COMPARING ORAL AND PHARYNGEAL CANCER RATES IN RURAL AND URBAN AREAS

BY

Catherine M. Womack

Submitted to the graduate degree program in Clinical Research
and the Graduate Faculty of the University of Kansas
in partial fulfillment of the requirements for the degree of
Master in Science

Dr. Jonathan Mahnken
Chairperson

Committee members: Dr. Sue-Min Lai
Dr. John Keighley

Date defended: December 3, 2008

The Thesis Committee for Catherine M. Womack certifies
that this is the approved Version of the following thesis:

COMPARING ORAL AND PHARYNGEAL CANCER RATES IN RURAL AND URBAN AREAS

Committee:

Jonathan D. Mahnken
Chairperson

Date approved: December 18, 2008

Abstract

Background: Risk factors for oral and pharyngeal cancers include tobacco use, alcohol use, poor diet, HPV infection, poor oral care, low socio-economic status, gender and genetics, and age. This analysis aims to discover whether or not differences exist in incidence and survival rates in oral and pharyngeal cancer patients in rural and urban areas.

Methods: Surveillance Epidemiology and End Results (SEER) data from 17 registries for the years 2000-2005 was used for this analysis. A Poisson regression and Survival analysis were performed.

Results: Rural or urban residency was not significant in either analysis. Race, gender, and age were all significant at the 0.05 level.

Conclusions: The dataset for this analysis was limited to variables in the SEER data and population data sets. Known risk factors could not be accounted for in this analysis, which could have had an impact on the results, especially in rural and urban differences.

Table of Contents

Tables and Figures

Introduction

Oral and pharyngeal cancers include cancers of the lip, tongue, salivary gland, floor, gum, and other parts of the mouth, nasopharynx, tonsil, oropharynx, hypopharynx, and other sites in the oral cavity and pharynx[1,2]. The overall incidence rate for oral and pharyngeal cancers of males and females of all races and ethnicities is 10.1 per 100,000[2]. In the United States, in 2004, the age-adjusted oral and pharyngeal cancer incidence rate for males was 15.7 per 100,000 (21,396 cases), and the age-adjusted rate for females was 5.9 per 100,000 (9,424 cases). The age-adjusted death rate from oral-pharyngeal cancers was 4.0 per 100,000 (5,312 cases) for males and 1.5 per 100,000 (2,514 deaths) for females. Oral-pharyngeal cancer had the 8th highest incidence rate among males in 2004. The incidence rate was not among the top ten most diagnosed primary sites for females in 2004[3].

Estimated age adjusted prevalence rate is 5.6 per 100,000 for all gender and racial/ethnic groups as of January 1, 2005. Age adjusted prevalence of oral and pharyngeal cancers is highest in white males (8.3 per 100,000) and lowest in Hispanic females (2.2 per 100,000). Overall, Hispanics have the lowest prevalence rates (3.1 per 100,000) when compared with whites (5.7 per 100,000) African Americans (4.4 per 100,000) and Asian/Pacific islanders (5.1 per 100,000)[2].

Mortality and incidence rates of oral and pharyngeal cancers have been steadily declining since the 1970's. The percentage change in age adjusted oral and pharyngeal cancer mortality decreased by 0.5% between 1975 and 1979, 1.7% between 1979 and 1993, 2.7% between 1993 and 2000, and 1.3% between 2000 and 2005[2]. Incidence rates have been declining as well, with the average annual percentage change in incidence rates decreasing 1.2% between the years of 1996 and 2000, as well as the years 2000 through 2005[2].

Survival rates for oral and pharyngeal cancers are largely dependent on the stage of the tumor at diagnosis. Patients with pharyngeal cancers are typically diagnosed with later stage carcinomas (T3-T4), especially when compared with patients diagnosed with oral cancers, who were more likely to be diagnosed in earlier stages[4]. However, no significant difference in tumor size was found in patients with pharyngeal cancers when compared with oral cancers[5]. A delay in seeking medical attention after the onset of symptoms is also greater in patients diagnosed with pharyngeal cancers when compared with oral cancers (45 days and 28 days, respectively)[4]. Most patients who delayed medical attention attributed their symptoms to an infection[4].

Detection of asymptomatic lesions is more likely to occur in the office of a dental professional[5]. Patients diagnosed at a non-symptom driven appointment were more likely to have a significantly smaller lesion ($p=0.033$) and a lesser stage ($p=0.007$)[5]. Patients diagnosed in a symptom-driven appointment by a dental professional were less likely to have metastases of the cervical area at time of diagnosis[5].

The survival rates of patients with oral and pharyngeal cancers steadily decline in the 5 years after initial diagnosis. Survival is highest in the first 12 months, at 79%, followed by a 24 month rate of 60%, a 36 month rate of 46%, a 48 month rate of 40%, and a 60 month rate of 39%[6]. Stage III and IV patients had the lowest 5 year survival rates (34% and 20%), especially when compared with early stage diagnosis. The stage one five-year survival rate is as high as 89%[6].

Risk factors for oral-pharyngeal cancers include tobacco use (both smoking and smokeless tobacco), alcohol use, poor diet, HPV infection, poor oral care and tooth loss, low socio-economic status, gender and genetics, and older age[7].

Literature Review

Alcohol and Tobacco Use

In industrialized areas such as Europe and the United States, it has been well-established that tobacco use and alcohol consumption account for roughly 75% of all cases of oral and pharyngeal cancers[8]. Tobacco and alcohol use commonly occur together, which makes it is difficult to attribute risk to either alcohol or tobacco alone. An international pooled analysis found that among smokers who have never used alcohol, the risk of oral and pharyngeal cancer is 2.13 times higher than persons who have never used alcohol or smoked[8]. The risk among those who have never smoked was higher among women when compared with men (2.33 vs. 1.65). However, this difference was not found to be statistically significant[8].

Among users of alcohol who have never smoked, the there was no significant association between alcohol use and oral and pharyngeal cancers. A statistically significant higher risk was detected, however, among those who never smoked and consumed three or more alcoholic drinks per day (OR 2.04)[8].when compared with those who had never smoked and consumed alcohol. The authors concluded that roughly 24% of oral and pharyngeal cancers could be attributed to tobacco use among those who have never consumed alcohol, while approximately

7% of cases can be attributed to alcohol use among those who have never used tobacco[8].

Alcohol

Alcohol use is a risk factor associated with oral and pharyngeal cancers. Consumption of all types of alcoholic beverages increases a person's risk for oral and pharyngeal cancer, however the type of spirit does impact the level of risk. As total alcohol consumption increases so does the risk for oral and pharyngeal cancers[9]. Those who consume 3-4 drinks a day are at a 2.1 times higher risk for oral and pharyngeal cancer, 5-7 drinks 5 times higher, 8-11 drinks 12.2 times higher, and 12 or more drinks a day 21.1 times higher risk. There is a significant trend across the levels of drinking ($p<0.0001$)[9].

However, among those who consume beer or spirits, and no wine, the increase in risk is markedly lower, with the highest risk being for those who consume three or more beers a day, with a 2.3 times higher risk. Wine has the most significant single effect on oral pharyngeal cancer risk, with those who consume 3-4 drinks a day are at a 2.2 times higher risk, 5-7 drinks 7.1 times higher, 8-11 drinks 11.8 times higher, and 12 or more drinks a day 16.8 times higher risk. There is again a significant trend across the levels of drinking ($p<0.0001$)[9].

Those in rural areas are more likely to have a current alcohol related disorder (OR=1.20, 95% CI 1.04 , 1.39) and exceed their specified daily limit for alcohol consumption (OR=1.14, 95% CI 1.02 , 1.27) compared with their urban and suburban counterparts[10].

Tobacco

Both smoking and smokeless tobacco use are risk factors for oral and pharyngeal cancers. The prevalence of smoking among adults in the United States was 20.8% (45.3 million people) in 2006, according to the National Health Interview Survey[11]. Among current smokers, 80% smoked every day, while the remaining 20% smoked only some days. Prevalence of smoking was lower among women (18%) compared with men (24%). Adults ages 18-44 had the highest rates of current smoking. In 2005, approximately 2.3% of American adults used smokeless tobacco[12], with the highest rate among males.

In 2000, the rate of smokeless tobacco use was 9.0% in rural (non MSA) area, while it was 3.3% in urban (MSA) areas[13]. Among women, the prevalence of smokeless tobacco use was only 0.3%, compared with a prevalence of 4.5% in males. Males ages 18-44 were most likely to use smokeless tobacco, over 5% of the population[13]. White males were three times more likely to use smokeless tobacco when compared with African American and Hispanic males[13].

Use of smokeless tobacco has been shown to increase the risk of oral cancer up to 2.6 times in a pooled analysis from the United States and Europe[14]. However, other studies have shown a less elevated risk. The odds of developing oral or pharyngeal cancer was found to be 1.94 times higher in users of smokeless tobacco when compared with never users[15].

Diet and Nutrition

Poor diet, namely low consumption of fruits and/or vegetables, is another risk factor for oral and pharyngeal cancers. Consumption of fruit and vegetables has been found to decrease the risk of developing oral cancer. Fruit (OR=0.4) and vegetables (OR=0.2) were found to cause a significant decrease in oral cancer risk, especially among citrus fruits and juices[16]. Consumption of milk was also found to have a protective effect (OR=0.38)[17]. Intake of red meats, pork, and processed meats are associated with a significantly increased risk of oral and pharyngeal cancers. Among people who consume more than 3.5 servings of red meat a week, the odds of oral and pharyngeal cancer are 2.14 times higher when compared with people who do not consume red meat. Consuming more than 4.5 servings or pork or processed meat per week increased the chances of oral cancers by 3.21 times[17].

Among adults in rural areas, very few have diets that meet the Healthy Eating Index (HEI) standards. Dietary guidelines are not being adhered to, with

24% having a poor diet, 75% needing improvement, and only 1% meeting dietary guidelines[18]. This disparity is especially true among older adults with oral health problems[19].

Human Papilloma Virus Infection and Exposure

Current research is increasingly showing a link between infection with the human papilloma virus and oral and pharyngeal cancers. A significant association has been found between cases of oral and pharyngeal cancers and exposure to HPV-16 over the course of a person's lifetime (OR=32.2)[20]. Among those with an oral HPV-16 infection present, there is a strong association with oral cancers (OR=14.6)[20]. Infection with any of 37 types of HPV also greatly increases the odds of oral cancer (OR=12.3)[20].

Smoking and drinking alcohol in concert with exposure or infection to HPV-16 is associated with greater odds of oral and pharyngeal cancers. Among those with a positive HPV-16 serum test who smoke and drink alcohol, the odds of oral and pharyngeal cancer are greatly increased (OR=44.8)[20]. The association between persons who have had an oral HPV infection and oral and pharyngeal cancers was also greatly increased (OR=43.7)[20]. Rates of HPV infection are significantly higher in rural areas when compared with urban areas (10.2 per 100,000 compared with 8.4 per 100,000)[21].

Of potentially HPV associated cancers of the oral cavity and pharynx, the majority were in the tonsil (43.6%), followed by the base of the tongue (38.4%), with the remaining 18% from the rest of the oral cavity and pharynx[21]. Of these cases, the age-adjusted incidence rates were higher in men (three times greater risk), and the age-adjusted incidence rates increased for both men and women combined between the years 1998-2003, with an annual percentage change of 3% in cases of the tonsil and base of the tongue[21].

Dentition and Oral Care

Oral care and associated problems with dentition may also be another risk factor for oral cancers. People who brush their teeth less than once per day have an increased risk of 3.2 compared with those who brush their teeth at least twice daily[16]. Denture wearers are six times more likely to suffer from oral cancer when compared with those who brush their teeth at least twice daily[16]. Frequency of dental visits and oral check-ups was found to have a protective effect. Those who have never visited a dentist were 12 times more likely to suffer from oral and pharyngeal cancer compared with people who had at least yearly dental visits[16].

Tooth loss and missing teeth is also associated with a higher risk for oral cancers. For people missing between 6 and 15 teeth, the risk of oral cancer is seven times higher than those with intact teeth[16]. Those missing 16 or more teeth are have nearly 10 times higher risk for oral cancer[16]. Alveolar bone loss in the tooth is also associated with a higher risk for tongue cancers. Every millimeter of alveolar bone loss leads to a 5.23 increase in risk for cancer of the tongue[16]. Mean measures of alveolar bone loss was found to be higher in cancer cases when compared with controls ($p<0.001$)[16].

Among those living in rural areas, the prevalence of unmet dental needs is higher in rural areas compared with urban areas (10.8% compared with 9.8%)[23]. The percent of people who report poor dental health is also higher in rural areas when compared with urban areas (38.2% compared with 31.8%)[23]. Also, the percent of people who report having a dental care visit in the past year is lower in rural areas when compared with urban areas (58.3% compared with 65.8%)[23].

Socio-economic Level

Socio-economic status is another possible risk factor for oral and pharyngeal cancers. The risk for those of low socio-economic status was 2.41 times higher when compared with those of high income[24]. Low socioeconomic occupation was also a risk factor for oral and pharyngeal cancers. People who fall into the lower socioeconomic occupational group are at 1.84 times greater risk of oral and pharyngeal cancers[24]. Standardized incidence rates are the highest among areas with the highest levels of poverty when compared with the least impoverished areas for both males (120 per 100,000 and 79 per 100,000, respectively) and females (108 per 100,000 and 87 per 100,000, respectively)[25].

Rural residents have a lower socioeconomic status than their urban counterparts. The United States Economic Research service reports that among rural residents, 14.2% live below the poverty threshold, compared with 12.1% of urban residents[26]. Disparities in household incomes also exist between rural and urban residents. Urban residents have a median income $13,500 higher than rural residents[27].

Gender and Genetic Factors

The incidence rate of oral and pharyngeal cancers among men and women is nearly 4:1[28]. There is increasing evidence to suggest that female hormones may play a role in oral and pharyngeal cancers. Gallus et al found a significant trend between age of menopause onset and risk for oral cancer (P<0.01). Women who experience menopause later in life (>50 years of age) have a decreased risk for oral and pharyngeal cancer[29]. Among cases of oral and pharyngeal cancer, the odds were 2.36 times higher for a woman to experience early menopause (<= 45 years of age)[28].

Among cases of oral and pharyngeal cancer, the risk is 2.6 times higher when a family member also has oral and pharyngeal cancer. The risk jumps to 7.1 times higher when two or more family members are affected, whether the relatives are immediate family or not[30]. Among current smokers who consume 21 or more alcoholic drinks per week with a family history of oral and pharyngeal, the risk is 46.2 times higher when compared with a nonsmoker nondrinker with no family history[30].

The primary aim of this research is to determine whether there are differences in incidence and survival rates between rural and urban cases of oral and pharyngeal cancers. As of this writing, there was no research that specifically looked at rural and urban disparities in oral and pharyngeal cancer rates. Given the aforementioned risk factors for oral and pharyngeal cancers, the majority of which show disparities between rural and urban areas, it is hypothesized that there would be differences in incidence and survival rates between rural and urban oral and pharyngeal cancer patients. Specifically, since rural residents have poorer outcomes among all risk factors than their urban counterparts, they should show higher rates of oral and pharyngeal cancers and shorter rates of survival.

Methods

A literature review of current published articles was performed using PubMed. An initial search of oral and pharyngeal cancer rates was performed to determine the underlying oral and pharyngeal cancer rates in the United States population. Data analysis from the SEER registries and the United States Cancer Statistics from the National Program of Cancer Registries was also examined to determine population based oral and pharyngeal cancer incidence and survival rates. A literature search was performed using the terms oral and pharyngeal cancer with the combinations rural and urban and metropolitan and non-metropolitan areas. No articles were found with research from the United States in respect to survival rates or incidence rates of oral and pharyngeal cancer patients.

Subsequent searches were performed for risk factors for oral and pharyngeal cancer, and the primary risk factors were determined to be tobacco use (both smoking and smokeless tobacco), alcohol use, poor diet, HPV infection, poor oral care and tooth loss, low socio-economic status, gender and genetics, and older age. Cross searches were performed to determine if disparities existed between risk factors and rural or urban residency status.

Since the research into the risk factors for oral and pharyngeal cancer rates showed disparities between rural and urban populations, the aim of this research was to examine whether or not disparities exist between rural and urban populations in respect to oral and pharyngeal cancer incidence and survival rates.

Data from the Surveillance Epidemiology and End Results Program (SEER) was used for this analysis. SEER is a population-based data source for cancer cases in the United States. The data includes cases from the 17 SEER registries and encompasses years 2000 through 2005. During the six-year span between 2000 and 2005, 47,136 new cases of cancer of the oral cavity and pharynx were diagnosed in SEER registry areas[2]. Population data was obtained through the SEER program data estimates based on 2000 United States census data[2].

The race/ethnicity variable includes non Hispanic White, non Hispanic Black, Hispanic, non Hispanic American Indian, and non Hispanic Asian/Pacific Islander. Anyone with Hispanic origin was considered Hispanic in this analysis.

Rural or urban designation was derived from the Federal Information Processing Standards (FIPS) rural-urban continuum coding system, using the rural designation of the United States Department of Agriculture[31]. Rural areas have a population of less than 2,500, whether in open country or in settlements. Rural-urban continuum codes are shown in table 1.

Table 1: Rural and Urban Continuum Codes

Rural-Urban Continuum Codes	
Code	**Description**
Metro counties:	
1	Counties in metro areas of 1 million population or more
2	Counties in metro areas of 250,000 to 1 million population
3	Counties in metro areas of fewer than 250,000 population
Nonmetro counties:	
4	Urban population of 20,000 or more, adjacent to a metro area
5	Urban population of 20,000 or more, not adjacent to a metro area
6	Urban population of 2,500 to 19,999, adjacent to a metro area
7	Urban population of 2,500 to 19,999, not adjacent to a metro area
8	Completely rural or less than 2,500 urban population, adjacent to metro
9	Completely rural or less than 2,500 urban population, not adjacent to metro

from http://www.ers.usda.gov/Briefing/Rurality/WhatIsRural/

The designations metro counties (rural urban continuum codes 1 through 3), and nonmetro counties adjacent or nonadjacent to a metro area with a population of 2,500 or more (codes 4 through 7) were considered urban, while counties with a population of less than 2,500 were considered rural (codes 8 and 9).

SEER registry identification number tells which of the seventeen SEER registries the case originated from. The SEER registries are geographically located across the United States to include population subgroups that are concentrated in a specific area. The SEER registries encompass roughly 26% of the United States population, with representative subgroups of different race and ethnicity designations[2]. SEER registry locations are shown in figure 1.

from http://seer.cancer.gov/registries/

Figure 1: SEER Registry Locations

.

Age groups were divided into 10 year groupings, consistent with the age groupings used by the SEER program. The age groups included were under 35 years of age, 35-44 years of age, 45-54 years of age, 55-64 years of age, 65-74 years of age, and older than 75 years of age. Age-adjusted rates by race and ethnicity were calculated for all races and ethnicities using direct age adjustment. The United States population from the 2000 census was used as the standard population.

Analysis

All analyses were run using SAS. Variables in the incidence rate dataset include oral and pharyngeal cancer cases, gender, race/ethnic group, SEER registry identification number, rural or urban designation, and age group at diagnosis. Oral and pharyngeal cancer cases were determined using the SEER site recode variable, which recodes the ICD-O-3 cancer locations. Sites are shown in table 2.

Table 2: Oral and Pharyngeal Cancer Sites

Oral Cavity and Pharynx	SEER Site Code
Lip	20010
Tongue	20020
Salivary Gland	20030
Floor of Mouth	20040
Gum and Other Mouth	20050
Nasopharynx	20060
Tonsil	20070
Oropharynx	20080
Hypopharynx	20090
Other Oral Cavity and Pharynx	20100

Frequency counts for oral and pharyngeal cancer cases were created for each combination of gender, race, SEER registry, rural or urban residency status, and age group. These counts were then merged with frequency counts of population data to determine the incidence rates of oral and pharyngeal cancers by population counts among SEER registry populations.

Poisson regression analysis was performed on this data to model the rate of oral and pharyngeal cancers. The Poisson model is a log-linear model that models the incidence rate of oral and pharyngeal cancers in the SEER registry area population. The primary variable of interest is the rural/urban residency status. Secondary variables of interest are race, gender, and age group. SEER registry identification number was also included in the models to account for any possible geographic effects, but was not included in the results as a variable of interest.

An initial Poisson model was run that included only the rural or urban residency status variable. A main effects model was run including all variables of interest (race, age, gender, rural or urban residency, SEER registry), to determine possible effects of secondary variables of interest on the primary outcome of interest. After model diagnostics were performed, a final Poisson regression model was run that included both main effect and interaction terms. Backward elimination including all possible interaction terms up to the four-way interaction between race, age, gender, and rural/urban residency was performed to determine the best possible model to fit the data. A 0.15 level p-value was used to determine inclusion in the model.

A survival rate data set included the variables age group, race, gender, rural or urban residency status, SEER registry identification number, and survival time for cases from the year 2001. Survival time was computed in months, using the SEER data set variable survival time recode, with times ranging from zero to 60 months. Subjects who survived the five year period were censored. Survival was determined from the SEER variable survival recode. All cause mortality was investigated in this analysis, using the SEER variables for survival and survival time.

Survival analysis was performed on the data, to determine if the rate of survival was different across the different groups. Five year survival data was computed in months for oral and pharyngeal cancer cases. Variables included in the survival analysis were rural or urban residency status, age, gender, and race. Kaplan-Meier curves were created to compare the survival functions across the age, gender, race, and rural/urban groups, and a log rank test was used to determine if there were significant differences in the strata. A Cox Proportional Hazards regression model was run to determine the hazard ratios for the different groups, initially including only the rural or urban residency variable. A main effects Cox Proportional Hazards model was run including all variables of interest (race, gender, age group, rural or urban residency, SEER registry). Model diagnostics were performed to check for violations of the proportional hazards assumptions.

Results

Baseline characteristics of SEER oral and pharyngeal cancer cases are shown below in table 3. Thirty-three SEER cases did not have proper FIPS coding and were not included in the baseline demographics. Eight cases were missing age group (less than 1% of the population), and 466 cases were missing race/ethnicity data (roughly 1.2% of the population). Gender and age distributions are similar at baseline in the rural and urban groups. Racial and ethnic distributions are not evenly distributed, with the majority of cases occurring among whites.

Unadjusted (crude) incidence rates by age group are shown below in table 3. Those in the 34 years and younger group have the lowest incidence rates, while those in the 75 years of age and older group have the highest incidence rates. The incidence rate increases across the age groups in both rural and urban areas, and is higher in males, which is consistent with current research.

Table 3: Baseline Characteristics for SEER cancer cases

Baseline Characteristics SEER cases								
		Urban (n=46464)			Rural (n=639)			
		n	%	incidence per 100K	n	%	incidence per 100K	
Age	<34	1407	0.3	0.62 (0.59,0.652)	13	2.0	0.65 (-0.25,1.55)	
	35-44	3305	7.1	4.69 (4.53,4.85)	44	6.9	6.87 (1.74,12.00)	
	45-54	9571	20.6	15.52 (15.21,15.83)	132	20.7	20.30 (11.55,29.05)	
	55-64	11659	25.1	29.25 (28.72,29.78)	143	22.4	29.20 (17.07,41.29)	
	65-74	10147	21.8	39.17 (38.41,39.93)	146	22.8	40.27 (23.74,56.80)	
	75+	10367	22.3	42.06 (41.25,42.87)	161	25.2	46.98 (28.63,65.32)	
Gender	Male	31600	68.0	14.27 (14.11,14.43)	461	72.1	20.67 (15.90,25.44)	
	Female	14864	32.0	6.56 (6.45,6.67)	178	27.9	7.87 (4.95,10.79)	
Race	White	35311	76.0	13.50 (13.36,13.64)	605	94.8	14.92 (11.92,17.92)	
	Black	4366	9.4	9.01 (8.74,9.28)	18	2.8	12.67 (-2.17,27.51)	
	Hispanic	3159	6.8	3.38 (3.26,3.50)	6	0.9	5.02 (-5.35,15.39)	
	Am. Indian	197	0.4	4.94 (4.25,5.63)	5	0.8	3.02 (-3.60,9.64)	
	Asian	2966	6.4	7.34 (7.06,7.60)	4	0.6	33.49 (-51.25,118.23)	
	missing	465	1.0		1	0.2		

Due to the disproportionate distribution of races, incidence rates in the different race/ethnic groups are more difficult to determine. Without stratifying by rural/urban status, the highest age-adjusted incidence is in blacks, with 11.44 per 100,000, followed by whites with 11.40 per 100,000, Asian/Pacific Islanders 8.14 per 100,000, American Indians with 7.24 per 100,000, and the lowest incidence rate is among Hispanics with 6.38 per 100,000.

Table 4: Age-Adjusted Incidence Rates

	Overall	Rural	Urban
	n=47103	n=639	n=46464
White	11.40 (11.11, 11.66)	12.42 (10.05,14.78)	11.72 (11.42,12.00)
Black	11.44 (10.64,12.23)	14.35 (0.57,28.64)	11.43 (10.63,12.22)
Hispanic	6.38 (5.86,6.89)	6.19 (-0.37,12.76)	6.37 (-0.027,12.78)
American Indian	7.24 (6.76,7.71)	3.51 (-2.25,9.54)	6.34 (4.27,8.41)
Asian/Pacific Islander	8.14 (7.43,8.84)	34.83 (-13.4,83.11)	8.13 (7.43,8.83)

Poisson Analysis

In the initial Poisson regression analysis, residency status was highly significant (p<0.0001), with a rate 1.38 times higher for rural residents. The incidence rate ratios from the main effects Poisson Regression model (age group, race, gender, rural or urban residency) are shown in table 3. Type III sums of squares analysis shows that all variables of interest are significant at the 0.0001 level, excluding residency status, which was not statistically significant (p=0.54).

Table 5: Risk Ratios Main Effects Model

		Rate Ratio	p value	95% CI
Age	<34	1.00	-	
	35-44	7.31	**<0.0001**	6.869-7.788
	45-54	23.70	**<0.0001**	22.404-25.072
	55-64	44.20	**<0.0001**	41.806-46.729
	65-74	60.39	**<0.0001**	57.085-63.883
	75+	68.18	**<0.0001**	64.669-71.889
Gender	Male	1.00	-	
	Female	0.39	**<0.0001**	0.382-0.398
Race	White	1.00	-	
	Black	1.01	0.6485	0.975-1.041
	Hispanic	0.52	**<0.0001**	0.504-0.543
	Am. Indian	0.54	**<0.0001**	0.466-0.630
	Asian	0.68	**<0.0001**	0.656-0.712
MSA	Urban	1.00	-	
	Rural	1.02	0.54	0.457-2.030

The Type III sums of squares gives a likelihood ratio test of the predictors, and given those that are statistically significant, there is a good chance that each is a good predictor for the model. A deviance test was performed to examine model fit, and it was found to be highly significant (p=0). This model does not appear to fit well.

In the final Poisson model, the primary variable of interest, residency status, was not statistically significant (p=0.54). Type III sums of squares analysis found that race, age, and the interaction terms race*age, gender*age, and gender*race*age were all significantly different from zero at the 0.0001 level. Gender was also statistically significant from zero (p=0.03).

Given the statistically significant values of the likelihood ratio tests of each predictor, they are most likely a good fit for the model, including the interaction terms. When all the terms are included in the model, each predictor is significant, except for rural/urban residency status (p=0.54). Model parameter estimates are given in appendix A.

A deviance test was performed on the final model including interaction terms, and the test was not significant (p=0.56) so the model including interaction terms appears to be a good fit for the data. Over or under dispersion is not a concern in the final model, given that the deviance value to degrees of freedom ratio is very close to 1 (0.99).

Using the parameter estimates shown in appendix B, to compare the incidence rate for a black female, aged 45-54, residing in a rural area with a black female, aged 45-54 in an urban area, the following equations:

$$IR = e^{\alpha+\beta(female)+\beta(black)+\beta(rural)+\beta(45-54)+\beta(45-54*female)+\beta(45-54*black)+\beta(female*black*45-54)}$$

$$IR = e^{\alpha+\beta(female)+\beta(black)+\beta(45-54)+\beta(45-54*female)+\beta(45-54*black)+\beta(female*black*45-54)}$$

Give the following computations:

$IR = e^{-11.95-0.12+0.021+0.025+3.6-1.095+0.11+0.1377}$ =9.45/100,000 rural black female age 45-54, 95% CI (8.24 , 10.83).

$IR = e^{-11.95-0.12+0.021+3.6-1.095+0.11+0.1377}$ =9.22/100,000 urban black female age 45-54, 95% CI (8.26 , 10.28).

The incidence rate is higher in rural black females ages 45-54 when compared with their urban counterparts. The incidence rate ratio is 1.02.

Survival Analysis

Kaplan Meier curves are shown in figures 2-5. Using the log rank test to test for equality between the strata, five year survival times for rural/urban residency status are not significantly different from each other (p=0.11, rural 5 year survival 59.78%, urban 5 year survival 51.80%). There was a significant difference in survival times between males and females (p=0.005). The survival time for females was longer than males after five years (54.47% compared with 50.63%).

Race groups were statistically different at $p<0.0001$ for the global test across the strata. Five-year survival rates were different for whites (52.49%), blacks (38.19%), Hispanics (53.11%), American Indians (58.14%), and Asian/Pacific Islanders (59.59%), but no inferences can be made between the survival time differences between the different groups given this global test result. Pairwise comparisons were not performed at this time.

Age groups were also statistically different at $p<0.0001$ for the global test across the strata. The highest rate of survival was among those 34 years of age and younger (79.91%), followed by 35-44 year olds (73.10%), 45-54 year olds (64.72%), 55-64 year olds (56.62%), 65-74 year olds (45.68%), and finally 75 years and older (31.74%). However, since pairwise comparisons were not performed, conclusions between the survival rates between the different age groups cannot be drawn.

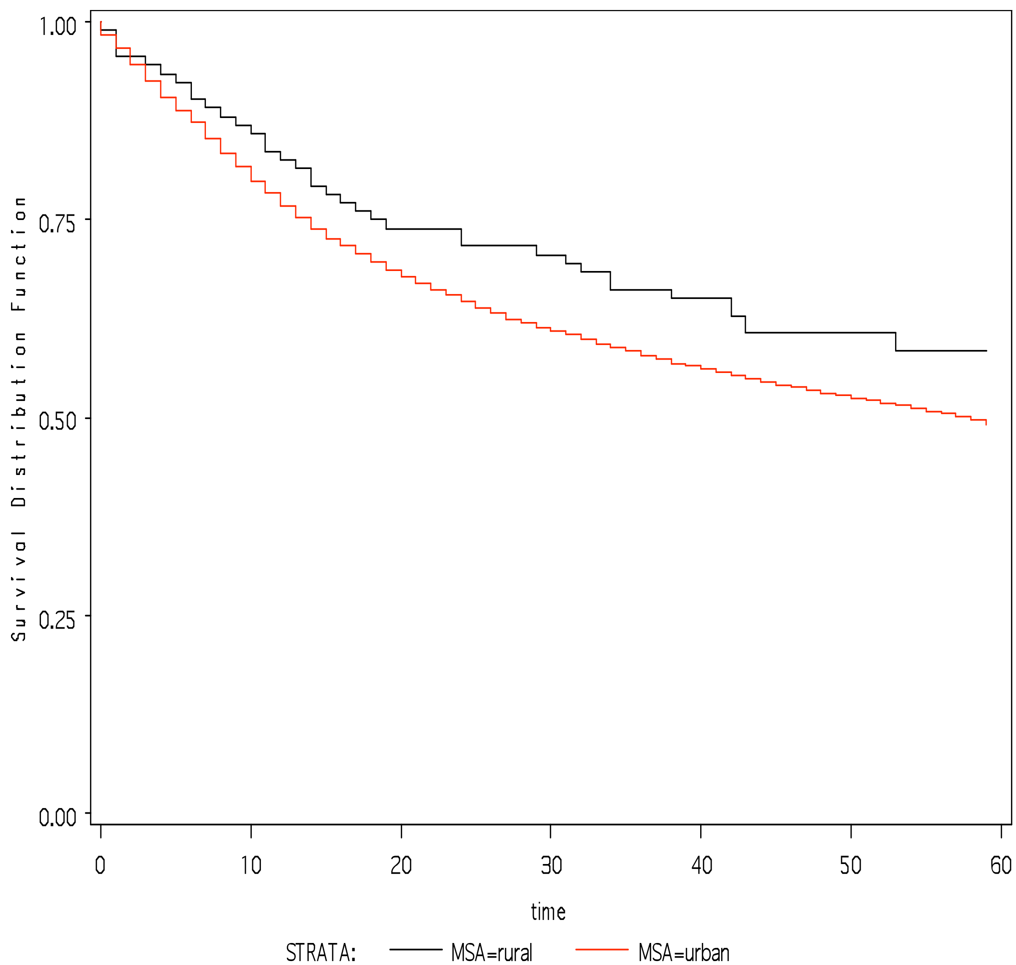


Figure 2: Five year survival rates in rural and urban areas

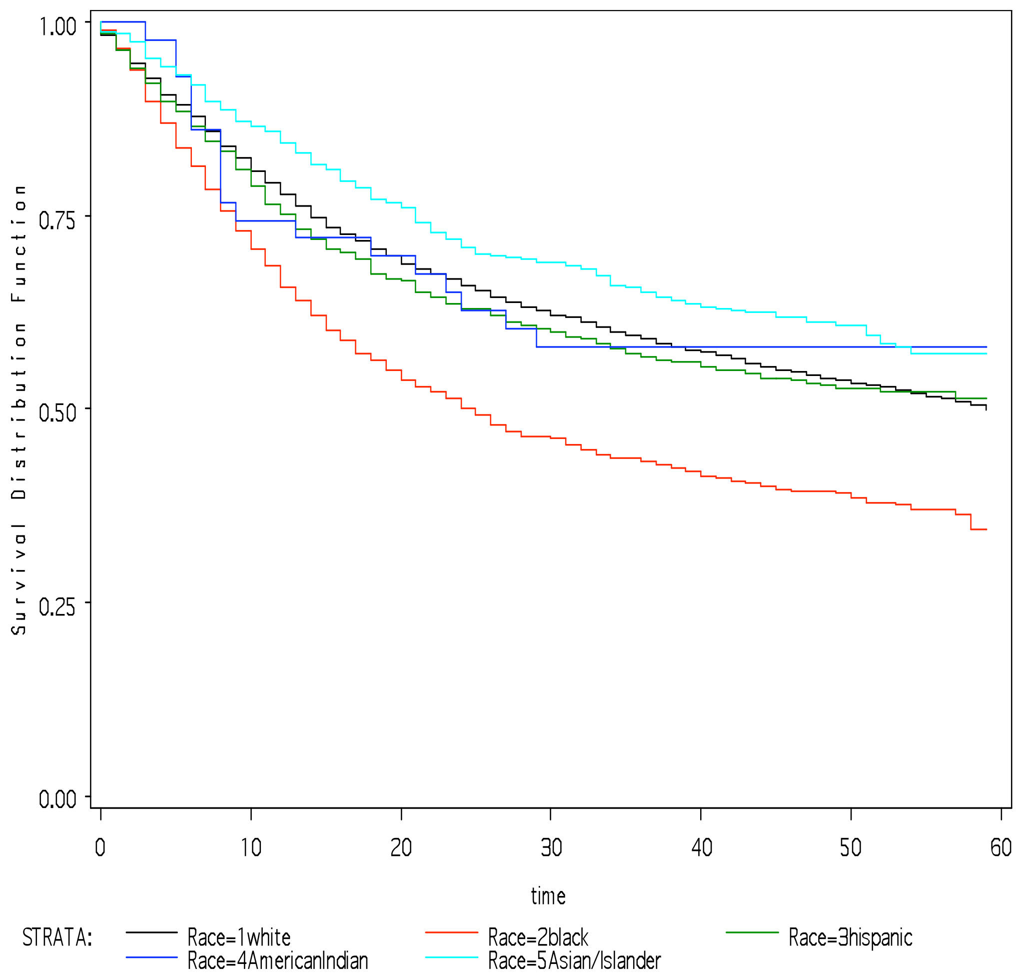


Figure 3: Five year survival rates in different races

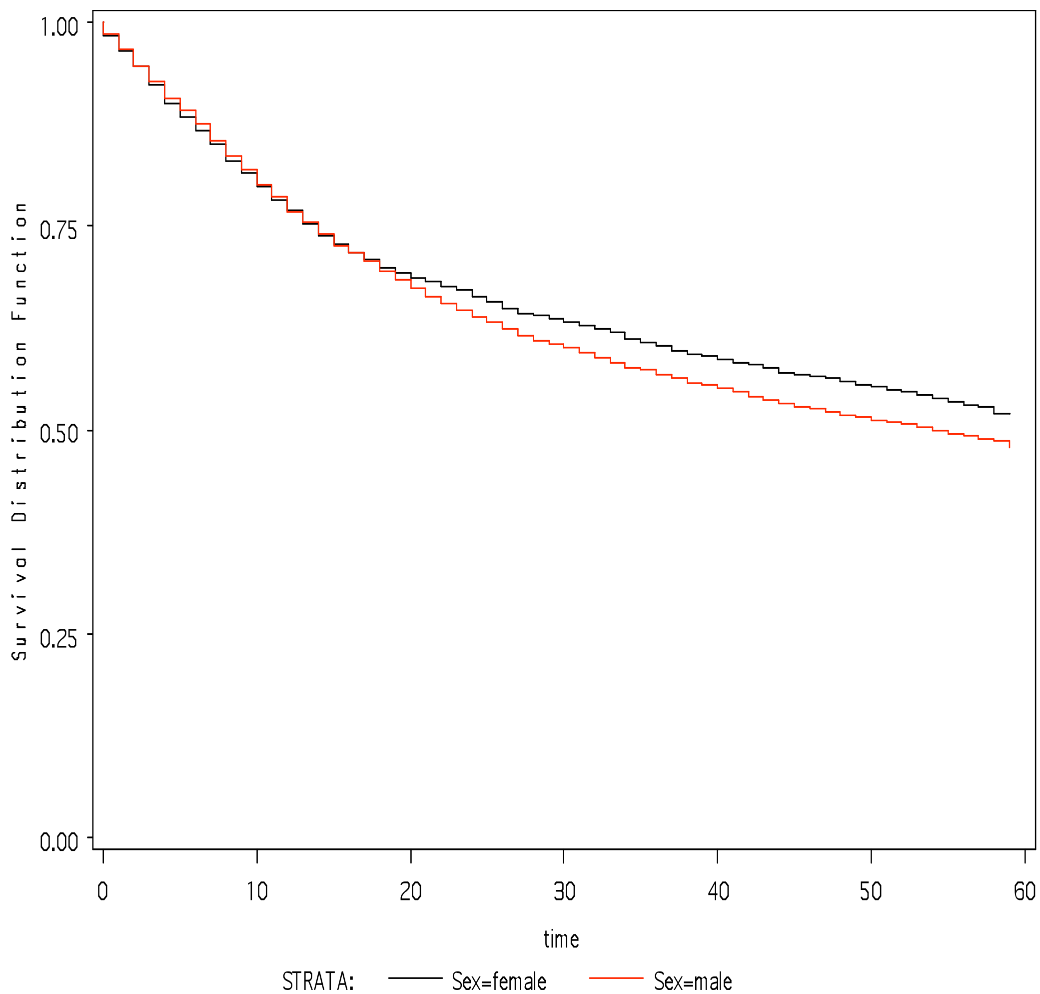


Figure 4: Five year survival rates across gender

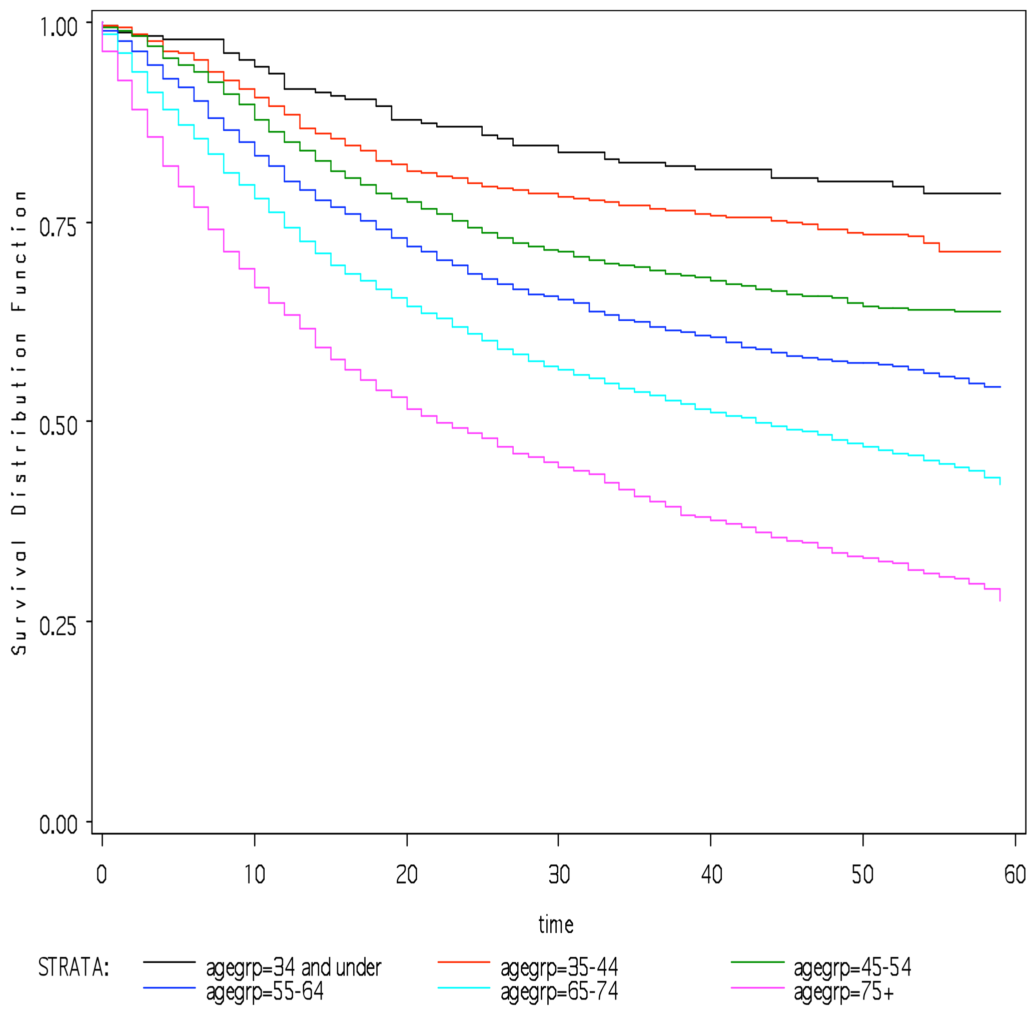


Figure 5: Five year survival rates across age groups

In the initial Cox Proportional Hazards model, rural and urban residency status was not found to be significant (p=0.12). In the main effects Cox Proportional Hazards model, the type III sums of squares showed that age, race, and gender were significant at the <0.0001 level, while rural/urban residency status was not significant (p=0.23). Hazard ratios are shown in table 5.

Table 6: Hazard Ratios Cox Regression Model

		Hazard Ratio	P value	95% CI
Race	White	1.00	-	
	Black	1.75	**<0.0001**	1.576-1.948
	Hispanic	1.20	**0.0074**	1.050-1.371
	Am Indian	1.17	0.56	0.679-2.039
	Asian/Islander	0.94	0.44	0.799-1.104
Gender	Male	1.00	-	
	Female	0.81	**<0.0001**	0.755-0.870
Age	<35	1.00	-	
	35-44	1.37	0.0621	0.984-1.896
	45-54	1.85	**<0.0001**	1.370-2.489
	55-64	2.47	**<0.0001**	1.836-3.312
	65-74	3.43	**<0.0001**	2.555-4.601
	75+	5.43	**<0.0001**	4.053-7.277
Metro	Urban	1.00	-	
	Rural	0.816	0.2312	0.584-1.139

To test the model fit, the correlation between the Schoenfeld residuals and time was tested for all observations that were not censored. The correlations between age (p=0.75), rural/urban residency status (p=0.66), race (p=0.096), and gender (p=0.98) with time were not statistically significant, and thus do not appear to violate the proportional hazards assumption. Residual plots are shown in appendix D.

Adjusted survival function estimates for urban black females age 45-54 (figure 6) and rural black females age 45-54 (figure 7) are shown below.

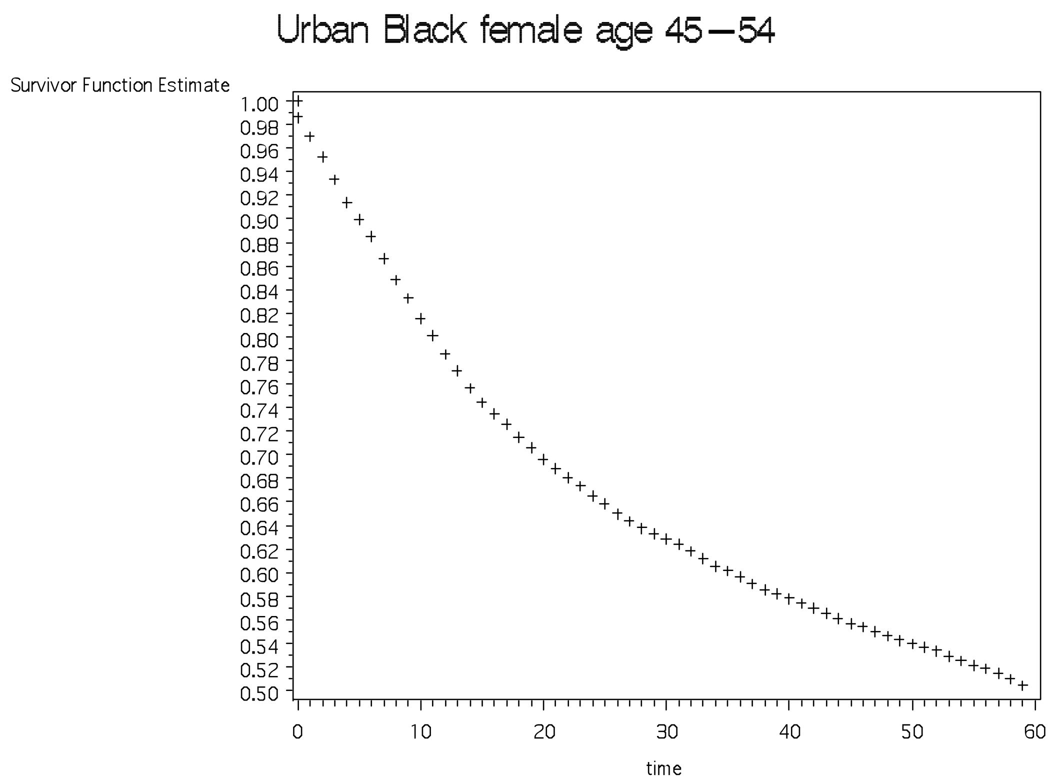


Figure 6: Survival function estimate for urban black female age 45-54

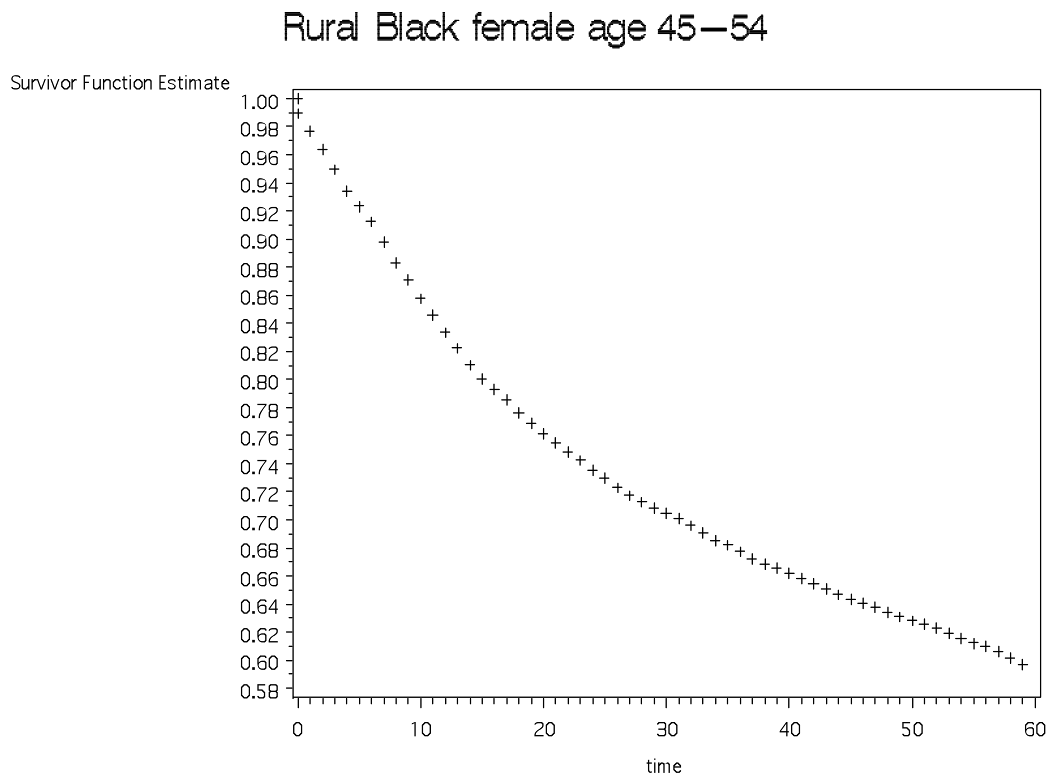


Figure 7: Survival function estimate for rural black female age 45-54

Discussion

The primary variable of interest, rural/urban residency status, was not found to be statistically significant in the final Poisson analysis. This could be due to the limited scope of the data set, and the possibility of confounding variables or covariates not tested in this type of analysis. Secondary variables of interest including race, age, and gender were all found to be statistically significant in the final Poisson model.

Interpretation of the final Poisson regression model is somewhat difficult given the inclusion of interaction terms. Since the interactions of race, gender, and age were all significant in varying combinations, this affects interpretation of the main effects in the final model. The risk ratios for main effects could be inflated due to the nature of the interaction terms, and the main effects cannot be discussed without including the interaction terms.

However, the analysis is consistent with current research in that race, gender, and age group were found to be significant predictors of oral and pharyngeal cancer rates. The significance of the interaction terms between race and age, age and gender, and race, age, and gender, show that the incidence rate for oral and pharyngeal cancers is also affected by the relationship between the risk factors.

The survival analysis data is consistent with current findings in that the five year survival times are not equal across the different strata for gender, age, and race. Survival times are higher in females, which is consistent with other current research. Though not significant, survival rates were found to be higher in rural areas. However, at the time of this writing, no research was available to corroborate this finding. The survival time for age appears to decrease as age increases, however this assumption cannot be confirmed without pairwise comparison data. This finding, although not statistically tested, is consistent with data from the SEER analysis.

From the Proportional Hazards Model, the main effects of gender and age are consistent with current findings. The hazard is lower in females compared with males, and as a person ages, the hazard increases. Hazard ratios for race show that the hazard is lowest in Asian/Pacific Islanders compared with their white counterparts, although this hazard ratio is not statistically significant. The hazard is higher in Hispanics, blacks, and American Indians when compared with whites.

The variable stage was not included in the Proportional Hazards model, which could have had an impact on the results. Since it is known that stage at diagnosis significantly impacts survival time, the inclusion of stage in the model could have possibly changed the significance of the other predictor variables. The exclusion of stage in the Proportional Hazards Model limits the generalizability of the model, especially when considering that stage at diagnosis is a significant predictor of survival time.

The present analysis has some known limitations which could affect the validity of the analysis. The variables included in the Poisson regression model were limited by the nature of the analysis. Since the rate of cancer in the population was modeled, only variables that were present in both the case and population data set were available for analysis. The population data set only includes demographic variables (age, gender, race, SEER registry identification number, and rural/urban residency status), which significantly limited the scope of the analysis. Known risk factors such as tobacco use (both smoking and smokeless tobacco), alcohol use, poor diet, HPV infection, poor oral care and tooth loss, and low socio-economic status could not be included in the model. The limitations of the data set have an obvious impact on the results, since the known factors that contribute to incidence and survival in oral and pharyngeal cancer cases cannot be controlled for. Since all of the known risk factors have been found to have differences between rural and urban areas, the significance of rural and urban residency could have been impacted by the limited data.

However, as stated in the introduction, discrepancies were found between the known risk factors in rural and urban areas. The expectation for these to be carried forth into incidence and survival rates was not found to be statistically significant in this analysis. Further research with a dataset of a much larger scope would be useful in running a more thorough analysis.

References

1. National Cancer Institute. SEER Site Recode ICD-O-3 (1/27/2003) Definition. Available at: http://seer.cancer.gov/siterecode/icdo3_d01272003/. Accessed July 19, 2008.

2. Ries LAG, Melbert D, Krapcho M, Stinchcomb DG, Howlader N, Horner MJ, Mariotto A, Miller BA, Feuer EJ, Altekruse SF, Lewis DR, Clegg L, Eisner MP, Reichman M, Edwards BK (eds). *SEER Cancer Statistics Review, 1975-2005*, National Cancer Institute. Bethesda, MD, http://seer.cancer.gov/csr/1975_2005/, based on November 2007 SEER data submission, posted to the SEER web site, 2008.

3. U.S. Cancer Statistics Working Group. *United States Cancer Statistics: 2004 Incidence and Mortality.* Atlanta: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute; 2007.

4. Brouha X, Tromp D, Hordijk G, Winnubst J, de Leeuw J. Oral and Pharyngeal Cancer: Analysis of Patient Delay at Different Tumor Stages. *Head and Neck.* 2005;27:939–945.

5. Holmes J,Dierks E, Homer L, Potter B. Is Detection of Oral and Oropharyngeal Squamous Cancer by a Dental Health Care Provider Associated With a Lower Stage at Diagnosis?. *J Oral Maxillofac Surg.* 2003;61:285-291.

6. Brandizzi D, Gandolfo M, Velazco M, Cabrini R, Lanfranchi H. Clinical Features and Evolution of Oral Cancer: A Study of 274 Cases in Buenos Aires, Argentina. *Med Oral Patol Oral Cir Bucal.* 2008;13(9):E544-E548.

7. Petti S. Lifestyle Risk Factors for Oral Cancer. *Oral Oncol*. In press.

8. Hashibe M, Brennan P, Benhamou S, et al. Alcohol Drinking in Never Users of Tobacco, Cigarette Smoking in Never Drinkers, and the Risk of Head and Neck Cancer: Pooled Analysis in the International Head and Neck Cancer Epidemiology Consortium. *J Natl Cancer Inst.* 2007;99:777–789.

9. Altieria A, Bosettia C, Gallusa S, et al. Wine, Beer and Spirits and Risk of Oral and Pharyngeal Cancer: A Case–Control Study from Italy and Switzerland. *Oral Oncol.* 2004;40:904–909.

10. Borders T, Booth B. Rural, Suburban, and Urban Variations in Alcohol Consumption in the United States: Findings From the National Epidemiologic

Survey on Alcohol and Related Conditions. *The Journal of Rural Health.* 2007;23(4):314-321.

11. Cigarette Smoking Among Adults — United States, 2006. *Morbidity and Mortality Weekly Report.* 2006;55(42):1145-1168.

12. Tobacco Use Among Adults — United States, 2005. *Morbidity and Mortality Weekly Report.* 2007;56(44);1157-1161.

13. Nelson D, Mowery P, Tomar S, Marcus S, Giovino G, Zhao L. Trends in Smokeless Tobacco Use Among Adults and Adolescents in the United States. *Am J Public Health.* 2006;96(5):897–905.

14. Boffetta P, Hecht S, Gray N, Gupta P, Straif K. Smokeless Tobacco and Cancer. *Lancet Oncol.* 2008;9:667–675.

15. Weitkunat R, Sanders E, Lee P. Meta-Analysis of the Relation between European and American Smokeless Tobacco and Oral Cancer. *BMC Public Health*. In press.

16. Lissowska J, Pilarska A, Pilarski P. Smoking, Alcohol, Diet, Dentition and Sexual Practices in the Epidemiology of Oral Cancer in Poland. *European Journal of Cancer Prevention*. 2003;12(1):25-33.

17. Levi F, Pasche C, La Vecchia C, Lucchini F, Franceschi S, Monnier P. Food Groups and Risk of Oral and Pharyngeal Cancer. *Int. J. Cancer*. 1998;77:705–709.

18. Vitolins M, Tooze J, Golden S, el al. Older Adults in the Rural South Are Not Meeting Healthful Eating Guidelines. *Journal of the American Dietetic Association.* 2007;107(2):265-272.

19. Bailey R, Ledikwe J, Smiciklas-Wright H, Mitchell D, Jensen G. Persistent Oral Health Problems Associated with Comorbidity and Impaired Diet Quality in Older Adults. *Journal of the American Dietetic Association.* 2004;104(4):1273-1276.

20. D'Souza G, Kreimer A, Viscidi R, et al. Case–Control Study of Human Papillomavirus and Oropharyngeal Cancer. *N Engl J Med*. 2007;356:1944-56.

21. Ryerson A, Peters E, Coughlin S, et al. Burden of Potentially Human Papillomavirus Associated Cancers of the Oropharynx and Oral Cavity in the US, 1998-2003. *CANCER Supplement.* 2008;113(10):2901-2909.

22. Hopenhayn C, King J, Christian A, Huang B, Christian W. Variability of Cervical Cancer Rates Across 5 Appalachian States, 1998-2003. *CANCER Supplement.* 2008;113(10):2974-2980.

23. Vargas C, Dye B, Hayes K. Oral health status of rural adults in the United States. *J Am Dent Assoc.* 2002;133:1672-1681.

24. Conway D, Petticrew M, Marlborough H, Berthiller J, Hashibe M, Macpherson L. Socioeconomic Inequalities and Oral Cancer Risk: A Systematic Review and Meta-Analysis of Case-Control Studies. *Int. J. Cancer*. 2008;122:2811–2819.

25. Polednak A. Geographic Pattern of Cancers Related to Tobacco and Alcohol in Connecticut: Implications for Cancer Control. *Cancer Detection and Prevention.* 2004;28:302–308.

26. USDA Economic Research Service. Rural Income, Poverty, and Welfare: Rural Poverty. Available at: http://www.ers.usda.gov/Briefing/incomepovertywelfare/ RuralPoverty/. Accessed August 16, 2008.

27. USDA Economic Research Service. Rural Income, Poverty, and Welfare: Rural Income. Available at: http://www.ers.usda.gov/Briefing/incomepovertywelfare/ RuralIncome/. Accessed August 16, 2008.

28. Suba Z. Gender-Related Hormonal Risk Factors for Oral Cancer. *Pathology Oncology Research.* 2007;13(3):195–202.

29. Gallus S, Bosetti C, Franceschi S, Levi F, Simonato L, Negri E, La Vecchia C. Oesophageal Cancer in Women: Tobacco, Alcohol, Nutritional and Hormonal Factors. *British Journal of Cancer.* 2001;85(3):341–345.

30. Garavello1 W, Foschi R, Talamini R. Family History and the Risk of Oral and Pharyngeal Cancer. *Int. J. Cancer*. 2008;122:1827–1831.

31. National Cancer Institute. Rural-Urban Continuum Codes. Available at: http://seer.cancer.gov/seerstat/variables/countyattribs/Rural.Urban.Continuum.Codes.1974.1983.1993.2003.xls. Accessed July 19, 2008.

Appendix A: Human Subjects Approval

The University of Kansas Medical Center

Human Research Protection Program

July 18, 2008

Project Title:	Comparing Oral Pharangeal Cancer Rates in Rural and Urban Populations
Sponsor:	None
Protocol Number:	N/A
Primary Investigator:	Catherine Womack
Status:	Not Human Subject Research

Dear Investigator:

Thank you for your submission. This is to certify that your research proposal has been reviewed by the KUMC Human Subjects Committee (HSC). The HSC has determined that your proposal is deemed to not involve human subjects and, as such, the HSC has no jurisdiction over your proposal and you may proceed with your activities.

Please note that if you revise your activities to include human subjects directly, or by collecting information identifying human subjects, you should contact our office immediately. The HSC must determine whether or not the revisions impact the risks to human subjects, thus affecting the project's status as not involving human subjects.

If you have any questions regarding the human subject protection process, please do not hesitate to contact our office.

Very truly yours,

Daniel J. Voss, M.S., J.D.
IRB Administrator

Appendix B: Dataset Index

Poisson Regression Analysis

Variable	Type	Description
Count	Continuous	Number of Cases per population frequency
MSA	Categorical	Rural or Urban Residency (taken from Fips Codes)
Race	Categorical	White, Black, Hispanic, American Indian, Asian/Pacific Islander
Sex	Categorical	Male or Female
AgeGrp	Categorical	10-year age groupings (>35, 35-44, 45-54, 55-64, 65-74, 75+)
Registryid	Categorical	SEER Registry Identification Number
lnpop	Continuous	Natural Log of Population frequencies

Survival Analysis

Variable	Type	Description
Survival Time	Continuous	Survival time in months (0-60)
MSA	Categorical	Rural or Urban Residency (taken from Fips Codes)
Race	Categorical	White, Black, Hispanic, American Indian, Asian/Pacific Islander
Sex	Categorical	Male or Female
AgeGrp	Categorical	10-year age groupings (>35, 35-44, 45-54, 55-64, 65-74, 75+)
Registryid	Categorical	SEER Registry Identification Number
Censor	Categorical	Death during survival analysis period

Appendix C: SAS Output

Analysis Of Parameter Estimates

Parameter				DF	Estimate
Intercept				1	-11.9501
MSA	rural			1	0.0253
agegrp	35-44			1	2.3071
agegrp	45-54			1	3.5993
agegrp	55-64			1	4.2264
agegrp	65-74			1	4.4882
agegrp	75+			1	4.5322
Sex	female			1	-0.1158
registryid	1			1	0.1586
registryid	2			1	0.0112
registryid	20			1	0.1855
registryid	21			1	0.3301
registryid	22			1	0.0826
registryid	23			1	0.0992
registryid	25			1	0.1045
registryid	26			1	-0.1903
registryid	27			1	0.0192
registryid	29			1	0.2187
registryid	31			1	0.0089
registryid	35			1	0.0845
registryid	37			1	0.2517
registryid	41			1	0.1275
registryid	42			1	0.1308
registryid	43			1	0.2524
Race	2black			1	0.0205
Race	3hispanic			1	-0.6121
Race	4AmericanIndian			1	-1.6965
Race	5Asian/Islander			1	0.3511
Race	6Unknown			0	0.0000
agegrp*Sex*Race	35-44	female	2black	1	0.1807
agegrp*Sex*Race	35-44	female	3hispanic	1	0.1034
agegrp*Sex*Race	35-44	female	4AmericanIndian	1	-0.0131
agegrp*Sex*Race	35-44	female	5Asian/Islander	1	0.1829
agegrp*Sex*Race	35-44	female	6Unknown	0	0.0000
agegrp*Sex*Race	45-54	female	2black	1	0.1377
agegrp*Sex*Race	45-54	female	3hispanic	1	0.1747
agegrp*Sex*Race	45-54	female	4AmericanIndian	1	0.6676
agegrp*Sex*Race	45-54	female	5Asian/Islander	1	0.3967
agegrp*Sex*Race	45-54	female	6Unknown	0	0.0000
agegrp*Sex*Race	55-64	female	2black	1	-0.2888
agegrp*Sex*Race	55-64	female	3hispanic	1	-0.0770
agegrp*Sex*Race	55-64	female	4AmericanIndian	1	0.3736
agegrp*Sex*Race	55-64	female	5Asian/Islander	1	0.2086
agegrp*Sex*Race	55-64	female	6Unknown	0	0.0000
agegrp*Sex*Race	65-74	female	2black	1	-0.3635
agegrp*Sex*Race	65-74	female	3hispanic	1	-0.2084

Analysis Of Parameter Estimates

Parameter				Standard Error
Intercept				0.0456
MSA	rural			0.0413
agegrp	35-44			0.0507
agegrp	45-54			0.0458
agegrp	55-64			0.0454
agegrp	65-74			0.0457
agegrp	75+			0.0460
Sex	female			0.0537
registryid	1			0.0234
registryid	2			0.0248
registryid	20			0.0227
registryid	21			0.0383
registryid	22			0.0254
registryid	23			0.0338
registryid	25			0.0233
registryid	26			0.0357
registryid	27			0.0289
registryid	29			0.1958
registryid	31			0.0319
registryid	35			0.0202
registryid	37			0.0973
registryid	41			0.0167
registryid	42			0.0231
registryid	43			0.0225
Race	2black			0.0806
Race	3hispanic			0.0758
Race	4AmericanIndian			0.5796
Race	5Asian/Islander			0.0792
Race	6Unknown			0.0000
agegrp*Sex*Race	35-44	female	2black	0.1145
agegrp*Sex*Race	35-44	female	3hispanic	0.1255
agegrp*Sex*Race	35-44	female	4AmericanIndian	0.4275
agegrp*Sex*Race	35-44	female	5Asian/Islander	0.1183
agegrp*Sex*Race	35-44	female	6Unknown	0.0000
agegrp*Sex*Race	45-54	female	2black	0.0701
agegrp*Sex*Race	45-54	female	3hispanic	0.0899
agegrp*Sex*Race	45-54	female	4AmericanIndian	0.2806
agegrp*Sex*Race	45-54	female	5Asian/Islander	0.0853
agegrp*Sex*Race	45-54	female	6Unknown	0.0000
agegrp*Sex*Race	55-64	female	2black	0.0729
agegrp*Sex*Race	55-64	female	3hispanic	0.0888
agegrp*Sex*Race	55-64	female	4AmericanIndian	0.2694
agegrp*Sex*Race	55-64	female	5Asian/Islander	0.0885
agegrp*Sex*Race	55-64	female	6Unknown	0.0000
agegrp*Sex*Race	65-74	female	2black	0.0801
agegrp*Sex*Race	65-74	female	3hispanic	0.0926

Analysis Of Parameter Estimates

Parameter				Wald 95% Confidence Limits
Intercept				-12.0395
MSA	rural			-0.0557
agegrp	35-44			2.2077
agegrp	45-54			3.5095
agegrp	55-64			4.1374
agegrp	65-74			4.3985
agegrp	75+			4.4420
Sex	female			-0.2211
registryid	1			0.1127
registryid	2			-0.0373
registryid	20			0.1410
registryid	21			0.2551
registryid	22			0.0327
registryid	23			0.0330
registryid	25			0.0589
registryid	26			-0.2603
registryid	27			-0.0374
registryid	29			-0.1651
registryid	31			-0.0537
registryid	35			0.0450
registryid	37			0.0610
registryid	41			0.0948
registryid	42			0.0856
registryid	43			0.2084
Race	2black			-0.1375
Race	3hispanic			-0.7607
Race	4AmericanIndian			-2.8326
Race	5Asian/Islander			0.1959
Race	6Unknown			0.0000
agegrp*Sex*Race	35-44	female	2black	-0.0437
agegrp*Sex*Race	35-44	female	3hispanic	-0.1427
agegrp*Sex*Race	35-44	female	4AmericanIndian	-0.8510
agegrp*Sex*Race	35-44	female	5Asian/Islander	-0.0491
agegrp*Sex*Race	35-44	female	6Unknown	0.0000
agegrp*Sex*Race	45-54	female	2black	0.0003
agegrp*Sex*Race	45-54	female	3hispanic	-0.0014
agegrp*Sex*Race	45-54	female	4AmericanIndian	0.1176
agegrp*Sex*Race	45-54	female	5Asian/Islander	0.2295
agegrp*Sex*Race	45-54	female	6Unknown	0.0000
agegrp*Sex*Race	55-64	female	2black	-0.4315
agegrp*Sex*Race	55-64	female	3hispanic	-0.2511
agegrp*Sex*Race	55-64	female	4AmericanIndian	-0.1545
agegrp*Sex*Race	55-64	female	5Asian/Islander	0.0351
agegrp*Sex*Race	55-64	female	6Unknown	0.0000
agegrp*Sex*Race	65-74	female	2black	-0.5205
agegrp*Sex*Race	65-74	female	3hispanic	-0.3899

Analysis Of Parameter Estimates

Parameter				Wald 95% Confidence Limits
Intercept				-11.8607
MSA	rural			0.1063
agegrp	35-44			2.4065
agegrp	45-54			3.6891
agegrp	55-64			4.3153
agegrp	65-74			4.5778
agegrp	75+			4.6223
Sex	female			-0.0105
registryid	1			0.2046
registryid	2			0.0598
registryid	20			0.2300
registryid	21			0.4052
registryid	22			0.1325
registryid	23			0.1654
registryid	25			0.1501
registryid	26			-0.1203
registryid	27			0.0758
registryid	29			0.6025
registryid	31			0.0714
registryid	35			0.1240
registryid	37			0.4425
registryid	41			0.1601
registryid	42			0.1760
registryid	43			0.2965
Race	2black			0.1786
Race	3hispanic			-0.4634
Race	4AmericanIndian			-0.5604
Race	5Asian/Islander			0.5063
Race	6Unknown			0.0000
agegrp*Sex*Race	35-44	female	2black	0.4051
agegrp*Sex*Race	35-44	female	3hispanic	0.3495
agegrp*Sex*Race	35-44	female	4AmericanIndian	0.8249
agegrp*Sex*Race	35-44	female	5Asian/Islander	0.4148
agegrp*Sex*Race	35-44	female	6Unknown	0.0000
agegrp*Sex*Race	45-54	female	2black	0.2751
agegrp*Sex*Race	45-54	female	3hispanic	0.3508
agegrp*Sex*Race	45-54	female	4AmericanIndian	1.2175
agegrp*Sex*Race	45-54	female	5Asian/Islander	0.5640
agegrp*Sex*Race	45-54	female	6Unknown	0.0000
agegrp*Sex*Race	55-64	female	2black	-0.1460
agegrp*Sex*Race	55-64	female	3hispanic	0.0971
agegrp*Sex*Race	55-64	female	4AmericanIndian	0.9017
agegrp*Sex*Race	55-64	female	5Asian/Islander	0.3820
agegrp*Sex*Race	55-64	female	6Unknown	0.0000
agegrp*Sex*Race	65-74	female	2black	-0.2064
agegrp*Sex*Race	65-74	female	3hispanic	-0.0268

Analysis Of Parameter Estimates

Parameter				Chi-Square
Intercept				68651.6
MSA	rural			0.37
agegrp	35-44			2068.58
agegrp	45-54			6170.64
agegrp	55-64			8671.32
agegrp	65-74			9627.53
agegrp	75+			9707.34
Sex	female			4.64
registryid	1			45.82
registryid	2			0.21
registryid	20			66.78
registryid	21			74.29
registryid	22			10.54
registryid	23			8.64
registryid	25			20.20
registryid	26			28.37
registryid	27			0.44
registryid	29			1.25
registryid	31			0.08
registryid	35			17.58
registryid	37			6.69
registryid	41			58.58
registryid	42			32.15
registryid	43			126.17
Race	2black			0.06
Race	3hispanic			65.13
Race	4AmericanIndian			8.57
Race	5Asian/Islander			19.66
Race	6Unknown			.
agegrp*Sex*Race	35-44	female	2black	2.49
agegrp*Sex*Race	35-44	female	3hispanic	0.68
agegrp*Sex*Race	35-44	female	4AmericanIndian	0.00
agegrp*Sex*Race	35-44	female	5Asian/Islander	2.39
agegrp*Sex*Race	35-44	female	6Unknown	.
agegrp*Sex*Race	45-54	female	2black	3.86
agegrp*Sex*Race	45-54	female	3hispanic	3.78
agegrp*Sex*Race	45-54	female	4AmericanIndian	5.66
agegrp*Sex*Race	45-54	female	5Asian/Islander	21.62
agegrp*Sex*Race	45-54	female	6Unknown	.
agegrp*Sex*Race	55-64	female	2black	15.71
agegrp*Sex*Race	55-64	female	3hispanic	0.75
agegrp*Sex*Race	55-64	female	4AmericanIndian	1.92
agegrp*Sex*Race	55-64	female	5Asian/Islander	5.55
agegrp*Sex*Race	55-64	female	6Unknown	.
agegrp*Sex*Race	65-74	female	2black	20.57
agegrp*Sex*Race	65-74	female	3hispanic	5.06

Analysis Of Parameter Estimates

Parameter				Pr > ChiSq
Intercept				<.0001
MSA	rural			0.5404
agegrp	35-44			<.0001
agegrp	45-54			<.0001
agegrp	55-64			<.0001
agegrp	65-74			<.0001
agegrp	75+			<.0001
Sex	female			0.0312
registryid	1			<.0001
registryid	2			0.6502
registryid	20			<.0001
registryid	21			<.0001
registryid	22			0.0012
registryid	23			0.0033
registryid	25			<.0001
registryid	26			<.0001
registryid	27			0.5062
registryid	29			0.2641
registryid	31			0.7809
registryid	35			<.0001
registryid	37			0.0097
registryid	41			<.0001
registryid	42			<.0001
registryid	43			<.0001
Race	2black			0.7992
Race	3hispanic			<.0001
Race	4AmericanIndian			0.0034
Race	5Asian/Islander			<.0001
Race	6Unknown			.
agegrp*Sex*Race	35-44	female	2black	0.1145
agegrp*Sex*Race	35-44	female	3hispanic	0.4102
agegrp*Sex*Race	35-44	female	4AmericanIndian	0.9756
agegrp*Sex*Race	35-44	female	5Asian/Islander	0.1223
agegrp*Sex*Race	35-44	female	6Unknown	.
agegrp*Sex*Race	45-54	female	2black	0.0496
agegrp*Sex*Race	45-54	female	3hispanic	0.0519
agegrp*Sex*Race	45-54	female	4AmericanIndian	0.0174
agegrp*Sex*Race	45-54	female	5Asian/Islander	<.0001
agegrp*Sex*Race	45-54	female	6Unknown	.
agegrp*Sex*Race	55-64	female	2black	<.0001
agegrp*Sex*Race	55-64	female	3hispanic	0.3860
agegrp*Sex*Race	55-64	female	4AmericanIndian	0.1656
agegrp*Sex*Race	55-64	female	5Asian/Islander	0.0184
agegrp*Sex*Race	55-64	female	6Unknown	.
agegrp*Sex*Race	65-74	female	2black	<.0001
agegrp*Sex*Race	65-74	female	3hispanic	0.0245

Analysis Of Parameter Estimates

Parameter				DF	Estimate
agegrp*Sex*Race	65-74	female	4AmericanIndian	1	-0.1378
agegrp*Sex*Race	65-74	female	5Asian/Islander	1	0.1781
agegrp*Sex*Race	65-74	female	6Unknown	0	0.0000
agegrp*Sex*Race	75+	female	2black	1	-0.1238
agegrp*Sex*Race	75+	female	3hispanic	1	-0.0279
agegrp*Sex*Race	75+	female	4AmericanIndian	1	0.3135
agegrp*Sex*Race	75+	female	5Asian/Islander	1	0.2882
agegrp*Sex*Race	75+	female	6Unknown	0	0.0000
agegrp*Sex	35-44	female		1	-0.7344
agegrp*Sex	45-54	female		1	-1.0954
agegrp*Sex	55-64	female		1	-1.0265
agegrp*Sex	65-74	female		1	-0.7916
agegrp*Sex	75+	female		1	-0.6209
agegrp*Race	35-44	2black		1	-0.0541
agegrp*Race	35-44	3hispanic		1	-0.1395
agegrp*Race	35-44	4AmericanIndian		1	1.4620
agegrp*Race	35-44	5Asian/Islander		1	-0.3515
agegrp*Race	35-44	6Unknown		0	0.0000
agegrp*Race	45-54	2black		1	0.1119
agegrp*Race	45-54	3hispanic		1	-0.1595
agegrp*Race	45-54	4AmericanIndian		1	0.9602
agegrp*Race	45-54	5Asian/Islander		1	-0.7782
agegrp*Race	45-54	6Unknown		0	0.0000
agegrp*Race	55-64	2black		1	0.1490
agegrp*Race	55-64	3hispanic		1	-0.0348
agegrp*Race	55-64	4AmericanIndian		1	1.0835
agegrp*Race	55-64	5Asian/Islander		1	-0.9474
agegrp*Race	55-64	6Unknown		0	0.0000
agegrp*Race	65-74	2black		1	0.0296
agegrp*Race	65-74	3hispanic		1	0.0497
agegrp*Race	65-74	4AmericanIndian		1	1.1375
agegrp*Race	65-74	5Asian/Islander		1	-1.0018
agegrp*Race	65-74	6Unknown		0	0.0000
agegrp*Race	75+	2black		1	-0.2765
agegrp*Race	75+	3hispanic		1	0.2137
agegrp*Race	75+	4AmericanIndian		1	0.3727
agegrp*Race	75+	5Asian/Islander		1	-0.9454
agegrp*Race	75+	6Unknown		0	0.0000
Scale				0	1.0000

Analysis Of Parameter Estimates

Parameter				Standard Error
agegrp*Sex*Race	65-74	female	4AmericanIndian	0.3441
agegrp*Sex*Race	65-74	female	5Asian/Islander	0.0912
agegrp*Sex*Race	65-74	female	6Unknown	0.0000
agegrp*Sex*Race	75+	female	2black	0.0900
agegrp*Sex*Race	75+	female	3hispanic	0.0899
agegrp*Sex*Race	75+	female	4AmericanIndian	0.5005
agegrp*Sex*Race	75+	female	5Asian/Islander	0.0911
agegrp*Sex*Race	75+	female	6Unknown	0.0000
agegrp*Sex	35-44	female		0.0715
agegrp*Sex	45-54	female		0.0608
agegrp*Sex	55-64	female		0.0590
agegrp*Sex	65-74	female		0.0587
agegrp*Sex	75+	female		0.0579
agegrp*Race	35-44	2black		0.1053
agegrp*Race	35-44	3hispanic		0.1024
agegrp*Race	35-44	4AmericanIndian		0.6252
agegrp*Race	35-44	5Asian/Islander		0.1053
agegrp*Race	35-44	6Unknown		0.0000
agegrp*Race	45-54	2black		0.0885
agegrp*Race	45-54	3hispanic		0.0885
agegrp*Race	45-54	4AmericanIndian		0.6043
agegrp*Race	45-54	5Asian/Islander		0.0926
agegrp*Race	45-54	6Unknown		0.0000
agegrp*Race	55-64	2black		0.0877
agegrp*Race	55-64	3hispanic		0.0876
agegrp*Race	55-64	4AmericanIndian		0.5994
agegrp*Race	55-64	5Asian/Islander		0.0928
agegrp*Race	55-64	6Unknown		0.0000
agegrp*Race	65-74	2black		0.0910
agegrp*Race	65-74	3hispanic		0.0908
agegrp*Race	65-74	4AmericanIndian		0.6077
agegrp*Race	65-74	5Asian/Islander		0.0969
agegrp*Race	65-74	6Unknown		0.0000
agegrp*Race	75+	2black		0.1003
agegrp*Race	75+	3hispanic		0.0955
agegrp*Race	75+	4AmericanIndian		0.6781
agegrp*Race	75+	5Asian/Islander		0.1009
agegrp*Race	75+	6Unknown		0.0000
Scale				0.0000

Analysis Of Parameter Estimates

Parameter				Wald 95% Confidence Limits
agegrp*Sex*Race	65-74	female	4AmericanIndian	-0.8121
agegrp*Sex*Race	65-74	female	5Asian/Islander	-0.0005
agegrp*Sex*Race	65-74	female	6Unknown	0.0000
agegrp*Sex*Race	75+	female	2black	-0.3002
agegrp*Sex*Race	75+	female	3hispanic	-0.2042
agegrp*Sex*Race	75+	female	4AmericanIndian	-0.6674
agegrp*Sex*Race	75+	female	5Asian/Islander	0.1095
agegrp*Sex*Race	75+	female	6Unknown	0.0000
agegrp*Sex	35-44	female		-0.8746
agegrp*Sex	45-54	female		-1.2146
agegrp*Sex	55-64	female		-1.1421
agegrp*Sex	65-74	female		-0.9068
agegrp*Sex	75+	female		-0.7343
agegrp*Race	35-44	2black		-0.2604
agegrp*Race	35-44	3hispanic		-0.3402
agegrp*Race	35-44	4AmericanIndian		0.2367
agegrp*Race	35-44	5Asian/Islander		-0.5579
agegrp*Race	35-44	6Unknown		0.0000
agegrp*Race	45-54	2black		-0.0616
agegrp*Race	45-54	3hispanic		-0.3330
agegrp*Race	45-54	4AmericanIndian		-0.2241
agegrp*Race	45-54	5Asian/Islander		-0.9597
agegrp*Race	45-54	6Unknown		0.0000
agegrp*Race	55-64	2black		-0.0230
agegrp*Race	55-64	3hispanic		-0.2065
agegrp*Race	55-64	4AmericanIndian		-0.0912
agegrp*Race	55-64	5Asian/Islander		-1.1294
agegrp*Race	55-64	6Unknown		0.0000
agegrp*Race	65-74	2black		-0.1488
agegrp*Race	65-74	3hispanic		-0.1283
agegrp*Race	65-74	4AmericanIndian		-0.0536
agegrp*Race	65-74	5Asian/Islander		-1.1918
agegrp*Race	65-74	6Unknown		0.0000
agegrp*Race	75+	2black		-0.4731
agegrp*Race	75+	3hispanic		0.0265
agegrp*Race	75+	4AmericanIndian		-0.9564
agegrp*Race	75+	5Asian/Islander		-1.1433
agegrp*Race	75+	6Unknown		0.0000
Scale				1.0000

Analysis Of Parameter Estimates

Parameter				Wald 95% Confidence Limits
agegrp*Sex*Race	65-74	female	4AmericanIndian	0.5366
agegrp*Sex*Race	65-74	female	5Asian/Islander	0.3568
agegrp*Sex*Race	65-74	female	6Unknown	0.0000
agegrp*Sex*Race	75+	female	2black	0.0526
agegrp*Sex*Race	75+	female	3hispanic	0.1483
agegrp*Sex*Race	75+	female	4AmericanIndian	1.2944
agegrp*Sex*Race	75+	female	5Asian/Islander	0.4668
agegrp*Sex*Race	75+	female	6Unknown	0.0000
agegrp*Sex	35-44	female		-0.5942
agegrp*Sex	45-54	female		-0.9762
agegrp*Sex	55-64	female		-0.9110
agegrp*Sex	65-74	female		-0.6765
agegrp*Sex	75+	female		-0.5075
agegrp*Race	35-44	2black		0.1522
agegrp*Race	35-44	3hispanic		0.0612
agegrp*Race	35-44	4AmericanIndian		2.6874
agegrp*Race	35-44	5Asian/Islander		-0.1451
agegrp*Race	35-44	6Unknown		0.0000
agegrp*Race	45-54	2black		0.2853
agegrp*Race	45-54	3hispanic		0.0139
agegrp*Race	45-54	4AmericanIndian		2.1446
agegrp*Race	45-54	5Asian/Islander		-0.5967
agegrp*Race	45-54	6Unknown		0.0000
agegrp*Race	55-64	2black		0.3209
agegrp*Race	55-64	3hispanic		0.1370
agegrp*Race	55-64	4AmericanIndian		2.2583
agegrp*Race	55-64	5Asian/Islander		-0.7654
agegrp*Race	55-64	6Unknown		0.0000
agegrp*Race	65-74	2black		0.2080
agegrp*Race	65-74	3hispanic		0.2276
agegrp*Race	65-74	4AmericanIndian		2.3286
agegrp*Race	65-74	5Asian/Islander		-0.8118
agegrp*Race	65-74	6Unknown		0.0000
agegrp*Race	75+	2black		-0.0800
agegrp*Race	75+	3hispanic		0.4008
agegrp*Race	75+	4AmericanIndian		1.7019
agegrp*Race	75+	5Asian/Islander		-0.7476
agegrp*Race	75+	6Unknown		0.0000
Scale				1.0000

Analysis Of Parameter Estimates

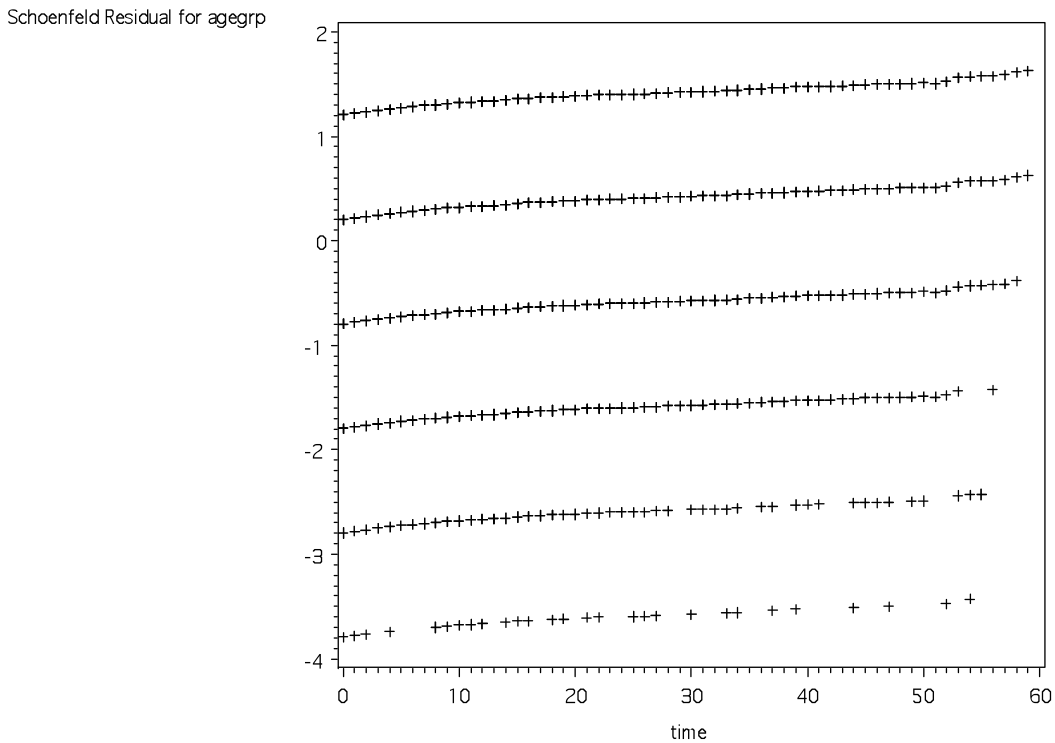
Parameter				Chi-Square
agegrp*Sex*Race	65-74	female	4AmericanIndian	0.16
agegrp*Sex*Race	65-74	female	5Asian/Islander	3.82
agegrp*Sex*Race	65-74	female	6Unknown	.
agegrp*Sex*Race	75+	female	2black	1.89
agegrp*Sex*Race	75+	female	3hispanic	0.10
agegrp*Sex*Race	75+	female	4AmericanIndian	0.39
agegrp*Sex*Race	75+	female	5Asian/Islander	10.00
agegrp*Sex*Race	75+	female	6Unknown	.
agegrp*Sex	35-44	female		105.42
agegrp*Sex	45-54	female		324.45
agegrp*Sex	55-64	female		302.99
agegrp*Sex	65-74	female		181.67
agegrp*Sex	75+	female		115.18
agegrp*Race	35-44	2black		0.26
agegrp*Race	35-44	3hispanic		1.86
agegrp*Race	35-44	4AmericanIndian		5.47
agegrp*Race	35-44	5Asian/Islander		11.14
agegrp*Race	35-44	6Unknown		.
agegrp*Race	45-54	2black		1.60
agegrp*Race	45-54	3hispanic		3.25
agegrp*Race	45-54	4AmericanIndian		2.53
agegrp*Race	45-54	5Asian/Islander		70.63
agegrp*Race	45-54	6Unknown		.
agegrp*Race	55-64	2black		2.88
agegrp*Race	55-64	3hispanic		0.16
agegrp*Race	55-64	4AmericanIndian		3.27
agegrp*Race	55-64	5Asian/Islander		104.14
agegrp*Race	55-64	6Unknown		.
agegrp*Race	65-74	2black		0.11
agegrp*Race	65-74	3hispanic		0.30
agegrp*Race	65-74	4AmericanIndian		3.50
agegrp*Race	65-74	5Asian/Islander		106.79
agegrp*Race	65-74	6Unknown		.
agegrp*Race	75+	2black		7.60
agegrp*Race	75+	3hispanic		5.01
agegrp*Race	75+	4AmericanIndian		0.30
agegrp*Race	75+	5Asian/Islander		87.75
agegrp*Race	75+	6Unknown		.
Scale				

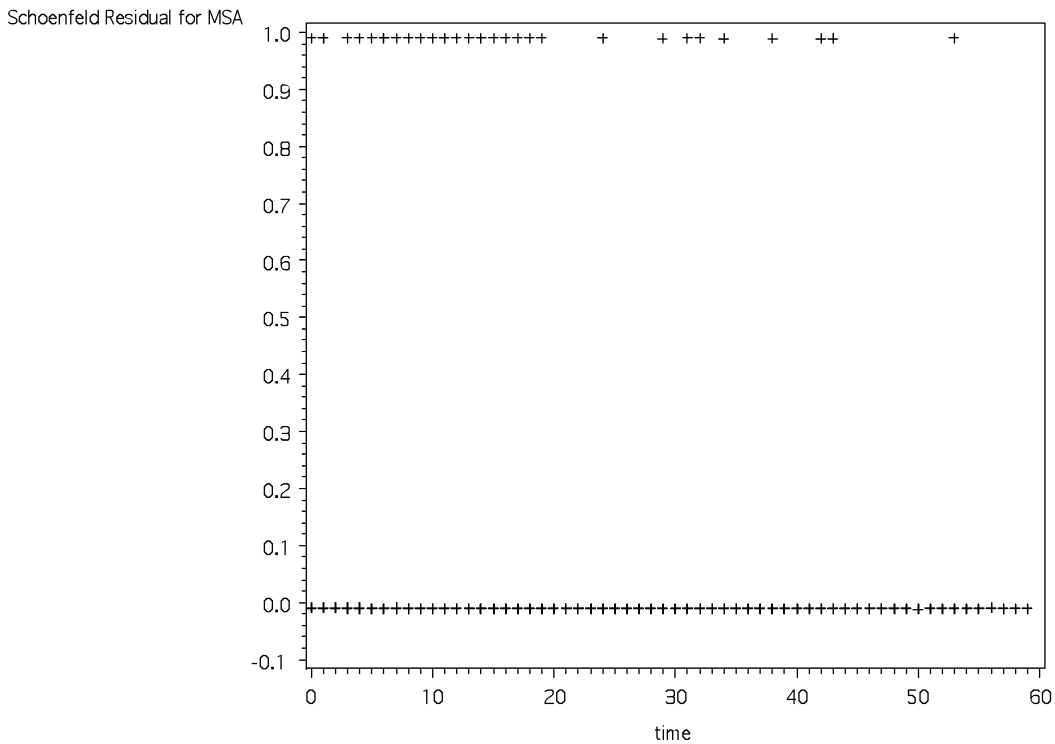
Analysis Of Parameter Estimates

Parameter				Pr > ChiSq
agegrp*Sex*Race	65-74	female	4AmericanIndian	0.6888
agegrp*Sex*Race	65-74	female	5Asian/Islander	0.0507
agegrp*Sex*Race	65-74	female	6Unknown	.
agegrp*Sex*Race	75+	female	2black	0.1691
agegrp*Sex*Race	75+	female	3hispanic	0.7561
agegrp*Sex*Race	75+	female	4AmericanIndian	0.5311
agegrp*Sex*Race	75+	female	5Asian/Islander	0.0016
agegrp*Sex*Race	75+	female	6Unknown	.
agegrp*Sex	35-44	female		<.0001
agegrp*Sex	45-54	female		<.0001
agegrp*Sex	55-64	female		<.0001
agegrp*Sex	65-74	female		<.0001
agegrp*Sex	75+	female		<.0001
agegrp*Race	35-44	2black		0.6071
agegrp*Race	35-44	3hispanic		0.1730
agegrp*Race	35-44	4AmericanIndian		0.0194
agegrp*Race	35-44	5Asian/Islander		0.0008
agegrp*Race	35-44	6Unknown		.
agegrp*Race	45-54	2black		0.2063
agegrp*Race	45-54	3hispanic		0.0714
agegrp*Race	45-54	4AmericanIndian		0.1120
agegrp*Race	45-54	5Asian/Islander		<.0001
agegrp*Race	45-54	6Unknown		.
agegrp*Race	55-64	2black		0.0896
agegrp*Race	55-64	3hispanic		0.6916
agegrp*Race	55-64	4AmericanIndian		0.0706
agegrp*Race	55-64	5Asian/Islander		<.0001
agegrp*Race	55-64	6Unknown		.
agegrp*Race	65-74	2black		0.7452
agegrp*Race	65-74	3hispanic		0.5845
agegrp*Race	65-74	4AmericanIndian		0.0612
agegrp*Race	65-74	5Asian/Islander		<.0001
agegrp*Race	65-74	6Unknown		.
agegrp*Race	75+	2black		0.0058
agegrp*Race	75+	3hispanic		0.0252
agegrp*Race	75+	4AmericanIndian		0.5826
agegrp*Race	75+	5Asian/Islander		<.0001
agegrp*Race	75+	6Unknown		.
Scale				

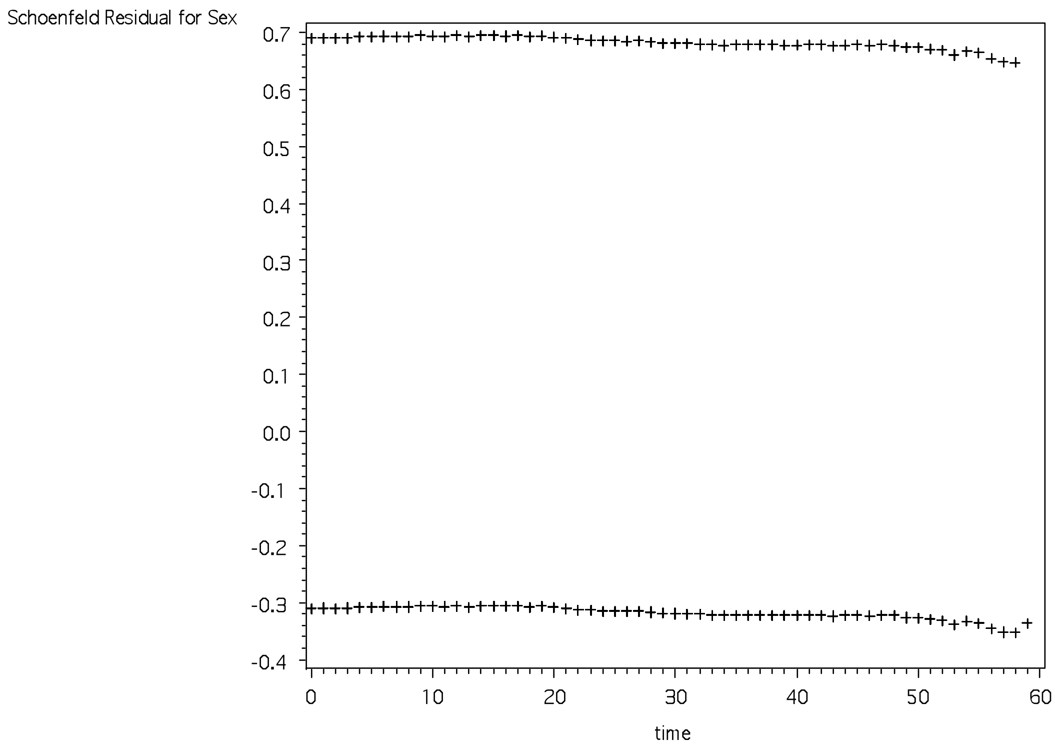
NOTE: The scale parameter was held fixed.

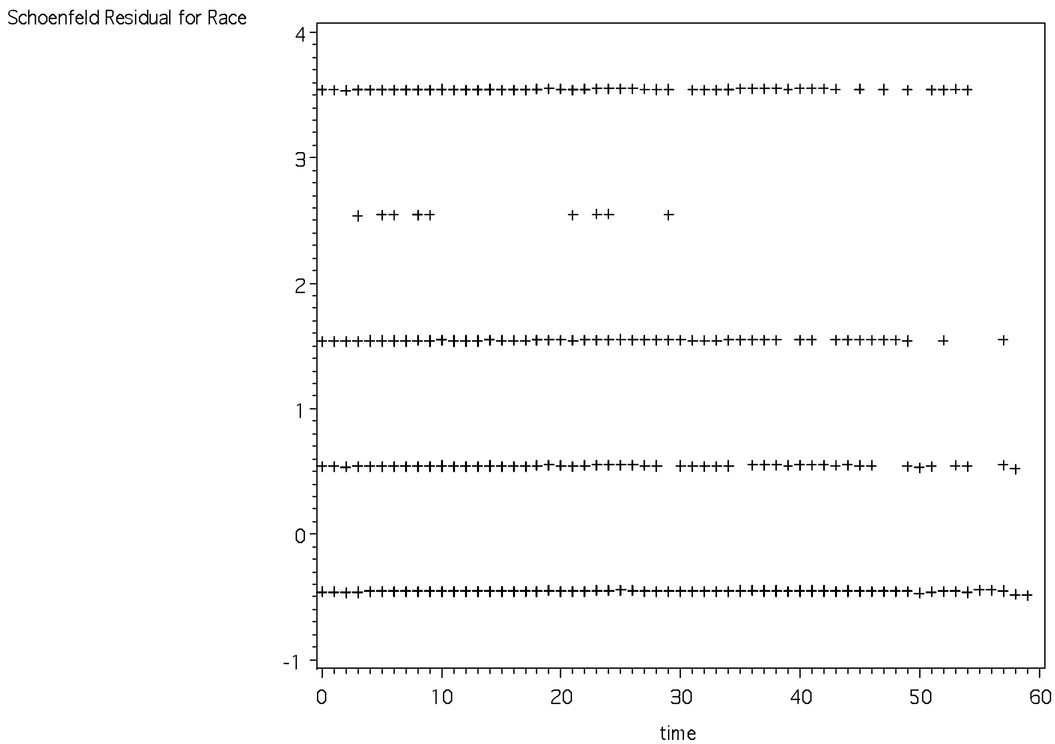
Appendix D. Schoenfeld Residual Plots





Schoenfeld Residual for MSA
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
-0.1
0
10
20
30
40
50
60
time


Schoenfeld Residual for Sex
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
-0.1
-0.2
-0.3
-0.4
0
10
20
30
40
50
60
time


Schoenfeld Residual for Race
4
3
2
1
0
-1
0
10
20
30
40
50
60
time