Disparities in Prevalence Rates for Lung, Colorectal, Breast, and Prostate Cancers in Medicaid

C. Daniel Mullins, PhD; Jesse L. Cooke, Jr; Junling Wang, MS; Fadia T. Shaya, PhD, MPH; Van Doren Hsu, PharmD; and Sandra Brooks, MD

Baltimore, Maryland

Background: Given previous reports of variations in prevalence of cancer in low-income individuals, we sought to determine if disparities in cancer prevalence existed in a similarly-insured Medicaid population.

Methods: Using Maryland Medicaid administrative claims data, prevalence rates of lung, colorectal, breast, and prostate cancers were calculated for Maryland Medicaid recipients who were continuously eligible during the period from January 1, 2000 to December 31, 2000. Chi-squared tests were used to test the differences across subgroups. Cancer prevalence data were age-adjusted using Maryland Medicaid enrollees as the standard population.

Results: The care prevalence rates for lung, colorectal, breast, and prostate cancers were 75/10,000, 63/10,000, 92/10,000, and 45/10,000, respectively. These rates were 1.2 to 5.2 times those reported at the national level. Generally, higher cancer prevalence rates in certain racial groups in Maryland Medicaid were consistent with previous studies. Regional differences in cancer prevalence existed for each cancer studied.

Conclusions: Limiting our study sample to a population of uniformly low socioeconomic individuals did not eliminate the disparity in prevalence rates between blacks and whites. Different patterns of racial disparity across regions reported by previous researchers might be due to small area variation in addition to socioeconomic status.

Key words: prevalence ■ lung neoplasms ■ colorectal neoplasms ■ breast neoplasms ■ prostatic neoplasms

 From the Center on Drugs and Public Policy (Mullins; Cooke Jr, doctoral candidate; Wang, doctoral candidate; Shaya); the Pharmaceutical Research Computing at the University of Maryland School of Pharmacy (Hsu); and the Department of Obstetrics and Gynecology and Reproductive Sciences, University of Maryland School of Medicine (Brooks), Baltimore, MD. Send correspondence and reprint requests for *J Natl Med Assoc.* 2004; 96:809–816 to: C. Daniel Mullins, Center on Drugs and Public Policy, University of Maryland School of Pharmacy, 515 W. Lombard St., Second Floor, Baltimore, MD 21201; phone: (410) 706-0879; fax: (410) 706-5394; e-mail: dmullins@rx.umaryland.edu

BACKGROUND

Cancers of the lung, colon, breast, and prostate accounted for 54% of all 1998 cancer deaths in Maryland,[1] compared with 52.7% of all 1998 cancer deaths in the United States.[2] Disparities in cancer stage, treatment intensity, and mortality between blacks and whites are well-documented; however, there is a paucity of data examining racial disparities in similarly insured, low-income populations.[3-5] Previous authors have postulated that low socioeconomic status is a stronger predictor of cancer outcome than race. A recent study by Sung et al. showed that in rural Georgia, differences between whites and blacks in prevalence rates for cervical carcinoma mostly disappeared in the Medicaid population, a population of homogenous economic status.[6] However, the disparities persisted among Medicaid enrollees in urban Georgia. We examined prevalence rates between whites and blacks for cancers of the lung, colorectum, breast, and prostate in a population of Maryland Medicaid recipients, an economically homogenous group. We also explored the potential reasons for differing patterns of racial disparities across regions.

If the disparities in prevalence rates between whites and blacks can be attributed to socioeconomic status, then disparities between whites and blacks should be largely eliminated after adjusting for socioeconomic status. Thus, our hypothesis is that disparities between whites and blacks should be largely eliminated in Maryland Medicaid, a homogenous group in terms of socioeconomic status.

METHODS

Study Population and Data Sources

This study utilized an historical cross-sectional study design. The data source for this study was Maryland Medicaid administrative claims data (including demographic, eligibility, managed care organization (MCO) enrollment data, medical, and

institutional fee-for-service claims) and MCO encounter data. In accordance with patient confidentiality concerns, this study was approved by the State of Maryland (Protocol # 01-16). It has also been reviewed and deemed to be exempt by the Institutional Review Board of the University of Maryland (Exemption No. CDM-040101).

To be included in our analysis, individuals needed to be Maryland Medicaid recipients 18 and older, with encounters, medical or institutional claims based on International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9 CM) diagnosis codes; prescription drug National Drug Codes (NDCs) for chemotherapy drugs, tamoxifen, analgesics, hematopoietics and Xeloda; or current procedure technology (CPT) codes for lung, colorectal, breast, or prostate cancers. Furthermore, they must have been continuously eligible for Medicaid between January 1, 2000 and December 31, 2000 to be included in the study cohort. See Figure 1 for the ICD-9CM Diagnosis Codes that were used to identify the cancers of interest. Demographic and enrollment information was extracted from each source file.

Analyses

Frequencies and crosstabulations were performed on all data to validate the completeness and integrity of the data. Algorithms were developed to evaluate claims for adjustment and duplications. Validation of these algorithms was conducted by reviewing raw claims for randomly selected recipients. The resultant data were unique with no duplication.

The prevalence for each cancer was calculated. The prevalence rates reflect the period from January 1, 2000 to December 31, 2000. The number of eligible Maryland Medicaid recipients 18 and older on January 1, 2000 was used as the denominator. Prevalence was calculated by race, region, age, and gender.

Figure 1. ICD-9CM Diagnosis Codes for Selected Cancers[a]

[a] A three-digit code followed by x (e.g. 174.x) indicates that all codes starting with the three-digit code before the "." are included.

Lung Cancer	Primary: 161.3, 161.8, 161.9, 162.x Secondary: 196.x, 197.x, 198.x In situ: 231.x
Colorectal Cancer	Primary: 153.x, 154.x Secondary: 196.x, 197.x, 198.x In situ: 230.3, 230.4, 230.5, 230.6, 230.7
Breast Cancer	Primary: 174.x, 175.x Secondary: 196.x, 197.x, 198.x In situ: 233.0
Prostate Cancer	Primary: 185, 189.3 Secondary: 196.x, 197.x, 198.x In situ: 233.4, 233.9

Table 1. Demographic Characteristics of Maryland Medicaid Enrollees (N=246,430)

Characteristic		Number of People	Percentage (%)[a]
Age	18–64	191,104	77.55
	65+	55,326	22.45
Gender	Female	186,636	75.74
	Male	59,794	24.26
Race	Black	120,577	48.93
	White	101,517	41.20
	Other	24,336	9.88
Regions	Urban	80,159	32.53
	Suburban	132,980	53.96
	Rural	32,914	13.36
	Unspecified[b]	377	0.15

[a] Percentages do not always add up to 100 because of rounding.
[b] Unknown or out-of-state on January 1, 2000.

Prevalence rates across regions were estimated to explore reasons behind the different patterns of racial disparities across regions reported by previous researchers. Prevalence rates across age and gender groups were calculated so that they could be compared with estimates of previous studies at the national level so that our study results could be validated.

There were three racial groups: black, white, and other. The racial group "other" was comprised of Hispanics, Asians, Native Americans, Pacific Islanders/Alaskans, and those of unknown ethnicity/race. Since each of these "other" racial groups individually accounted for less than 4% of the total Maryland Medicaid population, we decided that it was not appropriate to calculate prevalence rates for each individual group (combined, they account for less than 10% of the study population). Thus, only the differences between whites and blacks were analyzed in this study.

We defined geographic region as urban (Baltimore city); rural (Allegany, Garrett, Washington, Kent, Queen Anne's, Caroline, Talbot, Dorchester, Somerset, Wicomico, and Worchester counties); and suburban (the rest of Maryland) based on the proportion of agricultural populations in the total population in the regions. Each person was categorized to a geographic region (urban, suburban, rural) according to his/her county of residence on January 1, 2000. The patients were categorized into two age groups, those under 65 and 65 and older. The differences in subgroups were tested using Chi-squared tests.

Cancer prevalence data were age-adjusted using the direct standardization method. This was done by multiplying the age-specific rates in the target population by the age distribution of the standard population.[7] Maryland Medicaid enrollees were used as the standard population.

RESULTS

On January 1, 2000, Maryland Medicaid had 246,430 enrollees, with demographic characteristics as reported in Table 1. Most enrollees were under the age of 65 (77.55%). More females (75.74%) than males (24.26%) were Maryland Medicaid beneficiaries. Whites constituted 41.20% of the total Medicaid population, while blacks represented a slightly larger share of the Maryland Medicaid population (48.93%). Less than 10% (9.88%) of the total population represented people of other racial groups. More than half of the beneficiaries (53.96%) lived in suburban areas in Maryland. The next largest group of beneficiaries was urban, which represents Baltimore city. Enrollees from rural areas represented less than 14% of the total Medicaid population.

In our population, there were 1,836 lung cancer patients, 1,558 colorectal cancer patients, 2,255 breast cancer patients, and 1,098 prostate cancer patients. Since the number of Maryland Medicaid enrollees on January 1, 2000 was 246,430, the lung cancer prevalence rate was 75/10,000. The prevalence rates were 64/10,000 for colorectal cancer, 92/10,000 for breast cancer, and 45/10,000 for prostate cancer.

We reported the cancer prevalence rates across racial groups both before and after age adjustment in Table 2. Age-adjusted rates are being presented for the purposes of making comparisons in the rates of age-related health events (in this case, cancer). Age-adjusted rates are essential for events that vary with age (e.g., cancer deaths), when comparing populations with *different* age distributions. Age-adjusted rates should be used only for the purpose of comparison. Because an age-adjusted rate is based on an external standard population, it does not reflect the absolute frequency of the event in a population; therefore no significance tests are presented.

These rates varied across cancers and across racial groups. For lung cancer, whites had a higher prevalence rate than blacks, and the difference between them was statistically significant ($p < 0.0001$). After age adjustment, the prevalence rate for whites remained higher than blacks. For colorectal cancer, the prevalence rate in blacks was lower than in whites and the difference was significant ($p=0.0384$). After age adjustment, the prevalence rate for blacks was 1.1 times the rate for whites, in contrast to the comparison before age adjustment.

Table 2. Cancer Cases and Prevalence Rates (per 10,000) among Maryland Medicaid Enrollees across Races in 2000[a]

	Lung Cancer[b]			Colorectal Cancer[c]			Breast Cancer[d]			Prostate Cancer[e]		
	Number	Prev.	AP	Number	Prev.	AP	Number	Prev.	AP	Number	Prev.	AP
Black	810	67.18	103.19	724	60.04	84.94	1,092	90.56	116.04	616	51.09	71.31
White	872	85.90	116.38	681	67.08	76.15	994	97.91	104.07	361	35.56	37.93

[a] Prev. = unadjusted prevalence rate; AP = age-adjusted prevalence rate. Significance tests are for unadjusted prevalence rates; [b] $\chi^2 = 25.3070$, $p<0.0001$; [c] $\chi^2 = 4.2877$, $p=0.0384$; [d] $\chi^2 = 3.1404$, $p=0.0764$; [e] $\chi^2 = 30.0776$, $p<0.0001$

For breast cancer, before age adjustment, the prevalence rate for blacks was 0.9 times the rate for whites but the difference was not statistically significant (p=0.0764). After age adjustment, the prevalence rate was higher for blacks. As reported in Table 2, blacks had a higher prevalence rate for prostate cancer than whites both before and after age adjustment. The racial difference for prostate cancer before age adjustment was statistically significant.

Cancer prevalence rates across regions are reported in Table 3. The unspecified group was excluded from the analysis since the number of people was small, which makes calculation of the prevalence rate inappropriate. For lung, colorectal, and breast cancers, the unadjusted prevalence rates in urban and rural regions were closer than after age-adjustment and were both higher than the rates in suburban area. However, these geographic differences were statistically significant at $p<0.05$ only for lung cancer and colorectal cancer. For prostate cancer, the unadjusted prevalence rates for the three regions were similar. After age adjustment, the urban region had higher rates than the other two regions for all four cancers.

Table 4 presents the differences across age groups for the four cancers. As expected, the older group consistently had a higher level of prevalence rates. All these differences were statistically significant.

Table 5 includes the comparison between genders. Males had higher prevalence rates than females for two of the three cancers: lung and colorectal cancers. Females had higher rates for breast cancer compared with males. The differences across genders were significant for all three cancers. The trend was unchanged after age adjustment.

DISCUSSION

In this study, we determined that the prevalence rates for lung, colorectal, breast, and prostate cancers among Maryland Medicaid enrollees were 75/10,000, 63/10,000, 92/10,000, and 45/10,000, respectively. Prevalence rates are measures of the burden of diseases in a community for the purpose of setting public policy and allocating resources.[8] Cancer generally requires treatment over a period of time and, thus, is suitable for care prevalence studies. Cancer incidence calculations provide only a part of the epidemiology and disease burden story. Incidence rates do not capture the complex features of the healthcare required for cancer patients.

Using claims data to estimate prevalence rates is one of many ways to estimate cancer prevalence rates. The traditional method of estimating cancer prevalence rates is based on the Connecticut Tumor Registry model, which has recorded patients diagnosed since 1935. Using the number of survivors of cancer who had been diagnosed during a certain period, and mortality rates, researchers can directly compute a prevalence rate. When applied to the total U.S. population for selected years, a national estimate of cancer prevalence could be obtained.[9] The accuracy of estimation can be compromised; however, since patients can be lost to follow-up. In addition, the ascertainment of nonfatal cancers in the earlier years of the registry can be incomplete.[9] Another inaccuracy occurs if people diagnosed before a registry began are still alive at the reference time when the prevalence rate is estimated.[9] These limitations, if not offset by the inclusion of people who no longer reside in the registry area, would lead to underestimation of the prevalence rate.[9] One study by Capocaccia and DeAngelis confirmed the incompleteness of prevalence based on cancer registry data alone.[10]

Two other sources exist for estimating prevalence rates. The Surveillance, Epidemiology, and End Results (SEER) Program, an authoritative source of information from the National Cancer Institute (NCI) on cancer incidence and survival at the national level,[11] is also a cancer registry program. It has been used to estimate cancer prevalence rates.[12] The National Health Interview Survey (NHIS) is an alternative source for obtaining prevalence estimates. It has been used to estimate the prevalence of cancer among adults in the United States.[13] The NHIS collects data from a nationwide probability sample of noninstitutionalized people. Its information on chronic conditions is obtained from patients'

Table 3. Cancer Cases and Prevalence Rates (per 10,000) across Regions among Maryland Medicaid Enrollees in 2000[a]

	Lung Cancer[b]			Colorectal Cancer[c]			Breast Cancer[d]			Prostate Cancer[e]		
	Number	Prev.	AP	Number	Prev.	AP	Number	Prev.	AP	Number	Prev.	AP
Urban	681	84.96	123.35	556	69.36	91.34	783	97.68	119.10	358	44.66	57.87
Suburban	856	64.37	89.21	789	59.33	70.82	1,162	87.38	96.89	595	44.74	49.61
Rural	293	89.02	118.54	210	63.80	72.79	307	93.27	100.74	145	44.05	47.02

[a] Prev. = unadjusted prevalence rate. AP = age-adjusted prevalence rate. Significance tests are difference for unadjusted prevalence rates across three regions; [b] χ^2 =39.1541, p<0.0001; [c] χ^2 = 7.9311, p=0.0190; [d] χ^2 = 5.8685, p=0.0532; [e] χ^2 = 0.0283, p=0.9859

self-reports, which is considered to be less accurate than the cancer registry. Both under- and over-reporting are possible.[14]

We used an alternative to the aforementioned means for estimating cancer prevalence rates by measuring "care prevalence," which is "an estimate of prevalent cases that are still under care."[15] The NCI is carrying out a project estimating care prevalence for colorectal cancer using SEER-Medicare data.[15] The care prevalence is a valid prevalence measure though it has not been widely used. It capitalizes on criterion and construct validity—that is, the ability to predict events. Furthermore, care prevalence presents a good case for external validity, or generalizability. Estimating care prevalence is especially important when we are interested in estimating burden of cancer in a program such as Medicare or Medicaid. Other measures of prevalence attempt to include all patients (e.g., those in remission or "cured;" or those still under care) for a measure of total prevalence; however, insurers may only be concerned with the number of cancer patients within their plan and who are currently under some type of care. Other studies have used Medicare and Medicaid claims data to determine cancer prevalence rates.[6,16] Using claims data to estimate care prevalence rates also avoids the difficulties in differentiating prevalent cases and incident cases that previous researchers had when they tried to use claims data to identify cancer incident cases.[17]

One must be cautious when comparing our estimates of the cancer prevalence rates in Maryland Medicaid to estimates of the prevalence rates at the national level in previous studies. Cancer prevalence rate estimates can vary for multiple reasons.[18] Prevalence rates depend on the disease parameters, population characteristics, medical services, and the means of data collection. Variation in prevalence might be due to differences in incidences of the diseases and/or the mortality rates of the population selected for the analysis. The difference in prevalence can also result from the extent of screening programs, the stage distribution at diagnosis, and access to treatment facilities. In addition, prevalence rates can vary according to differences in coding conventions and data quality—namely, the follow-up rates from survey, the proportion of patients identified only from death certificates, and the number of patients who migrated in and out of the population.[18] Medicaid data differ from cancer registry and NHIS in several aspects. Medicaid enrollees are of uniformly lower socioeconomic status. In addition, some Medicaid-eligible enrollees are only enrolled in the program when a health problem is discovered. Others are enrolled only when they "spend down" their assets and become eligible for Medicaid.[6] However, the cancer registry and NHIS included patients of various socioeconomic status. Moreover, the administrative database collects data following different conventions than those of the cancer registry and NHIS.

We need to consider a reference framework despite the caveats of comparing our estimates of cancer prevalence rates among Maryland Medicaid population with those from previous studies at the national level. We have compared our estimates of overall cancer-specific prevalence rates in the Maryland Medicaid populations with those in the national population. Table 6 presents the comparison between our estimates and those based on the 1987 and 1992 NHIS.[13,14] Maryland Medicaid data generally indicated higher prevalence rates (1.2 to 5.2 times as high as those at the national level) for all four of these cancers, compared with the national level. The only exception was the estimate of breast cancer based on the 1987 NHIS, which was higher than our estimate. No previous prevalence studies on these four cancers in other Medicaid population are available for comparison to our estimates in Maryland. One anecdotal argument for high cancer prevalence rates in the Medicaid population is that some individuals become Medicaid-eligible because of their cancer diagnosis and the related physical and economic impact.[6] However, because of the significantly higher cancer prevalence rates in this population, further research should be done to determine whether our findings can be replicated in another Medicaid population, and, if so, to identify the reasons for much higher rates for cancer in the Medicaid population. One potential contributing factor to the observed higher prevalence rates relates to whether cancer costs would result in

Table 4. Cancer Cases and Unadjusted Prevalence Rates (per 10,000) across Age Groups among Maryland Medicaid Enrollees in 2000

	Lung Cancer[a]		Colorectal Cancer[b]		Breast Cancer[c]		Prostate Cancer[d]	
	Number	Prevalence	Number	Prevalence	Number	Prevalence	Number	Prevalence
18–64	1,001	52.38	595	31.13	1,038	54.32	284	14.86
65+	835	150.92	963	174.06	1,217	219.97	814	147.13

[a] $\chi^2 = 552.1205$, $p<0.0001$; [b] $\chi^2 = 1366.8999$, $p<0.0001$; [c] $\chi^2 = 1263.5857$, $p<0.0001$; [d] $\chi^2 = 1665.1059$, $p<0.0001$

individuals becoming impoverished to the extent that they qualify for Medicaid. A post hoc analysis of the Maryland Medicaid eligibility file does not support this hypothesis. In the overall Medicaid population, 9.48% qualified through "spend down"; in the cancer population, 8.17% qualified through "spend down."

The comparison across racial groups for each cancer was generally consistent with previous studies. Previous studies have found racial disparities between whites and blacks in the prevalence rates for lung cancer, colorectal cancer, breast cancer, and prostate cancer at the national level. For lung cancer, estimates based on SEER Program and Connecticut Cancer Registry among men reported that whites generally had lower rates for lung/bronchus cancer than blacks. A more complicated trend was reported among women. In older age groups, white women had higher prevalent rates for lung cancer than black women, according to both the Connecticut Cancer Registry and SEER Program. An opposite trend was identified in younger age groups.[12] Our findings of higher prevalence rates among whites do not contradict results from the SEER Program and Connecticut Cancer Registry, which demonstrated a higher prevalence rate *overall* among whites than blacks.[12] Site-specific rates were also higher in whites, with the exception of colorectal cancer, which was higher for black males through ages 59 and black females through age 64.

A study by Byrne using NHIS data reported that whites had higher rates for colorectal cancer than blacks.[13] Estimates based on the SEER program and Connecticut Cancer Registry reported that whites had higher prevalence rates for colorectal cancer than did blacks in older age groups; however, the rate in whites was lower than blacks in younger age groups.[12] A similar picture exists for breast cancer. Estimates by Byrne et al. reported that the rates for breast cancer were higher among whites than blacks.[13] For female breast cancer, estimates based on the Connecticut Cancer Registry and SEER program both reported that whites had higher prevalence rates than blacks in older age groups, but blacks had higher prevalence rates in younger age groups.[12] For colorectal cancer and breast cancer, our report of higher prevalence rates among whites than among blacks before age adjustment is consistent with reports in NHIS[13] and the estimates based on the Connecticut Cancer Registry and SEER program.[12] After age adjustment, our results showed higher prevalence rates in blacks than whites for both cancers.

For prostate cancer, both the study by Byrne[13] and estimates based on SEER and the Connecticut Cancer Registry[12] reported that blacks had higher prevalence rates than whites. We found higher rates in blacks than in whites both before and after age

Table 5. Cancer Cases and Prevalence Rates (per 10,000) across Genders among Maryland Medicaid Enrollees in 2000[a]

	Lung Cancer[b]			Colorectal Cancer[c]			Breast Cancer[d]		
	Number	Prev.	AP	Number	Prev.	AP	Number	Prev.	AP
Male	838	140.15	203.09	539	90.14	111.67	50	8.36	9.81
Female	998	53.47	72.85	1019	54.60	66.45	2205	118.14	135.42

[a] Prev. = Unadjusted prevalence rate. AP = age-adjusted prevalence rate. Significance tests are for unadjusted prevalence rates; [b] $\chi^2 = 451.2623$, $p<0.0001$; [c] $\chi^2 = 89.7650$, $p<0.0001$; [d] $\chi^2 = 594.3795$, $p<0.0001$

Table 6. Estimates of Cancer Prevalence Rates in Maryland Medicaid versus Rates Based on the National Health Interview Survey (per 10,000)

Cancers	Our Estimates	Estimates from 1987 NHIS[a]	Estimates from 1992 NHIS[b]
Lung	74.50	14.2[c]	17
Colorectal	63.22	34.3	35.5
Breast	91.51	133.2[d]	79.3[d]
Prostate	44.56	32.4	37.1[e]

[a] Byrne J, Kessler L, Devesa SS. The prevalence of cancer among adults in the United States: 1987. *Cancer.* 1992;69:2154-2159.
[b] Hewitt M, Breen N, Devesa S. Cancer prevalence and survivorship issues: analyses of the 1992 National Health Interview Survey. *J Natl Cancer Inst.* 1999;91:1480-1486.
[c] Includes lung/larynx cancers.
[d] Female breast cancer only.
[e] Includes prostate, testes, and other male genital organ cancers.

adjustment, consistent both with estimates based on NHIS[13] and estimates based on the Connecticut Cancer Registry and SEER program.[12]

Our study population is relatively homogeneous in socioeconomic status. However, most of the differences in cancer prevalence between whites and blacks are still significant. Thus, our study results do not support the hypothesis that when we control for socioeconomic status, the disparities between racial groups are fully eliminated. Although Medicaid eligibility is not a perfect measure for socioeconomic status, our study results suggest that factors in addition to socioeconomic differences contribute to racial disparities in cancer.

The study by Sung and colleagues on Georgia Medicaid cancer patients showed that differences in cancer prevalence rates between whites and blacks existed in metropolitan areas.[6] They offered a reasonable argument for higher rates in blacks than in whites—that is, blacks were still somehow in a more disadvantaged position even though blacks and whites were all in Medicaid program and were all financially distressed.[6] However, their argument cannot explain the higher prevalence rates for certain cancers among whites in our study. Thus, further studies are warranted to document the reasons for racial disparity in Medicaid populations.

Our study results also provide insight for the apparently conflicting racial disparity patterns in urban and rural regions reported by Sung and colleagues.[6] We found differences in prevalence rates for cancers in urban, rural, and suburban areas. This leads us to believe that small area variation plays a role in explaining different racial disparity patterns in urban and rural areas. Previous studies on small area variation reported that small area variation and socioeconomic status simultaneously help to explain patterns of health services utilization.[19] In the case of the study by Sung et al.,[6] there is still small area variation, although they studied a reasonably homogeneous socioeconomic group.

The variation across age and gender groups in our study was consistent with that reported by previous studies. We reported that people 65 and older compared with those under 65 consistently had higher rates for each of these four cancers. This pattern was also identified in report based on NHIS.[13] We reported that males had higher rates for lung cancer and colorectal cancer than females both before and after age adjustment. These estimates were consistent with estimates based on NHIS[13] and the SEER Program and Connecticut Cancer Registry.[12]

Limitations

Although the method of estimating cancer prevalence rates based on claims data does not have the drawbacks of the cancer registry and NHIS data, it has its own limitation, which is related to validation of claims data. Previous studies reported that there were discrepancies between claims and medical record data.[20,21] In our study, validation of the medical and prescription claims data with primary data from the medical record is not possible due to limited resources. This may result in false positives or negatives due to incorrect data entry or code assignment. However, validation of claims data has not become routine in previous studies estimating prevalence rates using claims data.[6,16]

The second limitation is that we may underestimate the prevalence of the cancers because we based our prevalence estimates solely on outpatient medical claims. If there were patients who received only inpatient services, they were not included in the prevalence estimates. However, we feel that the likelihood of a cancer patient receiving only inpatient services with no prior or subsequent outpatient services or physician's visits is extremely low.

CONCLUSION

This study found much higher prevalence rates for lung, colorectal, breast, and prostate cancers among Maryland Medicaid population compared with estimates at the national level. Further studies are warranted to investigate the reasons underlying the much higher cancer prevalence rates overall among the Maryland Medicaid population.

Limiting our study sample to a population of uniformly low socioeconomic status did not eliminate the disparity in prevalence rates between blacks and whites, but did reduce the magnitude of the disparity of colorectal and breast prevalence rates. We found geographic differences in cancer prevalence rates for the four cancers under study. Different patterns of racial disparity across regions reported by previous researchers might be due to small area variation. Because our prevalence estimates are based on claims data, i.e., care prevalence, regional variations in prevalence may be reflective of regional variations in access to care, treatment patterns, decisions to test or screen for cancer and patient proclivity to seek medical care (e.g., related to racial and/or cultural differences). Our findings with respect to cancer prevalence rates across age groups and genders are consistent with previous studies.

ACKNOWLEDGEMENTS

Supported by the University of Maryland, Maryland Statewide Health Network through the Maryland Cigarette Restitution Funds. The data in this study was provided through cooperation with the Maryland Department of Health and Mental Hygiene.

The authors want to thank Lisa Blatt for her help with the preparation of this manuscript.

REFERENCES

1. Glendening PN, Townsend KK, Benjamin GC, et al. Maryland Department of Health & Mental Hygiene. Annual Cancer Report. Cigarette Restitution Fund Program. Cancer Prevention, Education, Screening and Treatment Program. Baltimore, MD. September, 2001.

2. Howe HL, Wingo PA, Thun MJ, et al. Annual report to the nation on the status of cancer (1973 through 1998), featuring cancers with recent increasing trends. *J Natl Cancer Inst.* 2001;93:824-842.

3. Bradley CJ, Given CW, Roberts C. Disparities in cancer diagnosis and survival. *Cancer.* 2001;91:178-188.

4. Roetzheim RG, Pal N, Tennant C, et al. Effects of health insurance and race on early detection of cancer. *J Natl Cancer Inst.* 1999; 91:1409-1415.

5. Bradley CJ, Given CW, Roberts C. Race, socioeconomic status, and breast cancer treatment and survival. *J Natl Cancer Inst.* 2002;94:490-496.

6. Sung JF, Blumenthal DS, Alema-Mensah, E, et al. Racial and urban/rural differences in cervical carcinoma in Georgia Medicaid recipients. *Cancer.* 1997;80:231-236.

7. Fleiss, JL. Statistical Methods for Rates and Proportions (2nd edition). New York, NY: John Wiley & Sons, Inc. 1981.

8. Gordis L. Measuring the occurrence of disease. In: Gordis L. Epidemiology (2nd edition). Philadelphia, PA: WB Saunders Company. 2000.

9. Polednak AP. Estimating the prevalence of cancer in the United States. *Cancer.* 1997;80:136-141.

10. Capocaccia R, DeAngelis R. Estimating the completeness of prevalence based on cancer registry data. *Stat Med.* 1997;16:425-440.

11. National Cancer Institute. About SEER. Available at: http://seer.cancer.gov/about/. Accessed October 16, 2002.

12. Merrill RM, Capocaccia R, Feuer EJ, et al. Cancer prevalence estimates based on tumor registry data in the Surveillance, Epidemiology, and End Results (SEER) program. *Int J Epidemiol.* 2000;29:197-207.

13. Byrne J, Kessler L, Devesa SS. The prevalence of cancer among adults in the United States: 1987. *Cancer.* 1992;69:2154-2159.

14. Hewitt M, Breen N, Devesa S. Cancer prevalence and survivorship issues: analyses of the 1992 National Health Interview Survey. *J Natl Cancer Inst.* 1999;91:1480-1486.

15. National Cancer Institute. Prevalence Definitions. Available at: http://srab.cancer.gov/prevalence/definitions.html. Accessed September 5, 2002.

16. Joseph AK, Mark TL, Mueller C. The period prevalence and costs of treating nonmelanoma skin cancers in patients over 65 years of age covered by Medicare. Dermatol Surg. 2001;27:955-959.

17. McClish DK, Penberthy L, Whittemore M, et al. Ability of Medicare claims data and cancer registries to identify cancer cases and treatment. *Am J Epidemiol.* 1997;145:227-233.

18. Giles G. How important are estimates of cancer prevalence? *Ann Oncol.* 2002;13:815-816.

19. Wennberg JE. Unwarranted variations in healthcare delivery: implications for academic medical centres. *BMJ.* 2002;325:961-964.

20. Jollis JG, Ancukiewicz M, DeLong DR, et al. Discordance of databases designed for claims payment versus clinical information systems. Implications for outcomes research. *Ann Intern Med.* 1993;119:844-850.

21. Fowles JB, Lawthers AG, Weiner JP, et al. Agreement between physicians' office records and Medicare Part B claims data. *Health Care Financ Rev.* 1995;16:189-199. ■