

---

# Colon Cancer Treatment Costs for Medicare and Dually Eligible Beneficiaries

Zhehui Luo, Ph.D., Cathy J. Bradley, Ph.D., Bassam A. Dahman, and Joseph C. Gardiner, Ph.D.

---

*To estimate the cost attributable to colon cancer treatment 1 year after diagnosis by cancer stage, comorbidity, treatment regimen, and Medicaid eligibility, we extracted an inception cohort of colon cancer patients aged 66 and older diagnosed between 1997 and 2000 from the Michigan Tumor Registry. Patients were matched to non-cancer control subjects in the Medicare Denominator file. We used the difference-in-differences method to estimate costs attributable to cancer, controlling for costs prior to diagnosis. The mean total colon cancer cost per Medicare patient was \$29,196. The method can be applied to longitudinal data to estimate long term costs of cancer from inception where incident patients are identified from a tumor registry.*

## INTRODUCTION

The cost of colorectal cancer has recently been the subject of several scientific investigations (Wright et al., 2007; Yabroff et al., 2007a; Warren et al., 2008). These investigations were most likely spurred by recent screening initiatives and efforts to raise public awareness of colorectal cancer. Accurately estimating the direct medical cost of cancer is relevant to policymakers weighing new options for cancer prevention and control, screening

guidelines, and treatments. A descriptive review of cancer cost studies found significant heterogeneity in estimation methods, study settings, populations, and measurements of cost (Yabroff et al., 2007b). Past analyses of the cost of cancer treatment focused on long-term aggregate estimates (Brown et al., 1999; 2002; Etzioni et al., 2002) and were not designed to answer questions related to patient characteristics or treatment regimens.

In this study, we have two objectives: (1) to extend prior studies by estimating the cost attributable to colon cancer 1 year after diagnosis by cancer stage, comorbidity, treatment regimen, and other patient characteristics; and (2) to estimate the differences in 1-year cost between Medicare only and the dually eligible beneficiaries. Colon cancer usually occurs later in life (at age 60 to 70 years), and Medicare and Medicaid are the primary payers of cancer care. We focused on colon cancer instead of colorectal cancer because the cost of rectum cancer is usually higher and because colon cancer is among the cancer sites where screening, early detection, and effective treatment are feasible and proven to reduce mortality (Midgley and Kerr, 2005). Individuals who receive health care coverage from the Medicare and Medicaid Programs for at least 12 months prior to the diagnosis of cancer are defined as dual eligibles in this and our previous study (Bradley, Luo, and Given, 2008). Dually eligible beneficiaries are more likely to live under the Federal poverty level, reside in nursing homes or live alone, be from a minority population and unmarried, and

---

Zhehui Luo and Joseph C. Gardiner are with Michigan State University. Cathy J. Bradley and Bassam A. Dahman are with Virginia Commonwealth University. The research in this article was supported by National Cancer Institute Grant Number R01-CA101835-01. The statements expressed in this article are those of the authors and do not necessarily reflect the views or policies of Michigan State University, Virginia Commonwealth University, National Cancer Institute, or Centers for Medicare & Medicaid Services (CMS).

have lower education attainment (Murray and Shatto, 1998). Studies have found that Medicaid patients are less likely to receive cancer screening and more likely to be diagnosed at a later cancer stage than are Medicare only patients (Ward et al., 2008). An inquiry on cancer cost differentials by cancer stage, treatment procedure and comorbidity between Medicare only and dually eligible groups can shed light on disparity in healthcare utilization. Our method of estimating 1-year cost takes into account prior year non-cancer costs and treatment received.

## DATA AND METHODS

### Cancer Patients

We used statewide Medicaid and Medicare data merged with the Michigan Tumor Registry to extract a study sample of patients with a first primary colon cancer diagnosis in the years 1997 through 1999. The Michigan Cancer Surveillance Program, which maintains the Michigan Tumor Registry, is more than 95% complete based on external audit findings. For details of the linkage process, see Bradley et al. (2007). This study was approved by Institutional Review Boards at the Michigan Department of Community Health, Michigan State University, and Virginia Commonwealth University.

From statewide Medicare files, we extracted all claims for inpatient, outpatient, physician services, and hospice during the study period for all patients who correctly matched to the Michigan State segment of the Medicare Denominator file (approximately 89% of patients) and were enrolled in Parts A and B. Patients enrolled in Part A only were excluded for lack of physician office visit information.

We identified 8,157 Medicare Parts A and B beneficiaries aged 66 years and older who had a first primary colon cancer diagnosis from 1997 to 1999. Our database contains claims from January 1996 to December 2000 so that all patients had at least 12 months of data before and after the month of diagnosis. We excluded patients enrolled in managed care ( $n=512$ ) because their claims were not available. We also excluded cancer patients who had no claims ( $n=144$ ) or had zero cost ( $n=22$ ) during the study period. Patients with invasive but unknown stage of cancer were excluded because we could not assign these patients to a specific stage ( $n=782$ ). Patients of other or unknown race ( $n=128$ ) were excluded to avoid mismatch with controls (see non-cancer subject section below). Finally, 107 patients were excluded because they did not have a matched control subject or their matched controls had no claims or valid cost data in the study period. The remaining sample size was 6,462 of which 765 were continuously insured by Medicaid since the time of diagnosis in addition to Medicare.

Claims data were used to identify treatment. Surgery procedures were identified in the inpatient and outpatient files using International Classification of Disease, 9th Edition (ICD-9) codes.<sup>1</sup> Chemotherapy initiation was identified by at least one claim indicating the administration of chemotherapy within 6 months following diagnosis.<sup>2</sup> Hershman et al. (2006) found that 91% of elderly colon cancer patients initiate chemotherapy within 3 months of diagnosis.

<sup>1</sup> The ICD-9 codes were 45.71-45.79, 45.8, 48.41-48.49, 48.50, and 48.61-48.69.

<sup>2</sup> Chemotherapy was identified by the Current Procedural Terminology (CPT) codes 96400-96599; Health Care Common Procedural Codes Q0083-Q0085, J8510, J8520, J8521, J8530-J8999, J9000-J9999, J0640; and ICD-9 codes E0781, E9331, and V58.1.

## Non-Cancer Subjects

To attribute costs to a particular disease, researchers have examined and designated each claim as related to or not related to the disease under study (Finkelstein et al., 2003). However, disease causality and concurrence is a complex phenomenon. For example, depression has been found to be both a risk factor for cancer (Gallo et al., 2000) and a consequence (Polsky et al., 2005) of cancer, but including depression treatment as a “cancer cost” is questionable. Therefore, researchers have turned to matching cancer patients to non-cancer controls and comparing costs in each group to distinguish between cancer and non-cancer treatments. Various matching methods have been applied to match patients with and without the disease under study. We took a broader perspective to assess cancer costs by randomly selecting up to three control subjects to each cancer patient matched on age, race, sex, and health service area of residence. We used the cancer patient’s date of diagnosis as the reference date for the matched controls to establish a pre- and post diagnosis period.

## Outcome and Control Variables

The primary outcome of interest was the total cost of cancer treatment in the year after diagnosis or until death within 1 year of diagnosis. Previous research has shown that most short-term cancer cost occurs within the first year of diagnosis (Delco et al., 2005). Medicare covers inpatient services (Part A) and outpatient services (Part B). We used the sum of Medicare payment, patient deductible and coinsurance amount, and the third-party payer paid amount as a proxy for the value of medical services. All cost estimates are

in 2000 dollars deflated by the Medicare Economic Index.<sup>3</sup>

Closely associated with cost and treatment options is survival. Patients’ survival was ascertained through the Medicare Denominator file and National Death Index. Dually eligible breast cancer patients had poorer 8-year survival compared with Medicare only patients (Bradley et al., 2005). Patients who die within 1 year of diagnosis may have higher or lower costs depending on the length of survival and treatments received. Brown et al. (1999) found that the content of care for patients with short survival is more similar to that of the last year of life phase than that of the initial phase. Because the cost in the last year of life phase is much higher than the cost in the initial phase, we may expect higher cost in the year after diagnosis among those with short survival than those who survive more than 1 year.

We defined cancer stage using the Surveillance Epidemiology and End Results (SEER) summary stages (in situ, local, regional, and distant) and excluded patients with unknown stage. We constructed the Deyo, Cherkin, and Ciol (1992) and Klabunde et al. (2000) adaptation of the Charlson Comorbidity Index as comorbidity burden for cancer patients and their controls before and after the diagnosis or reference date. We used patients’ inpatient, outpatient, and physician claims to construct the Comorbidity Index, which was grouped into categories 0, 1, 2, and  $\geq 3$ .

Data on patient age, race, and sex were obtained from the Michigan Tumor Registry. Age was grouped into the following categories: 66 to 70 years, 71 to 75 years, 76 to 80 years, and older than 80 years. Based on patients’ address, we linked the census tract median household income

<sup>3</sup> We did not use the Hospital Wage Index to adjust for inflation in Part A costs because our data are from a single State.

and education level to each patient. The income categories were <\$25,000; \$25,001 to \$35,000; \$35,001 to \$45,000; and >\$45,000. Education in each census tract was measured by the percentage of the population with less than high school, high school but not college, and college or more education. Missing values in income and education were imputed using the mean imputation method.<sup>4</sup> Based on patients' county of residence, we obtained the number of short-term hospitals with oncology services and the number of colon/rectum surgical specialists as measures of county-level resource availability.

Adjuvant and palliative chemotherapy is the standard treatment for advanced stage cancer, and recent evidence suggests the use of chemotherapy in stage II cancer as well. Thus, we categorize cancer treatments to three groups: no resection (including those with no adjuvant treatment and those with chemotherapy only ( $n=1,177$  [18.21%]), one or more resection without chemotherapy ( $n=3,665$  [56.72%]), and one or more surgeries with chemotherapy ( $n=1,620$  [25.07%]).<sup>5</sup>

## Statistical Methods

Our first objective was to estimate the mean cost attributable to cancer 1 year after diagnosis. Three features of cost data presented themselves immediately. First, a substantial proportion of patients had zero cost in the 12 months before diagnosis and a substantial proportion of control patients had zero cost in both periods. Second, costs for cancer patients in the 12 months after diagnosis had a different distribution than costs for cancer

patients in the 12 months before diagnosis and for control patients in both periods. Finally, the expenditure data were highly skewed. Because of these features, we used strategies other than Ordinary Least Squares regression to estimate the marginal effect of patient characteristics on mean cancer costs.

First, we formulated a two-part model (Mullahy, 1998) for costs of control patients in both periods and of cancer patients in the 12 months before diagnosis, which contain many observations with zero cost. The first part of the two-part model estimates the probability of any cost, specified as a probit (Equation 1) or logit (Equation 2) model.

$$\Pr(y_{it} > 0 | \mathbf{x}_{it}) = \Phi(\mathbf{x}'_{it}\beta) \quad (1)$$

$$\Pr(y_{it} > 0 | \mathbf{x}_{it}) = \frac{\exp(\mathbf{x}'_{it}\beta)}{1 + \exp(\mathbf{x}'_{it}\beta)} \quad (2)$$

where  $\Phi$  denotes the cumulative density function for the standard normal distribution,  $y_{it}$  the direct medical cost for patient  $i$  in period  $t$ , and  $t = 0$  or  $1$  for the 12 months before or after the reference date.

Second, we considered alternatives for the second part of the two-part model and the estimation of the mean cost for cancer patients after diagnosis. Equation 3 represents a general specification for this part:

$$E(y_{it} | y_{it} > 0, \mathbf{z}_{it}) = f(\mathbf{z}'_{it}\gamma) \quad (3)$$

There are three general ways to address non-normal and skewed data. We can (1) transform the data using some functional forms (e.g., log transformation, square-root transformation, or Box-Cox transformation), (2) use parametric distributions in a generalized linear model (GLM), or (3) use nonparametric approaches. The first approach leads to difficulties

<sup>4</sup> The number of patients with imputed income and education value is 284 (4.4%) and 315 (4.9%) in the final sample. Excluding these patients did not change the results substantively.

<sup>5</sup> The number of patients with more than one resection is 147, among which 101 did not have chemotherapy, and 34 did. The sample is too small to provide separate estimates.

in retransformation to the original scale of costs. In addition, if the variance of the errors is related to covariates, then retransformed mean estimates could be biased (Manning, 1998; Duan, 1983). The last approach suffers from the dimensionality problem as well as difficulties in interpretation. We followed Manning and colleagues (2005) and systematically compared log-, square-root, Box-Cox transformation, and GLM with gamma distribution through a series of tests for distribution, nonlinearity, specification, goodness of fit, and overfitting. For the non-zero part of the two-part model, the Park test was used to gauge the selection of the distributions. The Pregibon Link test and the RESET test were used for nonlinearity of the specification, the modified Hosmer-Lemeshow test was used for goodness of fit, and the Copas test was used for overfitting using split sample cross-validation.<sup>6</sup> The best fitted models were used to estimate mean total medical cost in each period for cancer and control patients who had incurred any cost.

Combining the first and second part of the two-part model together and the stand-alone part for the cost of cancer patients after diagnosis (always positive), we estimated the expected values for all medical costs for cancer and control patients before and after the diagnosis or reference date (Equation 4 or 5):

$$E(y_{it}|\mathbf{x}_{it}, \mathbf{z}_{it}) = E(y_{it}|y_{it}>0, \mathbf{x}_{it}, \mathbf{z}_{it}) \Pr(y_{it}>0|\mathbf{x}_{it}, \mathbf{z}_{it}) = f(\mathbf{z}'_{it}\gamma) \Phi(\mathbf{x}'_{it}\beta) \quad (4)$$

$$E(y_{it}|\mathbf{x}_{it}, \mathbf{z}_{it}) = E(y_{it}|y_{it}>0, \mathbf{x}_{it}, \mathbf{z}_{it}) \Pr(y_{it}>0|\mathbf{x}_{it}, \mathbf{z}_{it}) = f(\mathbf{z}'_{it}\gamma) \frac{\exp(\mathbf{x}'_{it}\beta)}{1+\exp(\mathbf{x}'_{it}\beta)} \quad (5)$$

We then used the difference-in-differences method to estimate costs attributable to cancer. One-year total costs attributable to cancer were calculated as the difference of two differences: the difference between cancer patients and control subjects and the difference between the period before and after the diagnosis/reference date:

$$\text{DID} = [E(y|\text{cancer, post}) - E(y|\text{control, post})] - [E(y|\text{cancer, before}) - E(y|\text{control, before})]$$

This method is analogous to a quasi-experimental design (Card and Krueger, 1994) in that it reduces the contamination caused by temporal trends in increasing costs. The first difference eliminates the average cost attributable to other medical costs after the diagnosis of cancer, and the second difference eliminates the residual difference in medical costs before the diagnosis of cancer due to unmatched or unobserved characteristics of the cancer and control patients.

Prior to the formal diagnosis of cancer, some patients may have incurred costs for lab tests or “rule-out” visits. These costs can arguably be considered part of cancer costs. In this case, we carried out a sensitivity analysis by including costs incurred 1 to 3 months before diagnosis as cancer costs. Thus, we calculate the cancer costs in 12-, 13-, 14-, and 15-month periods separately and compare the range of estimated costs attributable to cancer.

In all estimation, we used heteroskedasticity- and cluster-robust standard errors because multiple observations for each cancer patient and for his or her multiple matched controls may lead to correlations between outcomes. To estimate the incremental cost between Medicare only and dually eligible patients and between

<sup>6</sup> References to these statistical tests are available upon request.

cancer patients with different diagnosis stage, treatment regimen, and comorbidity, we used the method of recycled

predictions (Basu and Rathouz, 2005). This method entails comparisons of two predictive margins where a particular

**Table 1**  
**Characteristic of Colon Cancer Patients and Control Subjects**

		Cancer Cases (N=6,462)		Control Subjects (N=11,483)	
		N	(%)	N	(%)
Age	66-70 years	1251	19.36	2203	19.18
	71-75 years	1629	25.21	2920	25.43
	76-80 years	1542	23.86	2777	24.18
	>80 years	2040	31.57	3583	31.2
Race	White	5711	88.38	10123	88.16
	African American	751	11.62	1360	11.84
Sex	Male	2841	43.96	5039	43.88
	Female	3621	56.04	6444	56.12
SEER Stage	In situ	275	4.26	n.a.	n.a.
	Local	2412	37.33	n.a.	n.a.
	Regional	2631	40.71	n.a.	n.a.
	Distant	1144	17.7	n.a.	n.a.
Census tract median annual income	≤\$25k	1853	28.68	3305	28.78
	\$25k to ≤\$35k	2053	31.77	3644	31.73
	\$35k to ≤\$45k	1449	22.42	2566	22.35
	>\$45k	823	12.74	1454	12.66
	Missing	284	4.39	514	4.48
Charlson Index		N	(%)	N	(%)
	0*	4264	65.99	8048	70.09
	1*	1271	19.67	2080	18.11
	2	500	7.74	805	7.01
	3+*	427	6.61	550	4.79
	Myocardial infarction	105	1.62	154	1.34
	Congestive heart failure*	646	10	895	7.79
	Peripheral vascular disease	203	3.14	347	3.02
	Cerebrovascular disease	340	5.26	587	5.11
	Obstructive pulmonary disease*	701	10.85	1011	8.8
	Dementia*	74	1.15	229	1.99
	Diabetes*	1008	15.6	1373	11.96
	Chronic renal failure	74	1.15	144	1.25
	Ulcer*	129	2	127	1.11
Rheumatism	96	1.49	172	1.5	
Alzheimer's disease*	123	1.9	363	3.16	

\* Indicates cases and controls had statistically significantly ( $p < 0.05$ ) different comorbid conditions based on likelihood ratio tests in conditional logistic regressions. Conditions with prevalence less than 1% in both groups are not presented: paralysis, cirrhosis, and liver disease.

SOURCE: Michigan Tumor Registry, Medicare and Medicaid fee-for-service claims from 1996 to 2000.

attribute (such as dual eligibility) is assumed present or absent (Graubard and Korn, 1999). Because of the complexity of the model, we obtained bootstrap standard errors and bias-corrected bootstrap confidence intervals of the predicted differences in total cost between Medicare only and dually eligible patients.

## Results

Table 1 reports the demographic and comorbid conditions of the cancer patients and the controls. Age, race, sex, and health service areas were distributed evenly due to matching. Cancer patients had higher comorbidity in the 12 months prior to their diagnosis of cancer. More cancer patients had congestive heart failure, obstructive pulmonary disease, diabetes, and ulcer, but fewer cancer patients had dementia and Alzheimer's disease. There was no statistically significant difference in the prevalence of other diseases between cancer patients and their matched control subjects.

Table 2 compares the demographic, comorbid conditions, survival, and treatment regimens among cancer patients by dual eligibility status. Compared with Medicare only patients, dually eligible patients were older, had higher proportions of African American and female individuals, lived in neighborhoods with lower income and education, and had similar cancer stage ( $p=0.121$ ) but worse survival (Table 3,  $p<0.001$ ). The dually eligible patients also had higher prevalence in 12 out of the 14 comorbid conditions in the Charlson Comorbidity Index.

As seen in Table 3, fewer dually eligible patients received combined resection and chemotherapy treatment as compared to the Medicare only patients (15% versus 26%). Among those who were diagnosed at the in situ/local or regional stage,

dually eligible patients were also more likely to receive no treatment or chemotherapy only and less likely to have combined resection and chemotherapy. Dually eligible patients had a higher fatality rate than the Medicare only group (23% versus 12% for in situ/local stage; 30% versus 21% for regional stage). Patients with distant stage of cancer had similar survival ( $p=0.584$ ) and similar treatments ( $p=0.133$ ) between the Medicare only and the dually eligible patients.

Our sensitivity analysis excluding 1, 2, or 3 months of claims before the diagnosis/reference date indicated that excluding costs in the month prior to the diagnosis/reference date led to comparable total, inpatient, and outpatient costs between cancer patients and their matched controls in the period before the diagnosis/reference date. Thus, our study estimated cost attributable to cancer in a period of 13 months: 1 month before the actual diagnosis date and 12 months after diagnosis.<sup>7</sup> Unadjusted direct medical costs in the 11-month base period and the 13-month post period for cancer patients and corresponding costs for controls subjects are summarized in Table 4. Compared to controls, cancer patients had similar total, inpatient, and outpatient costs in the baseline period. However, cancer patients had lower physician and hospice costs. In the post cancer period, the average total costs for cancer patients were \$28,832 higher than the total costs of control subjects, and the majority of this difference was due to inpatient costs (\$20,470). In addition, outpatient and physician office costs were higher in cancer patients (\$2,361 and \$5,522, respectively). In the base period, more cancer patients had inpa-

<sup>7</sup> Sensitivity analysis results excluding claims 0, 2, or 3 months before diagnosis are available upon request. These analyses did not change the final prediction of total cost attributable to cancer substantially.

**Table 2**  
**Characteristics of Colon Cancer Patients by Medicare/Medicaid Eligibility Status**

		Medicare Only (N=5,697)		Dually Eligible (N=765)	
		N	(%)	N	(%)
Age*	66–70 years	1125	19.75	126	16.47
	71–75 years	1438	25.24	191	24.97
	76–80 years	1384	24.29	158	20.65
	>80 years	1750	30.72	290	37.91
Race*	White	5187	91.05	524	68.5
	African American	510	8.95	241	31.5
Sex*	Male	2613	45.87	228	29.8
	Female	3084	54.13	537	70.2
SEER Stage	In situ	250	4.39	25	3.27
	Local	2140	37.56	272	35.56
	Regional	2292	40.23	339	44.31
	Distant	1015	17.82	129	16.86
Census tract median annual income*	≤\$25k	1477	25.93	376	49.15
	\$25k to ≤\$35k	1797	31.54	256	33.46
	\$35k to ≤\$45k	1380	24.22	69	9.02
	>\$45k	796	13.97	27	3.53
	Missing	247	4.34	37	4.84
		Mean	Std	Mean	Std
Census tract education	Percent <12 year*	0.23	0.1	0.3	0.13
	Percent <college*	0.6	0.09	0.57	0.09
	Percent ≥college*	0.17	0.14	0.13	0.1
Charlson Index		N	(%)	N	(%)
	0*	3848	67.54	416	54.38
	1*	1097	19.26	174	22.75
	2	419	7.35	81	10.59
	3+*	333	5.85	94	12.29
	Myocardial infarction	91	1.6	14	1.83
	Congestive heart failure*	526	9.23	120	15.69
	Peripheral vascular disease	168	2.95	35	4.58
	Cerebrovascular disease	275	4.83	65	8.5
	Obstructive pulmonary disease*	587	10.3	114	14.9
	Diabetes*	841	14.76	167	21.83
	Chronic renal failure*	57	1	17	2.22
	Ulcer*	105	1.84	24	3.14
	Rheumatism	83	1.46	13	1.7
Alzheimer's disease*	75	1.32	48	6.27	

\* Indicates statistical significant difference between the two groups ( $p < 0.05$ ). Conditions with prevalence less than 1% in either group are not presented: dementia, paralysis, cirrhosis and liver disease.

NOTE: The two-sample t-test was used for continuous variables, the Pearson chi-square test was used for categorical variables, and Fisher's exact test was used when the cell size is smaller than 5.

SOURCE: Michigan Tumor Registry, Medicare and Medicaid fee-for-service claims from 1996 to 2000.



tient claims as compared to the controls, but fewer cancer patients had outpatient claims or physician office visits.

The shape of the cost distributions between cancer and control subjects before and after the reference date were very different, indicating the need for separate estimations of the mean cost. Figure 1 displays the box plots of the total costs, square-root transformation of the total costs, and logarithm transformation of the total costs plus 0.05 for cancer and control patients in both periods. The shape for cancer and control subjects' costs in the 11 months before the diagnosis and reference date and the cost for

control subjects in the 13 months after the reference date are similar. However, the distribution for cancer patients' costs in the 13 months after the diagnosis date was very different. The Park test, the Pregibon Link test, the RESET test, the modified Hosmer-Lemeshow test, and the Copas test for the second part of the two-part model all favored the gamma distribution in a GLM over the other specifications. The modified Hosmer-Lemeshow test and the Copas test for the overall two-part model also favored the GLM gamma distribution over the log-normal distribution.

**Table 3**  
**One-Year Survival and Treatment Procedures by Medicare/Medicaid Eligibility Status**

	Medicare Only		Dually Eligible		p-value
	N	%	N	%	
Overall	N=5,697		N=765		
Death	1488	26.12	262	34.25	<0.001
None or chemo only	1011	17.75	166	21.7	
Resection no chemo	3184	55.89	481	62.88	
Resection with chemo	1502	26.36	118	15.42	<0.001
In Situ/Local Stage	N=2,390		N=297		
Death	292	12.22	67	22.56	<0.001
None or chemo only	483	20.21	82	27.61	
Resection no chemo	1667	69.75	203	68.35	
Resection with chemo	240	10.04	12	4.04	<0.001
Regional Stage	N=2,292		N=339		
Death	480	20.94	101	29.79	<0.001
None or chemo only	171	7.46	33	9.73	
Resection no chemo	1166	50.87	228	67.26	
Resection with chemo	955	41.67	78	23.01	<0.001
Distant Stage	N=1,015		N=129		
Death	716	70.54	94	72.87	0.584
None or chemo only	357	35.17	51	39.53	
Resection no chemo	351	34.58	50	38.76	
Resection with chemo	307	30.25	28	21.71	0.133

NOTE: The Pearson chi-square test was used for testing between Medicare only and dually eligible patients.

SOURCE: Michigan Tumor Resigtry, Medicare and Medicaid fee-for-service claims from 1996 to 2000.

Table 5 reports the average of predicted costs by cancer stage, age group, comorbidity, survival, and treatment received for all patients and for Medicare only and dually eligible patients separately estimated through a two-part model. The average direct medical costs attributable to cancer in 1 year after diagnosis were \$29,196. Treatment costs were not statistically significantly different between the Medicare only patients and the dually eligible patients ( $\Delta$ =\$1,272, 95% CI = [-\$357, \$2,769]). Patients with regional stage

cancer at diagnosis had the highest cost (\$30,748) followed by patients diagnosed with distant stage cancer (\$29,933) and patients with in situ or local stage cancer (\$27,551). The total costs for the dually eligible patients with regional and distant stage of cancer were lower than their Medicare only counterparts by \$2,050 ( $p$ <0.1) and by \$3,335 ( $p$ <0.1), respectively; and costs for in situ/local stage cancer were similar between the two groups.

Average total cancer costs were \$14,696, \$28,703, and \$42,523 for patients under-

**Table 4**  
**Average Costs and Percentage of Patients with Positive Costs for Cancer Patients and Control Subjects Before and After Diagnosis or Reference Date**

		Cancer Cases	Control Subjects	Difference	95% Confidence Interval	
11 months before reference date – 30 days <sup>a</sup>	Total	4757	4653	103	(-221,	428)
	Inpatient	2691	2451	240	(-16,	497)
	Outpatient	758	789	-31	(-96,	35)
	Physician*	1303	1387	-84	(-147,	-21)
	Hospice*	3	27	-23	(-35,	-12)
13 months after reference date – 30 <sup>a</sup>	Total*	34077	5249	28832	(28209,	29455)
	Inpatient*	23234	2767	20470	(19974,	20966)
	Outpatient*	3169	808	2361	(2232,	2490)
	Physician*	7143	1622	5522	(5361,	5683)
	Hospice*	530	52	479	(416,	541)
		Cancer Cases			Control Subjects	
		N	%		N	%
11 months before reference date – 30 days <sup>b</sup>	Total	5722	88.55		10219	88.99
	Inpatient*	1279	19.79		2038	17.75
	Outpatient*	4152	64.25		7815	68.06
	Physician*	5240	81.09		9908	86.28
	Hospice*	3	0.05		50	0.44
13 months after reference date – 30 <sup>b</sup>	Total	6462	100		10007	87.15
	Inpatient* <sup>b</sup>	6043	93.52		2214	19.28
	Outpatient* <sup>b</sup>	5923	91.66		7765	67.62
	Physician* <sup>b</sup>	6150	95.17		9778	85.15
	Hospice* <sup>b</sup>	770	11.92		105	0.91

\* Indicates statistically significant difference ( $p$ <0.05).

<sup>a</sup> Wald test for equality of means was used to compare costs between cancer patients and control subjects in each period. Robust standard errors were used.

<sup>b</sup> Conditional logistic regression for matched case-control groups was used to compare percent with zero cost.

SOURCE: Michigan Tumor Resigtry, Medicare and Medicaid fee-for-service claims from 1996 to 2000.

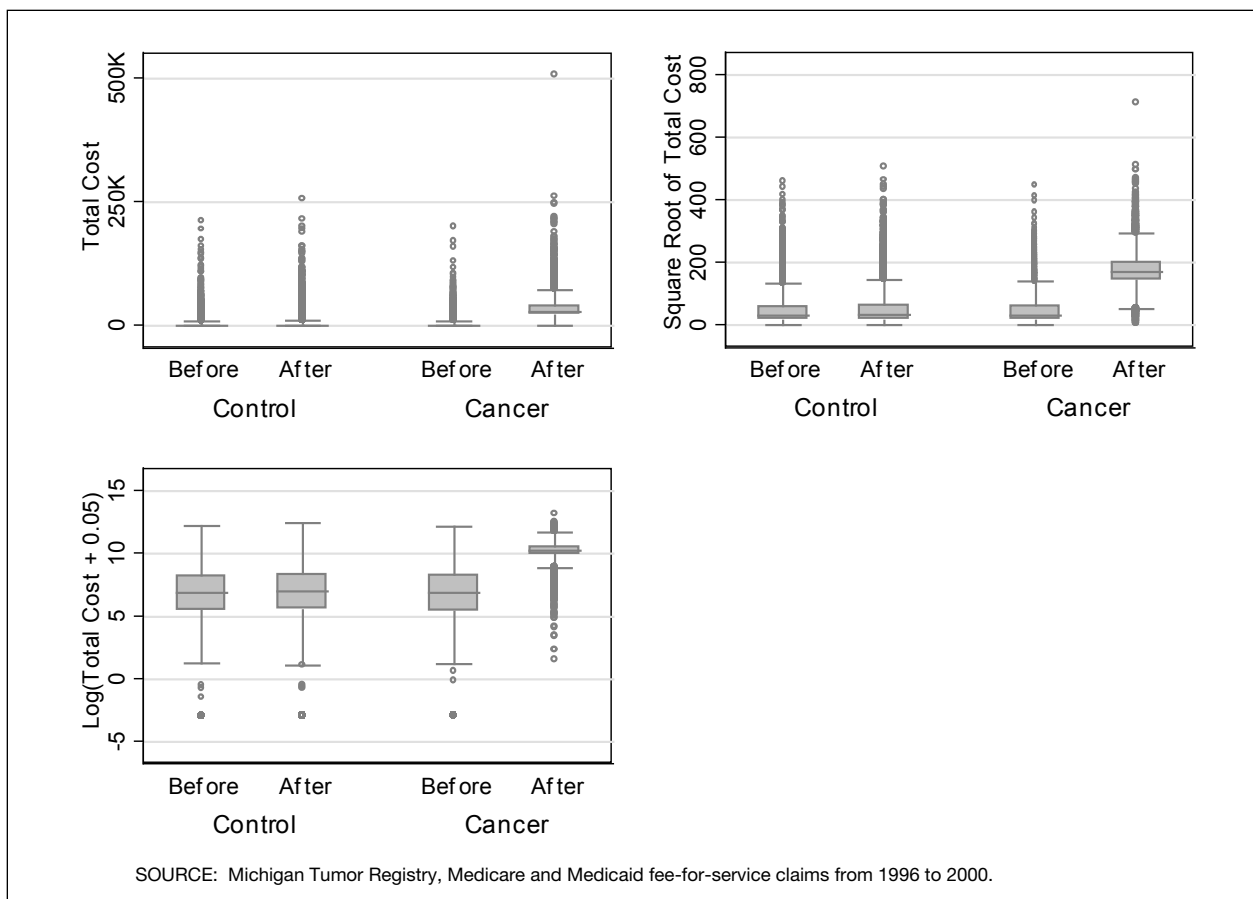
going no treatment or chemotherapy only, resection, and resection combined with chemotherapy, respectively. The treatment costs for resection combined with chemotherapy were much higher than the other treatment regimens. None of the differences between the Medicare only and the dually eligible patients by treatment procedures were statistically significant. However, the overall difference in costs for patients with combined resection and chemotherapy was substantial between the Medicare only and dually eligible patients. Given the same cancer stage, the differences were statistically significant for patients with regional or distant cancer and undergoing both resection and chemotherapy ( $\Delta=\$3,555$ ,

95% CI = [ $\$182$ ,  $\$6,677$ ] for regional cancer;  $\Delta=\$5,740$ , 95% CI = [ $\$223$ ,  $\$11,000$ ] for distant cancer). With more evidence on the benefits of adjuvant chemotherapy in the elderly, the gap between Medicare only and dually eligible patients in treatments needs to be addressed.

Patients with more comorbid conditions had higher costs than patients with few comorbid conditions. Medicare only cancer patients without comorbid conditions or with only one comorbid condition had higher costs than their dually eligible counterparts ( $\Delta=\$1,109$ ,  $p<0.1$ ;  $\Delta=\$1,280$ ,  $p<0.1$ ). The differences in average total costs for the other comorbid groups were statistically similar between the Medicare only and the dually eligible patients.

**Figure 1**

**Box Plots for Total Costs, Square-Root Transformation of Total Costs, and Logarithm Transformation of Total Costs Plus 0.05 for Cancer and Control Patients in Both Periods**



**Table 5**  
**Predications of Mean Cost Attributable to Cancer<sup>a</sup>**

	All (\$)	Medicare Only (\$)	Dually Eligible (\$)	Difference (\$) [95%CI] <sup>b</sup>
All patients	29196	29342	28070	1272 <sup>c</sup> [-357, 2569]
<b>Stage of Cancer</b>				
In situ/local	27551	27458	28063	-605 [-3866, 2126]
Regional	30748	30993	28943	2050 <sup>c</sup> [-341, 4109]
Distant	29933	30326	26992	3335 <sup>c</sup> [-64, 6438]
<b>Procedure</b>				
None or Chemo	14696	14760	14276	483 [-2246, 3376]
Resection alone	28703	28822	27922	900 [-1074, 2468]
Resection + Chemo	42523	42860	40322	2538 [-970, 6073]
<b>Comorbidity</b>				
0	25205	25329	24220	1109 <sup>c</sup> [-143, 2167]
1	27967	28110	26830	1280 <sup>c</sup> [-288, 2532]
2	28300	28465	26985	1480 [-573, 3044]
3+	32922	33125	31294	1832 [-1297, 4273]
<b>Stage of Cancer x Procedure</b>				
In situ/local x None or Chemo	10920	10844	11403	-560 [-3309, 1909]
In situ/local x Resection alone	27225	27105	27983	-878 [-4390, 1806]
In situ/local x Resection + Chemo	41426	41431	41403	28 [-5867, 4994]
Regional x None or Chemo	18101	18233	17238	995 [-2839, 4583]
Regional x Resection alone	30308	30543	28775	1768 [-1155, 4141]
Regional x Resection + Chemo	41624	42095	38540	3555 <sup>d</sup> [182, 6677]
Distant x None or Chemo	15332	15546	13908	1638 [-1519, 4578]
Distant x Resection alone	27925	28300	25423	2877 <sup>c</sup> [-644, 6166]
Distant x Resection + Chemo	45886	46634	40895	5740 <sup>d</sup> [223, 11000]
<b>Stage of Cancer x Comorbidity</b>				
In situ/local x 0	21583	21523	21960	-437 [-2734, 1542]
In situ/local x 1	23207	23141	23615	-474 [-2999, 1732]
In situ/local x 2	23272	23204	23685	-481 [-3002, 1730]
In situ/local x 3+	24239	24159	24683	-524 [-3295, 1779]
Regional x 0	27602	27789	26104	1685 <sup>c</sup> [-296, 3439]
Regional x 1	25916	29971	27939	2032 <sup>c</sup> [-303, 4165]
Regional x 2	28210	28467	26162	2305 <sup>c</sup> [-409, 4813]
Regional x 3+	33611	33969	30756	3213 <sup>c</sup> [-677, 6666]
Distant x 0	28090	28408	25557	2851 <sup>d</sup> [12, 5546]
Distant x 1	28522	28885	25635	3250 <sup>c</sup> [-76, 6322]
Distant x 2	30367	30816	26787	4030 <sup>c</sup> [-198, 7838]
Distant x 3+	25594	26103	21540	4563 <sup>c</sup> [-452, 9088]

<sup>a</sup>Recycled prediction approach was used except for the prediction for all patients (first row, first column).

<sup>b</sup>Bootstrap bias-corrected confidence intervals are used for testing equality in costs between Medicare only and dually eligible patients. 1,000 cluster bootstrapped samples were used where each cancer patient and the controls were considered one cluster in bootstrapping.

<sup>c</sup>The tests between Medicare only and the dually eligible patients were significant at  $p < 0.1$  using normal-based bootstrap confidence intervals.

<sup>d</sup>The tests between Medicare only and the dually eligible patients were significant at  $p < 0.05$  using normal-based bootstrap confidence intervals.

SOURCE: Michigan Tumor Registry, Medicare and Medicaid fee-for-service claims from 1996 to 2000.

Given the same stage of diagnosis, patients with more comorbid conditions had higher costs, with the exception of patients with distant stage cancer and three or more comorbid conditions who had lower costs than patients with distant stage cancer and fewer comorbid conditions. Consistent with the main effects of cancer stage on the differences of average total costs between Medicare only and dually eligible patients, only regional and distant stage cancer patients had marginally significant higher costs among the Medicare only patients. The gap widened by comorbidity for each cancer stage. For example, for patients with regional cancer stage, the differences in costs ranged from \$1,685 to \$3,213 when comorbidity increased from 0 to 3 or more. For patients with distant stage cancer and no comorbidity, the difference in average total costs was statistically significant ( $\Delta$ =\$2,851, 95% CI = [\$12, \$5,546]).

### Sensitivity to Cost Outliers

Two cancer patients and one control patient had total costs greater than \$250,000 in the period after diagnosis/reference date, and one cancer patient had total costs greater than \$500,000 in 1 year after diagnosis (Figure 1). These observations increase the skewness of the data. Because our goal is not to predict who had outlying costs but to estimate the mean costs for different patients, we re-estimated the model excluding those observations. Without the observation greater than \$500,000, the mean total cost for colon cancer was \$29,124; and the difference between the Medicare only and dually eligible patients was \$1,161 (95% CI = [-\$180, \$2,783]). Without the three observations greater than \$250,000, the mean total cost for colon cancer was \$29,100; and the difference between the

Medicare only and dually eligible patients was \$1,212 (95% CI = [-\$239, \$2,591]). Unsurprisingly, dropping the large cost observations lowered the standard errors of the estimates slightly. No estimates had changed substantially for any subgroups by stage, treatment, or comorbidity. The statistical significance also remained largely unchanged.

### Discussion

Our findings provide population-based estimates of 1-year costs attributable to colon cancer by stage of diagnosis, comorbidity, treatments, and dual eligibility status of Medicare beneficiaries. The mean total cost attributable to colon cancer 1 year after diagnosis was \$29,196. Patients diagnosed with in situ and local stage had the lowest costs (\$27,551), followed by patients with distant stage (\$29,933), and patients with regional cancer had the highest cost (\$30,748). Given the same stage of diagnosis, patients with more comorbid conditions had higher costs. Having one, two, or three and more comorbid conditions increased costs by \$2,762, \$3,095, and \$7,717, respectively, as compared to patients with no comorbidity.

Overall treatment costs were higher among Medicare only patients than among dually eligible patients, but the difference was not statistically significant. Dually eligible patients with regional or distant stage cancer who had both resection and chemotherapy consistently had lower costs than their Medicare counterparts.

Our assessment of colon cancer costs offers several insights. First, we provide fine-tuned estimates of cancer costs during the first year following diagnosis. Our method differs from Wright et al. (2007) in that we explicitly model the zero cost outcome by a two-part model

approach and compare the differential cost between Medicare only and dually eligible patients. The regression method is a meaningful alternative to the phase-of-care method (Brown et al., 1999; Yabroff et al., 2007a) and provides policy-relevant information regarding cancer stage and treatment costs along with cost information specific to patient characteristics such as age and comorbidity. Our method allows a prospective prediction of cancer cost by subpopulation, whereas the phase-of-care estimation depends on retrospectively segmenting survival into different periods and is not directly suitable for predicting future cost for a given patient.

Second, cancer treatment cost varies by stage of diagnosis and comorbid conditions. If recent screening initiatives are effective and result in fewer cancer cases diagnosed at later stages, then the long-term costs of colon cancer will be lower. This has implications for the longer-term forecast of Medicare costs.

Third, a recent study examining trends in the initial phase of cancer treatment found that there were significant increases in the proportion of colorectal cancer patients undergoing chemotherapy treatment and in the average Medicare payment for those patients (Warren et al., 2008). This is consistent with our findings of significantly higher costs for patients with combined resection and chemotherapy. The cost of chemotherapy will likely increase as newer and more expensive multidrug chemotherapy regimens emerge.

Our estimates (\$29,196) for cost attributable to colon cancer in 1 year are lower than the estimate in Yabroff et al. (2007a) who forecasted colorectal cancer cost for the elderly (age 65 and above) by phase of care (the initial phase, the continuing phase, and last year of life) through

the year 2020.<sup>8</sup> The two estimates are not directly comparable in that the initial phase in Yabroff et al. (2007a) does not include patients who survived less than 13 months. Our data included patients with survival within 12 months of diagnosis for whom the total cost was incurred in less than 1 year. In addition, the Yabroff et al. (2007a) estimates include rectal cancer, which is more expensive to treat than colon cancer. We did not have data on skilled nursing home facility, home health care or durable medical equipment costs. Murray and Eppig (1999) found that 7 percent of Medicare beneficiaries used a skilled or long-term care facility in 1996 and home health care expenditure accounted for 5 percent of the total Medicare expenditure. Excluding these claims led to an underestimate of the total cancer costs in our study. Finally our study focused on patients undergoing surgery and chemotherapy and excluded patients who had radiation alone. Our aim was to estimate cancer costs by most typical treatment regimens.

The review by Yabroff et al. (2007b) identified the measurement of cost—payment, charges, or expenditures—as one major source of variation in cancer cost studies and lamented the lack of published standards for conducting and reporting cost analyses. Brown et al. (2002) indicated Medicare payment was a good proxy for the economic cost of medical services compared to alternatives based on charges or cost-to-charge ratios. They relied on a “scale up” approach to account for deductibles and copayment when direct measures were not available. We thus used the sum of Medicare payments, patient deductible and coinsurance and third-party payer paid amount as the measure of medical care costs, which is the most

<sup>8</sup> The estimates were reported in 2002 dollars in Yabroff et al. (2007a) whereas our estimates were in 2000 dollars.

comprehensive and reliable measure of cost in the current literature.

Our study has several limitations. First, the study sample is confined to a single State and thus may not be generalizable to other States or regions. However, that would only be the case if Michigan physicians treated patients differently than physicians treat patients in other States. Second, the study sample is specific to patients aged 65 years and older and may not be applicable to younger patients who may opt for more aggressive treatments. Third, the sample does not include patients enrolled in a managed care plan; these patients may have a pattern of care that is different from patients enrolled in a fee-for-service plan. Nevertheless, the method we use can be applied to larger, nationwide datasets to estimate longer-term costs. Finally, our study period was from 1997 to 2000, which precedes the Medicare Prescription Drug, Improvement and Modernization Act of 2003 and as such does not contain any estimation of prescription drug costs for the Medicare only group. To calculate total costs for the dually eligible patients we did not include prescription payment because comparable information was not available for Medicare only patients. It remains unknown if the Medicare only and dually eligible groups have different prescription costs in the Part D era.

We estimated the 1-year costs of colon cancer by stage, treatment, and patient characteristics such as comorbid conditions, age, and dual eligibility status. By incorporating these characteristics into our model, we can address questions regarding the incremental costs of treating older patients, patients with advanced stage disease, patients with more comorbid conditions, and patients undergoing different treatment regimens. Finally, we applied a method of cost esti-

mation to colon cancer that can be applied to larger national datasets for a longer-term estimation of costs of cancer. Special considerations need to be given when a control patient develops cancer using methods similar to the nested case-control design (Barlow et al., 1999). This method complements methods that segment costs by disease stage and can be used prospectively for cost prediction. As Medicare costs continue to grow, it is important to understand the potential factors that affect the projection of future costs.

## References

- Barlow, W. E., Ichikawa, L., Rosner, D., et al.: Analysis of Case-Cohort Designs. *Journal of Clinical Epidemiology* 52(12): 1165-1172, 1999.
- Basu, A. and Rathouz, P. J.: Estimating Marginal and Incremental Effects on Health Outcomes Using Flexible Link and Variance Function Models. *Biostatistics* 6(1): 93-109, 2005.
- Bradley, C. J., Gardiner, J., Given, C. W., et al.: Cancer, Medicaid Enrollment, and Survival Disparities. *Cancer* 103(8): 1712-1718, 2005.
- Bradley, C. J., Given, C. W., Luo, Z., et al.: Medicaid, Medicare, and the Michigan Tumor Registry: A Linkage Strategy. *Medical Decision Making* 27(4): 352-363, 2007.
- Bradley, C. J., Luo, Z., and Given, C. W.: Cancer Incidence in Elderly Medicare and Dually Eligible Beneficiaries. *Health Services Research* 43(5): 1768-1778, 2008.
- Brown, M. L., Riley, G. F., Potosky, A. L., et al.: Obtaining Long-Term Disease Specific Costs of Care—Application to Medicare Enrollees Diagnosed with Colorectal Cancer. *Medical Care* 37(12): 1249-1259, 1999.
- Brown, M. L., Riley, G. F., Schussler, N., et al.: Estimating Health Care Costs Related to Cancer Treatment from SEER-Medicare Data. *Medical Care* 40(8): 104-117, 2002.
- Card, D. and Krueger, A. B.: Minimum-Wages and Employment—A Case-Study of the Fast-Food Industry in New-Jersey and Pennsylvania. *American Economic Review* 84(4): 772-793, 1994.

- Delco, F., Egger, R., Bauerfeind, P., et al.: Hospital Health Care Resource Utilization and Costs of Colorectal Cancer During the First 3-Year Period Following Diagnosis in Switzerland. *Alimentary Pharmacology & Therapeutics* 21(5): 615-622, 2005.
- Deyo, R. A., Cherkin, D. C., and Ciol, M. A.: Adapting a Clinical Comorbidity Index for Use with ICD-9-CM Administrative Databases. *Journal of Clinical Epidemiology* 45: 613-619, 1992.
- Duan, N.: Smearing Estimate—A Nonparametric Retransformation Method. *Journal of the American Statistical Association* 78(383): 605-610, 1983.
- Etzioni, R., Riley, G. F., Ramsey, S. D., et al.: Measuring Costs—Administrative Claims Data, Clinical Trials, and Beyond. *Medical Care* 40(6): 63-72, 2002.
- Finkelstein, E. A., Bray, J. W., Chen, H., et al.: Prevalence and Costs of Major Depression among Elderly Claimants with Diabetes. *Diabetes Care* 26(2): 415-420, 2003.
- Gallo, J. J., Armenian, H. K., Ford, D. E., et al.: Major Depression and Cancer: The 13-Year Follow-Up of the Baltimore Epidemiologic Catchment Area Sample (United States). *Cancer Causes Control* 11(8): 751-758, 2000.
- Graubard, B. I. and Korn, E. L.: Predictive Margins with Survey Data. *Biometrics* 55(2): 652-659, 1999.
- Hershman, D., Hall, M. J., Wang, X., et al.: Timing of Adjuvant Chemotherapy Initiation after Surgery for Stage III Colon Cancer. *Cancer* 107(11): 2581-2588, 2006.
- Klabunde, C. E., Potosky, A. L., Legler, J. M., et al.: Development of a Comorbidity Index Using Physician Claims Data. *Journal of Clinical Epidemiology* 53: 1258-1267, 2000.
- Manning, W. G.: The Logged Dependent Variable, Heteroscedasticity, and the Retransformation Problem. *Journal of Health Economics* 17(3): 283-295, 1998.
- Manning, W. G., Basu, A. and Mullahy, J.: Generalized Modeling Approaches to Risk Adjustment of Skewed Outcomes Data. *Journal of Health Economics* 24(3): 465-488, 2005.
- Midgley, R. and Kerr, D. J.: Adjuvant Chemotherapy for Stage II Colorectal Cancer: The Time Is Right! Nature Clinical Practice. *Oncology* 2(7): 364-369, 2005.
- Mullahy, J.: Much Ado about Two: Reconsidering Retransformation and the Two-Part Model in Health Econometrics. *Journal of Health Economics* 17(3): 247-281, 1998.
- Murray, L. A. and Eppig, F. J.: Health Expenditures for Medicare Beneficiaries. *Health Care Financing Review* 21(2): 281-286, 1999.
- Murray, L. A. and Shatto, A. E.: Dually Eligible Medicare Beneficiaries. *Health Care Financing Review* 20(2): 131-140, 1998.
- Polsky, D., Doshi, J. A., Marcus, S., et al.: Long-Term Risk for Depressive Symptoms after a Medical Diagnosis. *Archives of Internal Medicine* 165(11): 1260-1266, 2005.
- Ward, E., Halpern, M., Schrag, N., et al.: Association of Insurance with Cancer Care Utilization and Outcomes. CA: *A Cancer Journal for Clinicians* 58(1): 9-31, 2008.
- Warren, J. L., Yabroff, K. R., Meekins, A., et al.: Evaluation of Trends in the Cost of Initial Cancer Treatment. *Journal of the National Cancer Institute* 100(12): 888-987, 2008.
- Wright, G. E., Barlow, W. E., Green, P., et al.: Differences among the Elderly in the Treatment Costs of Colorectal Cancer: How Important Is Race? *Medical Care* 45(5): 420-430, 2007.
- Yabroff, K. R., Mariotto, A. B., Feuer, et al.: Projections of the Costs Associated with Colorectal Cancer Care in the United States, 2000-2020. *Health Economics: Early View*, 2007a.
- Yabroff, K. R., Warren, J. L., and Brown, M. L.: Costs of Cancer Care in the USA: A Descriptive Review. *Nature Clinical Practice. Oncology* 4(11): 643-656, 2007b.

---

Reprint Requests: Zhehui Luo, B601 West Fee Hall, Department of Epidemiology, Michigan State University, East Lansing, MI 48824. E-mail: zluo@msu.edu