the Division of Cancer Prevention and Control, and Dr. Susan Devesa of the Division of Cancer Etiology, National Cancer Institute, for their helpful advice.

Address for reprints: William P. McWhorter, MD, MPH, Division of Hematology-Oncology, University of Utah School of Medicine, Salt Lake City, UT 84132.

Accepted for publication September 20, 1988.

The contribution of SES to racial differences in cancer incidence can be tested directly, by seeing if adjustment for SES causes these differences to diminish. However, this has not often been done, because individual SES data (particularly on income) have been available only for relatively small studies.

The alternative to using SES data on individuals is to use aggregate data and apply these group characteristics to the individual cases. We have used a variation of this approach, utilizing SEER cancer incidence data from the years 1978 to 1982 and the SES variables available by census tract from the 1980 US Census. This report summarizes the results for the 12 most frequent cancer sites among blacks, with emphasis on the SES variable of poverty level.

Methods

Most of the nine registries in the SEER system provide information on the census tract of residence for each cancer case reported. In this analysis we utilized cases obtained from the SEER registries of Atlanta, Detroit, and San Francisco-Oakland, for the years 1978 to 1982. These registries were the only ones with large numbers of black cases coded for census tract of residence, and contained 79% of the black population in the SEER program. For all cancer sites, 156,393 cases were diagnosed in the three

registries for these 5 years. Approximately 16% of these cases occurred in blacks. We included in our analysis all cancer sites which had at least 500 black cases. Invasive cervical cancer was treated separately from *in situ*. The category of white race included whites with Spanish surname or origin, a group comprising 6.5% of the whites in these registry areas, but did not include Japanese, Chinese, American Indian or Alaskan native, Hawaiian, or "other" categories.

Population data were obtained from the 1980 US census; the adult population (age 20+ years) in these registries was approximately 5.8 million, 18% of whom were black. For each census tract in these registries, the population had been enumerated by race, sex, and 5-year age groups. Each census tract also had available, for blacks and whites separately, the following socioeconomic variables: (1) median years of education completed (adults), for males and females separately; (2) median family income; and (3) percent of persons below the designated poverty level. Poverty level takes into account household size as well as family income, and is designed to reflect economic status more accurately than family income alone. The Census Bureau suppresses SES data for very small population groups, to ensure that published data cannot be traced to individual households.

Race-specific tracts were designated by considering the black population of a census tract (with its SES values) separately from the white population of the same census tract (with different SES values). Since some census tracts consisted of persons of only one racial group, this resulted in a total of 1426 black tracts and 2068 white tracts. The average adult population (age 20+) was 781 for the black tracts and 2382 for the white tracts. Values for socioeconomic variables were then assigned to the cancer cases according to their race-specific tract of residence.

Approximately 10% of the SEER cases could not be assigned an SES level, because of missing information from either SEER files or census tract files. Likelihood of missing numerator data information was independent of race and sex, but was somewhat higher in the oldest age groups. Approximately 9% of the population did not have SES level data, because of missing or suppressed files. Older age was also the only predictor of missing denominator data; blacks did not have more missing data than whites. In any sex, race, or age group, the cases had approximately the same percentage of missing data as the population. For the 12 cancer sites selected, there were a total of approximately 20,000 black and 88,000 white cancer cases analyzed by SES level.

For each cancer site, five to seven age categories for adjustment were designated, after considering the age distribution of the cases. We used 10-year age groups ranging downward from 80+, 70 to 79, 60 to 69, etc. (providing each group contained at least 1% of the cases). The ex-

ception was *in situ* carcinoma of the cervix, which included the age groups 15 to 19 and 70+ years.

For each cancer site, the group of pooled black and white cases were divided into five equal SES quintiles for adjustment. These quintile cutpoints were then applied to the group of pooled black and white census tracts to get the race, age, sex, and SES-specific populations corresponding to the cases. Incidence rates could then be calculated for each of these specific populations.

Some cancers are more common in lower SES groups, whereas others are more common in upper SES groups; using a single set of SES quintile cutpoints across all cancer sites would have impaired the adjustment by seriously imbalancing the numbers of cases in the quintile groups for some cancer sites, resulting in attempted comparisons across relatively noninformative SES categories, as well as producing unstable variances of the rates. We therefore used quintile cutpoints specific for each cancer site for better racial comparisons. We used case-based (rather than population-based) SES quintiles for the same reasons; for several sites, we tested the alternative procedure of using population-based SES quintiles applied to the cases, but found negligible differences in the results. For three sites (rectum, bladder, uterine corpus) the distribution of cases by race and SES indicated that use of quintile cuts based on black cases rather than the pooled black and white cases would produce more stable variances of the rates, improving the racial comparisons; these figures are presented. There were no major differences in results between these two methods.

Valid adjustment for SES assumes consistency between blacks and whites in the relationship of SES to cancer incidence. To test this, graphs of age-adjusted black and white cancer incidence by SES groupings common to both races were produced, using direct age adjustment.¹⁰ In a second set of graphs (not presented here) each race was plotted over its separate race-specific range of SES quintiles. Although the races could not be directly compared on these graphs because of the different SES quintiles, we could check the shape and position of the curves over a wider range of SES levels, and thus check their compatibility with the curves which used race-common quintiles.

The primary method of adjustment for age and SES was Poisson regression,¹¹ which is particularly appropriate for modeling occurrences of rare events such as cancer. This procedure assumes that the number of cases occurring in each race-age-SES category follows the Poisson distribution, with the logarithm of the Poisson parameter being a linear function of race, age, and SES factors. The program CATMAX was obtained from G. Koch.¹² Binary dummy variables were used to model each race, age, and SES category separately. Males and females were considered separately within each site. (We checked this regression procedure against two alternate methods for selected

TABLE 1. Poverty Distribution, Population Age 20+ Years, by Race

Percent below poverty level	Blacks	Whites
	No.* (Percent)	No.* (Percent)
0-1	35,056 (3.6%)	313,132 (7.2%)
>1-4	9057 (0.9%)	1,801,779 (41.6%)
>4-8	65,489 (6.7%)	1,466,833 (33.9%)
>8-20	343,212 (35.3%)	659,399 (15.2%)
>20-100	519,396 (53.4%)	89,685 (2.1%)

* Total no. of persons in the race-specific tracts grouped by percent of the tract below poverty level.

sites: (1) logistic regression, using the SAS program PROC LOGIST,¹³ modeling the odds of being a case, weighted by numbers of persons in each category; and (2) log-linear regression, using the SAS program PROC GLM,¹³ modeling the log of the rates, weighted by number of cases in each category. The results were quite similar.)

Results

The races showed marked SES differences, as expected. As shown in Table 1, the percentage of blacks living below the poverty level was much higher than that of whites. (These poverty groupings were not the quintiles used in the analyses by cancer site.) Blacks also had lower median education levels and family incomes, although the racial differences were not as extreme as for poverty level.

We found that the three SES variables differed in the strength of their association with race and with cancer incidence. Accordingly, they also produced different results when used to adjust the black/white incidence ratios for SES. Median years of education produced the smallest adjustment, median family income was intermediate in strength, and percent below poverty resulted in the largest adjustment of the black/white cancer incidence ratios. For example, in the case of invasive cervix, adjustment for age plus education, age plus income, and age plus poverty

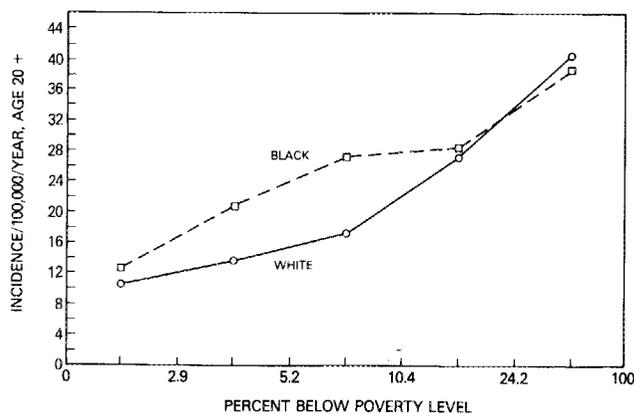


FIG. 1. Age-adjusted incidence of invasive cervical cancer versus poverty.

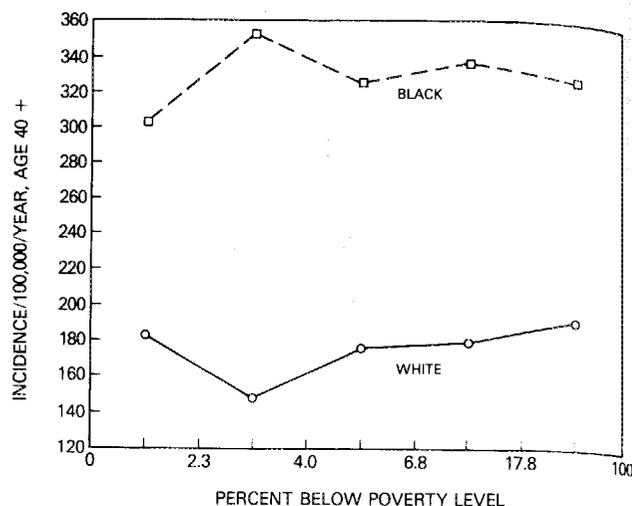


FIG. 2. Age-adjusted incidence of prostate cancer versus poverty.

produced black/white ratios of 1.66, 1.30, and 1.13, respectively. Since education and poverty were closely correlated variables, we found that the triple adjustment for age plus education plus poverty had no advantage over the double adjustment for age plus poverty; the corresponding ratios for invasive cervix were 1.17 and 1.13, respectively. Percent below poverty level was selected as the primary SES adjusting variable to present in this summary report.

For each cancer site, graphs of age-adjusted incidence rates by race-common SES quintile were used to compare the black and white SES gradients. Figure 1 demonstrates a site with strong poverty dependence, cervix. Figure 2 shows the example of a site with no appreciable poverty dependence, prostate. The curves varied from site to site, but for every site the curves for blacks and whites were essentially parallel, as required for valid SES adjustment. The only exception was for cancer of the rectum in males; in this one case, the incidence rates increased with poverty for whites but not for blacks. Therefore, the Poisson regression results for male rectal cancer are questionable.

The graphs (not presented here) which used race-specific SES quintiles rather than groupings common to both races were all consistent with the ones using race-common SES quintiles. This lends support to the belief that the results were not influenced by the methods of determining comparison groups.

Table 2 gives the results of the Poisson regression analyses, arranged in general order of decreasing strength of adjustment by poverty. Although there was a broad range of outcomes, the sites can be grouped into several categories: (1) those in which poverty explains all or nearly all the black/white difference; this group includes *in situ* cervix, invasive cervix, male lung, and female stomach; (2) those in which poverty explains much of the racial difference, breast, male stomach, esophagus, and pancreas;

TABLE 2. Effect of Adjustment for Poverty Level on Black/White Cancer Incidence: Poisson Regression Analysis

Site	Case (B/W)	Adjusted for	High risk	Ratio	95% CI
Cervix	750/1643	Age	Black	2.17	1.99-2.36
		Age + poverty		1.13	0.99-1.29
<i>In situ</i>	2214/5689	Age	Black	1.45	1.38-1.53
		Age + poverty		0.94	0.87-1.01
Lung	3268/12636	Age	Black	1.43	1.38-1.49
		Age + poverty		1.03	0.98-1.09
Male	1091/6271	Age	White	1.07	1.00-1.14
		Age + poverty		1.32	1.21-1.44
Female	2622/16489	Age	White	1.21	1.16-1.26
		Age + poverty		1.12	1.06-1.18
Stomach	564/1751	Age	Black	1.84	1.67-2.03
		Age + poverty		1.55	1.35-1.78
Male	279/1128	Age	Black	1.53	1.34-1.75
		Age + poverty		1.16	0.95-1.41
Female	508/690	Age	Black	4.07	3.63-4.56
		Age + poverty		2.32	1.94-2.78
Esophagus	178/366	Age	Black	2.77	2.32-3.32
		Age + poverty		1.77	1.33-2.36
Pancreas	424/1619	Age	Black	1.48	1.33-1.64
		Age + poverty		1.25	1.07-1.45
Male	374/1616	Age	Black	1.39	1.24-1.56
		Age + poverty		1.39	1.18-1.63
Female	991/5583	Age	Black	1.02	0.96-1.10
		Age + poverty		1.04	0.95-1.14
Colon	1244/6202	Age	Black	1.23	1.15-1.30
		Age + poverty		1.22	1.12-1.33
Rectum	373/2683	Age	White	1.27	1.14-1.41
		Age + poverty*		1.33	1.14-1.54
Male	316/2302	Age	White	1.23	1.09-1.39
		Age + poverty*		1.41	1.19-1.66
Female	458/4856	Age	White	1.98	1.80-2.18
		Age + poverty*		2.13	1.86-2.44
Uterine corpus	377/4206	Age	White	1.96	1.76-2.18
		Age + poverty*		1.93	1.68-2.21
Bladder	181/1574	Age	White	1.45	1.24-1.69
		Age + poverty*		1.70	1.37-2.11
Multiple myeloma	252/570	Age	Black	2.53	2.18-2.93
		Age + poverty		2.24	1.79-2.81
Male	226/588	Age	Black	2.30	1.97-2.68
		Age + poverty		2.37	1.86-3.02
Female	2956/9556	Age	Black	1.82	1.75-1.90
		Age + poverty		1.83	1.72-1.94

B: black; W: white.

Adjusted by Poisson regression for age (10-year age groups) and percent of race-specific census tract of residence below poverty level. Ordered

by approximate strength of adjustment by poverty.

* Poverty quintiles based on black cases. All others based on black + white cases.

and (3) those in which poverty explains little or none of a large racial difference, prostate, uterine corpus, myeloma, and bladder.

Colon and rectum rate ratios were not significantly affected by adjustment, but the initial ratios were not large. In female lung cancer a black/white difference was seen only after we adjusted for poverty; both black and white

rates rose with poverty, but whites had higher rates in each poverty quintile.

Discussion

We found that socioeconomic status, when measured by race-specific census tract poverty level, was responsible

for the black/white cancer incidence differential in varying degree, depending on cancer site. Poverty was responsible for all or much of the black/white difference for many of the sites studied. This supports the hypothesis that SES is a major, perhaps predominant cause of the racial differences. For several sites, however, large racial differences were independent of poverty.

The major strengths of this study are as follows: (1) the large community-based populations studied, (2) the standardized methods of cancer case detection and coding used in the SEER registry system, (3) the standardized SES measures used in the US census, and (4) the large number of cancer sites analyzed by the same methods.

There are several potential limitations of this study: (1) Individual SES data were not used. In this study, a person's SES was defined on the basis of the income or educational levels of that person's race-specific census tract of residence. Such aggregate variables can be interpreted at two levels: they may be considered as approximations to SES variables obtained on individuals; however, they may also capture some exposures that operate primarily at the census tract or "neighborhood" level, such as neighborhood peer-group health behaviors, access to health care, or possible local environmental pollution. Our findings indicate that this type of variable, whether representing individual or group phenomena, can capture some or all of the exposures responsible for racial differences in cancer incidence, since for some sites, all the racial differences were accounted for by poverty level. (2) A potential limitation in the adjustment for poverty was the relatively limited overlap in the black and white poverty ranges. We attempted to minimize this problem by choosing the most appropriate poverty groupings for comparing the races, and by verifying that these results were compatible with the poverty gradients seen over the separate black and white poverty ranges. (3) This study was based on cancer cases in three large cities, and may not be representative of the entire US population. However, these registries contained 79% of the black population in SEER, and our age-specific incidence rates and age-adjusted black/white ratios matched closely the ones based on the entire program.¹

We compared our results with those of Devesa, who did an earlier, somewhat comparable study¹⁴ which used different data sources (Third National Cancer Survey and 1970 US census), different methods of adjustment, and different SES indicators. Results have been published only for lung¹⁵ and for breast and cervix.¹⁶ The racial differences we obtained after adjusting for age (only) were generally larger than in the analysis 10 years previously. Some of this difference may be methodologic; however, it may reflect a widening of the racial gap for some sites. The two studies were generally similar in the direction of SES adjustment.

We found that for the majority of sites tested (invasive and *in situ* cervix, male lung, stomach, esophagus, pan-

creas, breast), poverty accounted for much or all of the racial differences. These findings are compatible with the hypothesis that racial differences in incidence for these sites largely reflect SES-dependent exposures to carcinogens or cofactors, such as smoking, occupation, child-bearing patterns, or others.

For four sites, however (uterine corpus, bladder, myeloma, prostate), large racial differences were found to persist after adjustment for poverty level. We found insignificant poverty gradients in these sites, in both races and both sexes. Several explanations for these findings are possible:

1. The apparent racial differences in incidence rates could be due to under-ascertainment of cases in a particular race. However, this explanation is hard to reconcile with the results found for different sites, such as a white excess in bladder cancer, but a black excess in prostate cancer.
2. There may actually be SES dependence for these sites, but our variables may not have captured it. Our findings are consistent with most prior studies, which have shown little or no association between SES and bladder cancer,¹⁷ myeloma,¹⁸ or prostate cancer,¹⁹ and weak or inconsistent associations between SES and cancer of the uterine corpus.^{20,21} However, the pertinent exposure factors for some cancers may have occurred so long ago that any current SES measures might not adequately reflect those past relationships. In this case, special studies would be required to discover such an effect. Moreover, less commonly used measures of SES, such as occupation, might better reflect pertinent exposures for these sites.
3. The racial differences for these sites may be due to genetic factors. This explanation should be invoked with much caution²²⁻²⁴; not for lack of an alternative explanation, but only on the basis of positive evidence (*i.e.*, racial differences in the frequency of cancer-determining genes). We are unaware of any such evidence of genetic factors in the racial differences for these sites.
4. Finally, these differences may be due to environmental (as opposed to genetic) exposures, specifically those that vary with race but are unrelated to SES. Classification by race is a complex, primarily socially derived phenomenon: a designation based on skin color but carrying ethnic, historical, and other implications as well as economic ones.²²⁻²⁴ Some environmental factors (cancer-related behaviors, for instance) may not vary with income or education, but may still be more common within certain racial groupings. Attempts to find factors of this type may be a particularly fruitful approach to the etiology of these cancers.

In summary, this study supports the usefulness of race-specific census tract data in investigating racial differences in cancer incidence. For the majority of sites studied, the racial differences were found to be largely or completely due to poverty. For several sites, large differences persisted

after adjustments for cancers.

1. Baquet, Other Minc 2785. Beth
2. US Bu istics of the US Govern
3. Nation the Nation United Stat
4. Unite Attitudes, 1966. Rock 1969.
5. Willia and alcohol study from 58:525-547
6. Blot V United Sta miology an
7. Buell of cancer r
8. Kitag States: A st vard Unive
9. Logar 1851-197 Research o
10. Arm England: B

after adjustment for poverty; these results suggest directions for future investigations into the etiology of these cancers.

REFERENCES

1. Baquet CR, Ringen K, Pollack ES *et al.* Cancer Among Blacks and Other Minorities: Statistical Profiles, DHEW publication no. (NIH) 86-2785. Bethesda, MD: National Institutes of Health, 1986.
2. US Bureau of Census. General, Social, and Economic Characteristics of the United States, 1980, series PC80-1-C1. Washington, DC: US Government Printing Office, 1982.
3. National Center for Health Statistics. Highlights From Wave I of the National Survey of Personal Health Practices and Consequences, United States, 1979, Vital and Health Statistics series 15, no. 1. Hyattsville, MD: National Center for Health Statistics, 1980.
4. United States Public Health Service. Use of Tobacco: Practices, Attitudes, Knowledge, and Beliefs, United States, Fall 1964 and Spring 1966. Rockville, MD: National Clearinghouse for Smoking and Health, 1969.
5. Williams RR, Horm JW. Association of cancer sites with tobacco and alcohol consumption and socioeconomic status of patients: Interview study from the Third National Cancer Survey. *J Natl Cancer Inst* 1977; 58:525-547.
6. Blot WJ, Fraumeni JF. Geographic epidemiology of cancer in the United States. In: Schottenfeld D, Fraumeni JF, eds. *Cancer Epidemiology and Prevention*. Philadelphia: WB Saunders, 1982; 179-183.
7. Buell P, Dunn JE Jr, Breslow L. The occupational-social class risks of cancer mortality in men. *J Chronic Dis* 1960; 12:600-621.
8. Kitagawa EM, Hauser PM. Differential Mortality in the United States: A study in Socio-Economic Epidemiology. Cambridge, MA: Harvard University Press, 1973.
9. Logan WPD. Cancer Mortality by Occupation and Social Class 1851-1971, IARC publication no. 36. Lyon: International Agency for Research on Cancer, 1982.
10. Armitage P. *Statistical Methods in Medical Research*. Oxford, England: Blackwell Scientific Publications, 1971; 388-391.
11. Frome EL, Checkoway H. Use of Poisson regression models in estimating incidence rates and ratios. *Am J Epidemiol* 1985; 121:309-323.
12. Stokes ME, Koch GC. A macro for maximum likelihood fitting of log-linear models to Poisson and multinomial counts with contrast matrix capability for hypothesis testing. Proceedings of the Eighth Annual SAS User's Group International Conference, 1983; 795-800.
13. SAS Institute, Inc. SUGI Supplemental Library User's Guide, 1983 ed. Cary, NC: SAS Institute, 1983.
14. Devesa SS. A study of the association of cancer incidence with income and education among whites and blacks (PhD thesis). Baltimore: Johns Hopkins University, 1979.
15. Devesa SS, Diamond EL. Socioeconomic and racial differences in lung cancer incidence. *Am J Epidemiol* 1983; 118:818-831.
16. Devesa SS, Diamond EL. Association of breast cancer and cervical cancer incidences with income and education among blacks and whites. *J Natl Cancer Inst* 1980; 65:515-528.
17. Morrison AS, Cole P. Urinary Tract. In: Schottenfeld D, Fraumeni JF, eds. *Cancer Epidemiology and Prevention*. Philadelphia: WB Saunders, 1982; 925-937.
18. Johnston JM, Grufferman S, Bourguet CC, Delzell E, DeLong ER, Cohen HJ. Socioeconomic status and risk of multiple myeloma. *J Epidemiol Commun Health* 1985; 39:175-178.
19. Ernster VL, Selvin S, Sacks ST, Austin DF, Brown SM, Winkelstein W. Prostatic cancer: Mortality and incidence rates by race and social class. *Am J Epidemiol* 1978; 107:311-320.
20. Elwood JM, Cole P, Rothman KJ, Kaplan SD. Epidemiology of endometrial cancer. *J Natl Cancer Inst* 1977; 59:1055-1060.
21. de Waard F. Uterine corpus. In: Schottenfeld D, Fraumeni JF, eds. *Cancer Epidemiology and Prevention*. Philadelphia: WB Saunders, 1982; 901-908.
22. Cooper R. A note on the biologic concept of race and its application in epidemiologic research. *Am Heart J* 1984; 108:715-723.
23. Montagu A. *The Concept of Race*. Toronto: Collier Macmillan, 1964.
24. Gould SJ. *The Mismeasure of Man*. New York: WW Norton, 1981.