TRENDS IN CANCER AND DEVELOPMENT OF AIDS IN HIV-INFECTED INDIVIDUALS BEFORE AND AFTER IMPLEMENTATION OF HIGHLY ACTIVE ANTIRETROVIRAL THERAPY

By ROSE MARIE MARTIN

A dissertation submitted to the

School of Public Health

University of Medicine and Dentistry of New Jersey

and the

Graduate School – New Brunswick

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Written under the direction of

Kitaw Demissie, M.D., Ph.D.

And approved by

_____

_____

_____

_____

_____

Piscataway/New Brunswick, New Jersey
May 2010

ABSTRACT OF THE DISSERTATION

Trends in Cancer and Development of Aids in HIV-Infected Individuals Before and After Implementation of Highly Active Antiretroviral Therapy

By ROSE MARIE MARTIN

Dissertation Director:

Kitaw Demissie, M.D., Ph.D.

Research Objectives

To analyze the effect of highly active anti-retroviral therapy on cancer incidence and mean age at cancer diagnosis in HIV-infected persons; on progression to AIDS related to cancer; and risks of cancer, pre- and post-HAART.

Methods

Records on New Jersey HIV-infected residents were used for a retrospective cohort study of HIV and cancer. Lung, colorectal, prostate, Hodgkin lymphoma, and anal cancer were studied. Incidence density pre- and post-HAART was tested for change in Standardized Incidence Ratio (SIRs). Relative risk was approximated by the ratio of the SIRs from the two periods. Difference in mean age at diagnosis pre- and post-HAART for each type was tested by a Student's t. Risks of developing AIDS, pre- and post-HAART, in HIV-infected persons with and without a cancer diagnosis were evaluated through proportional hazards

regression. Cox Proportional Regression was used to assess the hazard ratios associated with cancer after HIV seroconversion in the two periods.

Results

Cumulative incidence rates increased substantially for prostate cancer, and slightly for anal cancer and Hodgkin lymphoma pre- to post-HAART, but not for lung and colorectal cancer. Cumulative incidence rates for HIV-infected persons observed only pre- HAART were lower when compared to those HIV-positive only post-HAART for all five studied cancer types. SIRs were significant for all but pre-HAART colorectal cancer. Relative risks post-HAART compared to pre-HAART for each of the cancers, other than Hodgkin lymphoma, were significantly low. Age-specific incidence rates post-HAART were higher in the oldest age groups for colorectal and prostate cancer in addition to Hodgkin disease, but decreased in every age group for lung cancer. Mean age at cancer diagnosis was significantly higher in the post-HAART period for lung and colorectal cancer and Hodgkin lymphoma, but unchanged for prostate cancer. The regression analysis indicated an increased risk of AIDS development if a cancer was diagnosed after HIV seroconversion, but before AIDS, although the risk decreased post-HAART. The hazard ratio for cancer development was not significantly different pre- and post-HAART.

Conclusions

HAART availability and an increase in survival have significantly changed trends in HIV infection and associated morbidity, particularly for cancer in HIV-infected persons.

ACKNOWLEDGEMENTS

Trends in Cancer and Development of AIDS

In HIV-Infected Individuals

Before And After Implementation of

Highly Active Antiretroviral Therapy

By Rose Marie Martin

Sincere thanks go to New Jersey Department of Health and Senior Service staff who expressed the necessary support for dissertation goals and provided the data files for conducting the analyses. These staff include Betsy Kohler, the Director of the Office of Cancer Epidemiology, Karen Pawlish, Research Scientist with the Cancer Registry Program, Dr. Helene Cross, Director of Epidemiological Services for the Division of HIV/AIDS Services, and Dr. Abdel Ibrahim, epidemiologist with the HIV/AIDS Reporting System.

Special thanks are also due to the chair and members of my Dissertation Committee, for the many helpful comments and suggestions which made this final document possible. The Dissertation Committee was comprised of: Kitaw Demissie, M.D., Ph.D. (Chair); Shou-En Lu, Ph.D.; David August, M.D.; Sindy Paul, M.D.; and Mark Fulcomer, Ph.D.

TABLE OF CONTENTS

Trends in Cancer and Development of AIDS in HIV-Infected Individuals Before and After Implementation of Highly Active Antiretroviral Therapy

by Rose Marie Martin

TABLES AND FIGURES

Trends in Cancer and Development of AIDS In HIV-Infected Individuals Before and After Implementation of Highly Active Antiretroviral Therapy

By Rose Marie Martin

Chapter I. Background

Since 1993, Kaposi sarcoma, non-Hodgkin lymphoma, and invasive cervical cancer have been the only malignancies classified as defining AIDS by the CDC (1992). However, relationships between a number of other malignancies and human immunodeficiency virus (HIV) infection are poorly understood. Questions remain about the susceptibility of HIV-infected persons to various cancers and whether these may be a result of either extended survival or the effect of highly active antiretroviral therapy (HAART), including the effects of age at diagnosis of these other cancers. Bower, Palmieri, and Dhillon (2006) suggested that the good news from the administration of HAART on the AIDS-defining cancers may be balanced by the increasing numbers of non-AIDS defining cancers. These authors noted that there are both toxicity and pharmacokinetic drawbacks to the concomitant administration of chemotherapy and HAART.

The research questions being investigated in this dissertation encompass selected aspects of the role of HIV-induced immunosuppression related to availability of HAART in susceptibility to cancer. The investigation extends from all cancer cases in HIV-infected persons and, in particular, to five non-AIDS defining cancers either known to be associated with HIV infection or on the basis of biologic possibility are likely to be associated with HIV infection. These five cancers are Hodgkin lymphoma, lung, colorectal, prostate, and anal cancer.

There have been a number of studies aimed at providing insight into the effects of HAART therapy for persons with HIV infection in changes in life expectancy, susceptibility to other non-AIDS defining illnesses, cancer incidence rates and median age at diagnosis of cancers, among other relevant issues. Most of these studies suffered from a lack of complete information on person-years of observation with HIV infection, as the registries in most states and municipalities developed to gather data on this condition only registered people once the diagnosis of AIDS had been made. The State of New Jersey not only has a cancer registry that has captured all reportable cases of cancer diagnosed in New Jersey residents since October 1, 1978, but has also required reporting of diagnosed cases of AIDS to a central registry since 1985. The AIDS registry was expanded in late 1991 to encompass reporting of all HIV cases in New Jersey residents. This expansion allows calculation of person-years of experience from the beginning of the infection to cancer diagnosis, to death, or survival to a selected date. The unlinked, de-identified file of resident New Jerseyans who had HIV/AIDS, but not cancer, as well as the matched file of HIV/AIDS and cancer cases, was made available to this research so that observation time from all individuals positive for HIV/AIDS could be included in the analysis. This permits the more accurate estimation of person-years of observation and thus the expected incidence of cancers in the HIV-infected population.

HIV/AIDS and Cancer

The American Cancer Society (ACS) (2007a) estimated that 7.6 million people worldwide died from cancer and more than 12 million people received cancer diagnoses in 2007. The ACS (2007b) also estimated that 559,650 of the deaths and 1,444,920 of the new cancer cases in 2007 occurred in the U.S. These figures exclude basal and squamous cell skin and *in situ* carcinomas except urinary bladder. Cancer is currently the second leading cause of death nationwide. Experts at the Division of Cancer Prevention and Control (DCPC) (2004) predict that cancer will become the leading cause of death in this country within the decade. Although heart disease is currently the leading cause of death, the gap between age-adjusted heart disease and neoplasm death rates has narrowed over the past three decades, due primarily to the decline in the heart disease death rate (Jemal, Ward, Hao, & Thun, 2005). There has also been a slight decline in the age-adjusted cancer death rates which has been attributed to progress in cancer prevention, early detection, and treatment (DCPC, 2004).

Hessol et al. (2007) stated that people with HIV infection are at increased risk of cancer and projected that an estimated 30-40 percent will develop a malignancy at some point during their infection. The causes and associations of cancer development in HIV-infected persons are still unclear, but have been the subject of investigation by experts. New Jersey had just under 35,000 persons living with HIV/AIDS in June, 2009 (New Jersey Division of HIV/AIDS, 2009) and a reported 1,700 – 1,900 newly identified positive cases each year since 1996.

Patel et al. found (2008) that, by 2003, anal cancer was 59 times more common among HIV-infected individuals than in the general population. These authors reported that Hodgkin disease was 18 times more

common, liver cancer seven times more common, lung cancer 3.6 times more common, melanoma and throat cancer each three times more common, and colorectal cancer 2.4 times more common in HIV-positive persons. The study also found that HIV-infected individuals had a small reduction in risk for prostate cancer.

The burden of cancer, both incidence and mortality, differs by site, demographics, and geography. The ACS (2007c) reported that more people in the U.S. die from lung cancer than from any other type of cancer. Lung cancer is the leading cause of cancer death in both men and women and in the most recent data year, 2007, accounted for more deaths than colon cancer, breast cancer, and prostate cancer, combined. In that year, an estimated 89,510 men and 70,880 women died of lung cancer and an estimated total of 213,380 new cases were diagnosed.

The demographics of lung cancer incidence in HIV-infected persons differ from those of lung cancer in the general population. Engels (2001) found overall incidence was higher in persons with AIDS than in the general population matched by age, sex, race, and calendar year ($P<.0001$), with the difference particularly notable for persons aged 30 through 59. Other studies have noted the younger age of HIV-infected lung cancer cases compared to the general population of lung cancer cases. Cooley (2003) reported a number of studies in which the median age of lung cancer diagnosis in persons with HIV/AIDS was less than 40 years in the majority. Brock et al. (2006) found that HIV-infected individuals were younger with more advanced cancer and CD4 counts and HIV-1 RNA levels that indicated preserved immune function. Additionally, many of the studies of lung cancer in HIV-infected persons reported more adverse outcomes for these patients, with shortened survival (Tirelli et al., 2000; Brock et al., 2006; Spano et al., 2004; & Sridhar, Flores, Raub, & Saldana, 1992), although Hakimian, Fang, Thomas, & Edelman, (2007) found in their

retrospective series that patients with advanced-stage non-small cell lung cancer had a survival time that approached that of HIV-negative lung cancer patients (their results did not reach statistical significance). In the early years of the HIV epidemic, researchers did not detect an increased incidence of lung cancer among infected persons (Biggar, Burnett, Miki, & Nasca, 1989 & Chan, Arada, & Rom, 1993), although more recently, a number of researchers have found elevated lung cancer risks in HIV-infected populations. Frisch, Biggar, Engels, and Goedert (2001) surmised that the excess of lung cancer in persons with HIV/AIDS found in their study resulted from heavy smoking.

Cooley (2003) suggested that the association between lung cancer and HIV infection remains uncertain and available data do not link the incidence of lung cancer with the degree of immunosuppression. He also stated that other possible mechanisms, in addition to smoking, may be responsible for the increased incidence of lung cancer in HIV-positive people, but have not been clearly defined. The clinical course of lung cancer in HIV-infected persons differs from that of the general population. Engels et al. (2006a) echoed Cooley's results (2003) when they found that lung cancer incidence was unrelated to HIV-induced immunosuppression. Incidence remained high after adjustment for smoking, also suggesting the involvement of additional factors.

Colorectal cancer is the second leading cause of cancer-related deaths in the United States and is also one of the most commonly diagnosed cancers. In 2007, more than 153,760 United States residents were diagnosed with colorectal cancer and an estimated 52,180 people died of this cancer (ACS, 2007d). Colorectal cancer incidence and deaths have been declining in recent years. During the period 1998 through 2004, the nationwide incidence of colorectal cancer decreased by 2.3 percent per year in the total population while

deaths decreased by 4.7 percent per year over the period 2002 to 2004. Both of these decreases were statistically significant (U.S. Department of Health and Human Services ((USDHSS), 2007).

Colorectal cancer is a malignancy that has not received thorough study in HIV-infected populations due to low incidence in the age groups affected with HIV. Few large studies of the incidence and natural history of colon cancer among the HIV-infected population have been conducted to date (International Collaboration on HIV and Cancer, 2000). A recent study on neoplastic lesions in asymptomatic HIV-infected subjects (Bini, Green, & Poles, 2009) found a higher prevalence of colonic neoplasms in HIV-infected persons and an earlier age of development of adenocarcinomas. This study also found that in their study population, the neoplasms were more advanced in HIV-positive individuals than in uninfected subjects. The authors conclude that screening colonoscopy should be offered to persons positive for HIV infection, but thought that more study is needed of the age of initiation and the frequency of screening for this population.

Except for skin cancer, prostate cancer is the most frequently occurring malignancy among men in the U.S. and it is the second leading cause of death from cancer in men, after lung cancer (ACS, 2007e). Overall, prostate cancer is the sixth leading cause of death in males in this country. In 2007, an estimated 218,890 men nationwide were diagnosed with and 27,050 died from prostate cancer. More than 70 percent of all prostate cancers are diagnosed in men aged 65 and over and death rates are higher in African-American men than in any other racial/ethnic group.

Crum, Spencer, & Amling (2004) reported in their findings that prostate cancer is common in older men who are HIV-infected and that it is associated with African-American race and duration of HIV infection. There is disagreement on the factors associated with the development of prostate cancer. In a study of 44,788 pairs of twins in Sweden, Denmark, and Finland, Lichtenstein et al. (2000) concluded that 42 percent

of prostate cancer cases were due to inheritance. Ghadirian, Howe, Hislop & Maisonneuve (1997) reported that epidemiologic studies have consistently reported the existence of familial associations of prostate cancer, but the authors went on to state that common environmental or lifestyle exposures could play a role in familial patterns. Nelson, De Marzo, & Isaacs (2003) attributed the majority of prostate cancer to environmental factors, such as diet, lifestyle-related factors, and androgens, but they also reported that two inherited susceptibility genes may have roles in responses to infections, raising the possibility that infection or inflammation begins the carcinogenesis process in prostate cancer.

Manfredi et al. (2006) noted that an increased incidence of prostate cancer in HIV-infected men has not been established although it appeared to be more frequent at younger ages among the population with HIV infection. Biggar et al. (2004) reported a low overall risk for prostate cancer in their study of cancer risks in the elderly with HIV/AIDS, while Vianna, Lo, & Klein (2006) found that prostate-specific antigen (PSA) levels increased with age, but did not differ by HIV status. Amid calls for expanded PSA testing in the elderly (Manfredi et al., 2006 & di Gennaro, et al. 2005) and studies that recommended the use of American Cancer Society guidelines for men in general (Catalona, Antenor, & Roehl, 2002) or extension to age 40 for men at high risk (Levinson, Nagler, & Lowe, 2005), the need for clearer indications of the trends in age of diagnosis of prostate cancer in HIV-infected men is evident.

The ACS (2007f) defines Hodgkin disease as a type of cancer which starts in lymphatic tissue, encompassing the lymph nodes and other organs that comprise the body's immune and blood-forming system. Hodgkin disease is relatively rare in the general population, with only an estimated 8,190 new cases diagnosed in the country in 2007. The ACS also estimated that there were 1,070 deaths caused by Hodgkin disease in 2007.

Kinlen (2004) reported that an infective, mostly viral, basis has been found in an increasing number of different human cancers. Cameron and Hagensee (2007) say that "how HPV and HIV interact is still not known but is more likely to be linked to immune suppression rather than a direct interaction between viruses." Barbaro and Barbarini (2007) thought that while it is still "unclear whether HIV-1 acts directly as an oncogenic agent, it may contribute to the development of malignancies through several mechanisms (e.g., infection by oncogenic viruses, impaired immune surveillance, imbalance between cellular proliferation and differentiation)".

A number of authors have speculated about the role of EBV in the pathophysiology of HIV and Hodgkin disease (Hoffman et al., 2004). Goedert et al. (1998) suggested that, in addition to the possibility of behavioral risk factors, immunological failure to control herpes or other viral infections may contribute to certain malignancies including Hodgkin disease. Frisch, Biggar, Engels, and Goedert (2001) also reported that Hodgkin disease may be influenced by immunosuppression. More than a few studies have suggested an association between Hodgkin disease and HIV infection and questioned whether Hodgkin disease should be considered an AIDS-defining illness (Serraino et al., 1997; Grulich, Wan, Law, Coates, & Kaldor, 1999; & Clarke & Glaser, 2001).

Carcinoma of the anal canal is also relatively rare, responsible for only 1.5 percent of gastrointestinal tract cancer in the country as a whole (Ryan, Compton, & Mayer, 2000). Data published by the ACS (2007g) indicated that, in the U.S., an estimated 1,900 men and 2,750 women were diagnosed with anal cancer in 2007 and 690 residents died of the malignancy in that year. Incidence of anal cancer increased significantly over the period in which Surveillance Epidemiology and End Results (SEER) data are available, from 1973 to the present, in both men and women, with the increase being greater among men (Johnson, Madeleine,

Newcomer, Schwartz, & Daling, 2004). The findings indicated the highest rate of increase over the period in anal cancer incidence occurred among black men. Black women had the highest race-specific and gender-specific rates of anal cancer incidence in the first half of the time period studied, but had the lowest incidence rate in the most recent period.

In past decades, anal cancer was thought to be the result of chronic irritation from conditions such as hemorrhoids, fissures, and fistulae. There was also thought to be an association with inflammatory bowel disease (Ryan et al., 2000, Uronis & Bendell, 2007). Frisch, Olsen, Bautz, & Melbye (1994) concluded after a large-scale study of linked records in Denmark that in "…most cases in which presumably benign lesions (hemorrhoids, fistulas, fissures, or abscesses) precede anal cancer, the lesions probably represent the initial symptoms of an undetected anal cancer." Additionally, a case-control study by Holly et al. (1989) showed little, if any, risk of anal cancer associated with hemorrhoids, fissures, or fistulae.

A number of risk factors for anal cancer have been identified through epidemiologic studies (Ryan et al., 2000, Uronis & Bendell, 2007). There is strong evidence which identifies human papillomavirus infection (anogenital warts), history of receptive anal intercourse, a history of sexually transmitted disease, more than 10 sexual partners, history of cervical, vulvar, or vaginal cancer, and immunosuppression after solid-organ transplantation as risk factors for anal cancer. In addition, there is moderately strong evidence that human immunodeficiency virus (HIV) infection, long-term use of corticosteroids, and cigarette smoking are also anal cancer risk factors.

The association between anal cancer and HIV infection has been difficult to separate due to confounders, according to Uronis and Bendell (2007). The situation is complicated by the fact that HIV-positive patients

are more likely to be infected with HPV, often with more than one subtype. Ryan et al. (2000) reported that HIV-positive persons are two to six times as likely as HIV-negative persons to have anal human papillomavirus infection, regardless of sexual practices. This risk is inversely proportional to the CD4 lymphocyte count. Also, HIV-infected patients with low-grade anal intraepithelial neoplasia are twice as likely as HIV-negative persons to progress to high-grade anal intraepithelial neoplasia within two years, a risk that is again inversely proportional to the CD4 lymphocyte count (Palefsky et al., 1998).

Chiao and Krown (2003) reported studies which show that, compared with the same neoplastic processes in HIV-negative patients, some malignancies tend to be of higher grade and present with a more aggressive clinical course in HIV-positive patients. The role of chemotherapy, radiation therapy and surgery for treatment of non-AIDS defining malignancies remains unclear.

HIV and Cancer in the Pre-HAART and Post-HAART Eras

The advent of highly active antiretroviral therapy (HAART) is commonly assumed to have begun in 1996 and has led to an increase in years of survival for persons infected with HIV (Palella et al., 1998; Mocroft et al, 1998; Lloyd-Smith et al., 2006; Lai & Hardy, 2004; Hogg et al, 1997; & Hogg et al., 1998). Hogg and colleagues (1998) found that HIV-infected individuals who received initial therapy with regimens including stavudine or lamivudine had significantly lower mortality and longer AIDS-free survival than those who received initial therapy with regimens limited to zidovudine, didanosine, and zalcitabine. Palella, et al. (1998) concluded that the "recent declines in morbidity and mortality due to AIDS are attributable to the use of more intensive antiretroviral therapies. Combination antiretroviral therapy was associated with the most benefit: the inclusion of protease inhibitors in such regimens conferred additional benefits".

One of the recent studies of trends in the incidence of both AIDS-defining and non-AIDS-defining cancers in HIV-infected persons during pre- and post-HAART eras (Crum-Cianflone et al., 2009) concluded that HAART use is protective for AIDS-defining cancers but had no significant impact on non-AIDS-defining cancers. The Swiss Cohort Study (Clifford et al., 2005) investigated behavioral risks and HAART use in patients matched from HIV and cancer registries in Switzerland and found that HAART use could prevent excess risk of Kaposi's sarcoma and non-Hodgkin lymphoma, but not that from Hodgkin lymphoma and other non-AIDS-defining cancers in HIV-infected persons. A recent study of cancer risk in HIV-infected persons (Engels et al., 2008) also found a decline in Kaposi's sarcoma and non-Hodgkin lymphoma and an increase of some non-AIDS-defining cancers most likely due to the use of HAART. A meta-analysis of cancer incidence data from 23 prospective studies of HIV-positive individuals from North America, Europe, and Australia (International Collaboration of HIV and Cancer, 2000) also found substantial reductions in the incidence of Kaposi's sarcoma and non-Hodgkin lymphoma in persons infected with HIV, but no significant change in the incidence of other cancers.

The more substantial change in cancer incidence may occur as a result of increasing life expectancy of HAART users. As HAART prolongs the lives of HIV-infected persons it is expected that the cumulative incidence of these non-AIDS-defining cancers will rise. As HIV-infected persons age, there will be an increasing incidence of cancers of the colon, breast and prostate (Engels and Goedert, 2005). The etiology of these and other age-related cancers is complex and poorly understood, but these authors pointed out that they mostly arise through a series of genetic mutations acquired over a lifetime.

Clifford et al. (2005) reported that the standardized incidence ratios (SIRs) of certain non-AIDS-defining cancers may have increased in the post-HAART era (1996 and after) and surmised that the improvement in

life expectancy made possible by HAART along with an only partial reconstitution of immune status, may allow a larger number of cancers with long latent periods to manifest clinically.

Lim and Levine (2005) reported that the role of immunosuppression in the pathogenesis of non-AIDS-defining cancers is controversial. The relation between cancer risk and degree of immune deficiency seems to differ by cancer type (Clifford and Franceschi, 2007). Lederman and Valdez (2000) reported that although immune restoration due to suppression of HIV replication is a critical determinant of the trend in decreased mortality related to AIDS, the magnitude of immune restoration seen after treatment with HAART varies substantially among treated persons and is generally incomplete. Frisch, Biggar, and Goedert for the AIDS-Cancer Match Registry Study Group (2000a) concluded that HPV-associated malignancies occur at increased rates in persons with HIV/AIDS. Increasing relative risks for *in situ* cancers to and beyond the time of AIDS onset may reflect the gradual loss of control over HPV-infected keratinocytes with advancing immunosuppression. However, the lack of a similar increase for invasive HPV-associated cancers suggests that late-stage cancer invasion is not greatly influenced by immune status.

Bonnet et al. (2004) found that "Malignant disease has been a major cause of death among HIV-infected patients in industrialized nations since the introduction of HAART. Whereas lethal hemopathies and Kaposi sarcoma are associated with advanced immunosuppression, lethal solid tumors can occur in patients with controlled HIV infection". In a summary of studies that attempted to evaluate the association of cancer with level of immunosuppression, Frisch and colleagues (2001) defined three criteria that suggested that non-AIDS-defining malignancies were associated with immuno-suppression: (1) elevated overall Relative Risk in the period from 60 months before to 27 months after AIDS; (2) elevated Relative Risk in the 4- to 27-month post-AIDS period; and (3) increasing trend in Relative Risk from before to after AIDS onset They

found the following individual cancers met all three criteria: Hodgkin disease, lung cancer, penile cancer, soft tissue malignancies, lip cancer, and testicular seminoma. Also, Gallagher, Wang, Schymura, Kahn and Fordyce (2001) showed that cancer of the rectum, rectosigmoid, and anus; trachea, bronchus, and lung; skin; and connective tissues among males were associated with increasing immunosuppression. Metabolic abnormalities associated with HAART (including redistribution of fat from peripheral and buttock sites to the breasts and abdomen, elevated waist-hip ratios and body mass indices and undesirable serum lipid levels) have also been associated with breast cancer risk and survival. Frisch et al., for the AIDS-Cancer Match Registry Study Group (2001) went on to suggest that "most non-AIDS-defining cancers do not appear to be influenced by the advancing immunosuppression associated with HIV disease progression." These researchers reported that "Some cancers that met our criteria for potential association with immunosuppression may have occurred in excess in persons with HIV/AIDS because of heavy smoking (lung cancer), frequent exposure to human papillomavirus (penile cancer), or inaccurately recorded cases of Kaposi sarcoma (soft tissue malignancies) in these persons." However, these authors found that Hodgkin disease, notably of the mixed cellularity and lymphocytic depletion subtypes, and possibly lip cancer and testicular seminoma may be genuinely influenced by immunosuppression. "The incidence of Kaposi's sarcoma decreases and regression can occur after immune reconstitution: in patients with HIV on HAART or in transplant recipients after removal of immunosuppressive therapy. By contrast, Hodgkin lymphoma occurs most often at only moderate levels of immunosuppression." Also, people who used HAART had lower risks of Kaposi sarcoma and non-Hodgkin lymphoma compared with those who did not use HAART, although even with HAART these tumors occurred 20 times more frequently than they do in the general population without HIV/AIDS. HAART use was not associated with lower risks of Hodgkin lymphoma or other cancers.

The incidence of anal carcinoma among people with HIV and men who have sex with men is substantially increased (Bower et al., 2004). There is apparently no correlation between the relative risk of developing invasive anal cancer and the CD4 cell count. Also, HAART "…does not seem to lead to resolution of anal intraepithelial neoplasia". Chiao, Krown, Stier, and Schrag (2005) reported that risk of persistent HPV infection and anal HSIL rates are significantly higher in HIV-infected men, despite HAART. Sobhani et al. (2001) found the prevalence of high-grade dysplasias and cancer of the anal canal to be higher in HIV-positive than in HIV-negative patients, probably because of HPV activity. Diamond, Taylor, Aboumrad, Bringman, and Anton-Culver (2005) thought that the rising incidence of anal cancer among men with AIDS may be related to increased longevity with HAART and the consequent increased time at risk for the development of malignancy and/or the greater use of cytologic screening. The prevalence of high-grade dysplasias and cancer of the anal canal has been found to be significantly higher in HIV-infected than in HIV-negative patients, probably because of HPV activity (Sobhani et al., 2002).

There have been a number of studies designed to evaluate the relationship between HIV infection and anal cancer. These have been difficult to interpret because of the use of different timeframes and the effect of the availability of HAART. Frisch & Goodman (2000b) found a higher risk for anal cancer in both men and women with AIDS, relative risks of 6.8; 95% CI, 0.27-14.0 and 37.9; 95% CI 33.0-43.4, respectively. The authors found that patients with CD4 counts less than 200/mm^3 did not have higher risks than patients with CD4 counts greater than 200/mm^3. This led to the conclusion that while HPV-related malignancies are present in excess, this may be related to unknown cofactors rather than to HIV immunosuppression. Moreover, a population-based study (Chiao, Krown, Stier, & Schrag, 2005) found that the incidence of anal cancer increased from the years before the HIV epidemic, to the years of HIV incidence without HAART, and to the years of HAART. Some experts in the field have suggested that this is a result of longer life

spans of HIV-positive individuals who have a greater period of exposure to HPV, leaving more time for eventual development of squamous cell carcinoma. This points to a relationship between HPV and anal cancer and not to one between HIV and anal carcinoma. The effect of these changes in the trends in anal cancer in HIV-infected persons should be investigated, as part of establishing the true nature of the relationship between HIV infection and anal carcinoma. Most of the research involving the incidence of anal carcinoma in HIV-infected populations is limited to study of infected persons who have progressed to AIDS. With the advent of HAART, HIV-infected persons are experiencing longer survival prior to developing AIDS. At the same time, these individuals may be at greater risk than non-infected populations to the development of anal cancer.

As with many of the less frequent malignancies among the HIV-infected population, data from large-scale studies of the incidence and natural history of colon cancer in HIV-infected persons are limited (International Collaboration on HIV and Cancer, 2000). Most available studies of matched cancer and HIV/AIDS registries have not shown an increased incidence pre- to post-HAART of colon cancer in persons infected with HIV (Cooley, 2003; International Collaboration on HIV and Cancer, 2000; Clifford et al., 2005; Biggar, Kirby, Atkinson, McNeel, & Engels, 2004; Herida et al., 2003; Grulich, Wan, Law, Coates & Kaldor, 1999; & Chiao & Krown, 2003); however, a few studies have reported increased colorectal cancer incidence over the period of implementation of HAART (Bedimo et al., 2004 & Huynh et al., 2007). Additionally, many of the existing published studies on malignancies in HIV-infected persons are based on data from AIDS-based registries. A number of the recent cohort and linked HIV-cancer registry studies have shown that HIV-infected patients are at significantly higher risk than the general population for certain non-AIDS-defining malignancies. However, these studies often lack the power to detect less frequent cancers and may only include selected groups of HIV-infected persons (Chiao & Krown, 2003).

Clark and Glaser (2001) found that although immunodeficiency is important in HIV-associated Hodgkin disease development, research has shown that Hodgkin disease precedes the AIDS diagnosis in most patients and occurs with a relatively high CD4 count. Engels et al. (2008) reported that Hodgkin lymphoma incidence "increased substantially following an AIDS diagnosis but also, somewhat paradoxically, increased over time with the introduction of HAART". These findings create difficulty in interpreting the results of HAART therapy on HIV-related Hodgkin lymphoma. In a recent study of males in the San Francisco Bay Area, Glaser et al. (2003) found that HIV-related Hodgkin lymphoma (HL) had distinctive demographic features, more aggressive clinical characteristics, strong EBV association, and poorer survival. However, patients with HIV-related HL who were diagnosed after HAART was introduced appeared to have less aggressive disease and better survival.

The optimal treatment for Hodgkin disease is controversial according to Tirelli et al. (2000). These authors reported current thinking is that the outcome of AIDS-related Hodgkin disease might improve with some optimal combination of chemotherapy and antiretroviral therapy, specifically with the use of HAART. This is thought to have the potential ability to control the underlying HIV infection during treatment with chemotherapeutic agents. Due to the improvements in survival of patients with HIV and Hodgkin disease experienced after the implementation of HAART, Grulich et al. (1999) warned that the improved survival in HIV-infected persons might lead to increases in the number of cases of selected malignancies including Hodgkin disease.

Hessol et al. (2007) in their research concluded that the effect of potent antiretroviral therapy on the full spectrum of HIV-related cancers is still unknown but if treatment for HIV succeeds at immune restoration, then the incidence of some cancers may decline or at least be delayed. However, the reduced morbidity due

necessary for the development of certain cancers, including those not previously identified as HIV-related. Disentangling the effect of HAART from the effect of cancer treatment is difficult and is made especially difficult because HAART use does not appear to substantially reduce non-AIDS-defining cancer risk overall, and the impact of HAART on specific non-AIDS-defining cancer incidence rates and survival time is not uniform.

di Gennaro, Cinelli, Vaccher, Spina, & Tirelli (2005) suggested that "The impact of HAART on cancer incidence rates is still to be defined and the scenario of HIV and non-HIV-related neoplasms continues to evolve." They also reported that the incidence of invasive cervical cancer in HIV-seropositive women has increased since the introduction of HAART. Herida et al. (2003) found in their study that the advent of HAART has had no measurable impact on the overall incidence of non-AIDS-defining cancers, suggesting that these are not triggered by immune suppression alone. Because they found a slight increase in the incidence of Hodgkin disease between the pre-HAART and HAART periods, they would not exclude an oncogenic effect of HAART on Hodgkin disease.

Bedimo et al. (2004) found a decrease in AIDS-defining malignances for 1997-2002 versus 1989-1996 and a significant increase in non-AIDS defining malignancies. The mean CD4 cell count was lower among those with AIDS-defining malignancies than those with non-AIDS defining malignancies. The authors thought a longer duration of survival during HAART might explain the increasing incidence of non-AIDS defining malignancies. These authors proposed that recent studies that show no associations between the patients' CD4 cell counts and the risk of acquiring a non-AIDS-defining malignancy may mean that prolonged exposure to HAART and the attendant reconstitution of the immune system may not decrease but may actually increase the risk of non-AIDS-defining malignancies. Engels and Goedert (2005) quoted the

Swiss HIV Cohort Study's finding that corroborates the increased risk for cancers of the lung; liver; and cervix, anus and other anogenital sites; and the lack of association with CD4 cell count or AIDS onset.

At least one investigator thought that HAART use decreased the incidence of cancer. "The use of HAART appeared to be beneficial in protecting against the development of malignant disease" (Burgi, et al., 2005). In summary of this topic, Engels (2001) thought that "As HIV-infected persons age, new patterns in cancer incidence may emerge."

AIDS and the Diagnosis of Cancer Pre-HAART and Post-HAART

The trends in the effects of infection with HIV have been studied primarily through investigation of the progression of HIV seroconversion to diagnosis of AIDS or from HIV infection to cancers, both AIDS-defining cancers and those not considered AIDS-defining. Some recent studies have encompassed an analysis of the effects of the availability of HAART therapy.

Engels and colleagues (2006b) reported for the HIV/AIDS Cancer Match Study on trends in cancer risk among people with AIDS from 1980 through 2002. Findings included declines in Kaposi's sarcoma and non-Hodgkin lymphoma after the introduction of HAART, as well as an increase in Hodgkin lymphoma. The study focused on cancer risk relative to the general population in the two years after AIDS development in the pre- and post-HAART eras. In a study of the risk of cancer in elderly persons with AIDS for the AIDS Cancer Match study group (Biggar et al., 2004), the group found lower relative risks than in younger adults with AIDS, but a similar profile of cancer risks.

The three cancers considered as AIDS-defining are Kaposi's sarcoma, non-Hodgkin lymphoma, and invasive cervical cancer. Data provided by the New Jersey Cancer Registry staff from the 2004 AIDS/Cancer match show that for the period before HAART (1979-1995) there were 1,248 Kaposi sarcoma diagnoses in HIV-infected New Jersey residents. The number of similar cases in the period 1996 through 2002 was 298. Data from the SEER 9 areas (2009) indicate that the annual age-adjusted incidence rates for Kaposi sarcoma began to decline in the early 1990s, dropping from a rate of 4.37 per 100,000 population in 1992 to a rate of 0.67 per 100,000 in 2002. Non-Hodgkin lymphoma cases in HIV positive persons in New Jersey also declined, but not as dramatically (from 851 to 579). The New Jersey non-Hodgkin lymphoma incidence rate increased approximately 40% in both males and females between 1979 and 2002, in a pattern that was similar to that in the U.S. through the middle 1960s. Since 1996, the New Jersey male incidence rates have been higher than the U.S. rate, as have the rates in females since 1993. During the two periods, 1979 through 1995 and 1996 through 2002, cancer of the cervix uteri decreased from 401 to 125 diagnosed cases in New Jersey HIV-infected women. The state's cervical cancer incidence rate declined more than one-third from 1979 to 2002, although the New Jersey rates were higher than U.S. rates throughout the period 1979-2001 (Cancer Epidemiology Services, 2005).

In order to provide a parametric model for the estimation of incubation time from infection with HIV to the development of AIDS, Munoz and Xu (1996) compared the fit of Weibull, log-normal and three-parameter logistic models to data from the Multicenter AIDS Cohort (MAC) Study of 1,649 seroprevalent individuals and 476 seroconverters, representing 1,022 and 177 AIDS cases, respectively. Tassie et al. (2002) used these findings and those of other researchers to estimate the median incubation time in three calendar periods, through application of the log-normal distribution to the MAC data. The resulting estimates were 8.0 years from January 1992 to June 1995, (CI: 6.0,10.6); 9.8 years from July 1995 to June 1996 (CI:

8.5,11.2); and 20.0 years from July 1996 to June 1999 (CI: 17.1,23.3). Confirmation of the equality of incubation times was established for the potential mix of transmission mode in the study files. Hendriks, et al. (1998) found that the median time from HIV seroconversion to AIDS, conditional on survival to an AIDS diagnosis, from the end of 1985 through the end of 1996 was 8.2 years. In the study, the authors found that the progression to AIDS conditional on survival to an AIDS diagnosis was essentially identical in both the IDU cases and homosexual men under study (medians of 8.2 and 8.3 years, respectively). Hessol and Herminia (1996) reviewed studies of progression to AIDS and to death after HIV seroconversion that assessed results for varying cofactors, including gender. They found no differences in progression to AIDS by gender between injecting drug users and transfusion recipients or, between injecting drug use and heterosexual contact.

Data are limited regarding AIDS risk in HIV-infected persons with and without a cancer diagnosis. In an Australian study that linked HIV, AIDS, and cancer registries (Grulich et al., 2002), cancer risks were assessed for persons with HIV who did not develop AIDS or were more than 5 years prior to development of AIDS compared to persons from the AIDS registry. Major findings were that HIV-infected persons with mild immune deficiency (who had not developed AIDS) were at increased risk of anal cancer, while a significant trend of increasing relative risk of cancer with increasing time since HIV infection was found for Hodgkin disease and multiple myeloma. The authors note that larger studies with longer periods of follow-up are needed to confirm these findings.

Time to Development of Cancer after HIV Seroconversion in the Pre- and Post-HAART Eras

Although the issue has screening and treatment implications, the question of whether the incubation period of HIV infection to cancer diagnosis would change after implementation of HAART, compared to the pre-HAART era, is one that has not received much attention. A number of studies have investigated the change in time from HIV infection to AIDS affected by the introduction of HAART and the changes in cancer incidence in HIV/AIDS-infected populations before and after implementation of HAART. Ron Brookmeyer (1990) has studied the natural history of disease, reporting on the strengths and limitations of various epidemiologic study designs and sources of data for estimating the duration of the asymptomatic period of disease incidence. Based on his study, an estimate of the incubation period distribution of AIDS is given. The Medical Research Council Clinical Trials Unit in London, UK, has produced a study of the potential biases in estimates of the AIDS incubation period (Porter, Johnson, Phillips, & Darbyshire; 1999), including the recommendation that these must be documented and controlled for whenever possible. These authors also warned that comparison of results from one cohort with those of another may be misleading because of use of different methods.

The issues presented in study of the incubation of HIV infection to AIDS diagnosis may have implications for analysis of possible differences in risks of development of cancer pre- to post-HAART.

Originality of the Research

There have been a number of studies aimed at providing insight into the effects of HAART therapy for persons with HIV infection in changes in life expectancy, susceptibility to other non-AIDS defining

illnesses, cancer incidence rates and median age at diagnosis of cancers, among other issues. Most of these studies suffered from a lack of complete information on person-years of observation with HIV infection, as the registries in most states and municipalities developed to gather data on this condition only registered people once the diagnosis of AIDS had been made. The State of New Jersey expanded its AIDS registry, begun in 1985, to include the required reporting of all resident cases of HIV infection in late 1991. This expansion allows calculation of person-years of observation from the beginning of HIV infection to cancer diagnosis, to death, or survival to a selected date. The unlinked, de-identified file of resident New Jerseyans who had HIV/AIDS, but not cancer, as well as the matched file of HIV/AIDS and cancer cases, was made available to this research so that observation time for all individuals positive for HIV/AIDS could be included in the analysis. This permits the more accurate estimate of person-years of observation and, from this, the expected incidence of cancers in the HIV-infected population for calculation of Standardized Incidence Ratios and incidence rates per 100,000 person-years. Studies that use the national or state AIDS/Cancer Match Registry to investigate the effects of HAART are usually constrained in their analyses by the lack of access to data regarding person-years of observation and characteristics of the HIV-infected individuals who have not received a cancer diagnosis.

The Centers for Disease Control and Prevention recently published the first national HIV incidence figures from a revised methodology that more directly measures the number of HIV-infected persons in the country (CDC, 2009). The initial analysis provided an estimate of 56,000 new cases of HIV/AIDS in the U.S. in 2006. This is considerably higher than the previously accepted estimate of 40,000 new cases per year. CDC points out that the new estimate does not represent an actual increase in incidence, but instead is a result of a more accurate way of measuring newly infected persons. The CDC suggests that the annual

number of new infections was never as low as 40,000 and has been approximately unchanged since the late 1990s.

In a MMWR report from CDC (2005), the agency provided an estimate of more than a million people in the U.S. living with an HIV infection (CDC, 2005). A report on 33 states with name-based HIV data stated that the annual rate of HIV/AIDS diagnoses per 100,000 population did not change significantly from 2001 to 2004 (from 22.8 per 100,000 to 20.7). The report provided data on changes in incidence rate by race/ethnicity, indicating a 9.0% rate decrease among blacks (from 88.7 to 76.3 per 100,000) and a 9.0% average annual increase among Asian/Pacific Islanders. Both of these changes were significant.

Chapter II. Research Objectives

One of the primary goals of this dissertation is to examine the relationship between HIV infection and five malignancies. These five malignancies – lung, colon, prostate, Hodgkin lymphoma, and anal cancer - will be analyzed for changes in incidence after the introduction of highly active antiretroviral therapy (HAART) in HIV/AIDS cases with and without a cancer diagnosis using resident New Jersey data. These data will also be used to investigate changes in mean age at diagnosis of cancer, post-HAART compared to pre-HAART. The hazard of development of AIDS for HIV-infected persons who develop cancer, compared to HIV-infected persons who do not have a cancer diagnosis, pre- and post-HAART, will also be studied. Finally, the change in time to diagnosis of cancer after HIV-infection, pre- and post-HAART, will be analyzed.

Objective 1.

To determine whether the cumulative incidence of each of the five selected cancer types changed in HIV-positive persons from the pre- to post-HAART periods. The null hypothesis is that the SIR ratios did not change from the pre-HAART to post-HAART period, for each of the five cancer types under study. The alternative hypothesis is that one or more of the SIRs changed pre-HAART to post-HAART.

Objective 2.

The second objective is to assess whether the mean age at diagnosis of cancer changed from the pre-HAART to the post-HAART period, for each of the five selected cancer types. The null hypothesis is that

for each of the five cancer types under study, the mean age at diagnosis was not different pre-HAART and post-HAART. The alternative hypothesis is that the mean age at cancer diagnosis was different pre- and post-HAART for one or more of the cancer types being studied.

Objective 3.

The third objective is to determine whether the risk of development of AIDS in the population of HIV-infected persons, when adjusted for covariates, is different for those who have been diagnosed with cancer than for those who have not been given a cancer diagnosis. The null hypothesis is that for the population of HIV-infected persons reported to HARS, a diagnosis of cancer is not related to the development of AIDS, when adjusted for covariates. The alternative hypothesis is that a diagnosis of cancer in HIV-infected persons is related to the development of AIDS.

Objective 4.

The fourth objective is to evaluate if the implementation of HAART changed the risk of HIV infection to cancer diagnosis. The null hypothesis is, for the population of HIV-infected persons who have been diagnosed with cancer, implementation of HAART is not related to the risk of cancer diagnosis, when adjusted for covariates. The alternative hypothesis is that the implementation of HAART is related to risk of cancer diagnosis in HIV-infected persons and differs pre- and post-HAART.

Chapter III. Study Design and Methods

The protocol for this study was determined by both the New Jersey Department of Health and Senior Services' Human Research Ethics Program and UMDNJ's Institutional Review Board (IRB) to be Non-Human Subject Research and therefore not in need of IRB review.

Study Data Files

The AIDS Cancer Match Registry Study is a continuing joint project of the National Institutes of Health/National Cancer Institute (NIH/NCI) and, most recently, eleven U.S. sites (states and metropolitan areas). The project has linked AIDS and cancer registry data in these sites. In three states, one of which is New Jersey, data are collected on persons with HIV infection, in addition to persons who have developed AIDS. This data collection allows the study of people with HIV infection prior to development of AIDS.

Data from the New Jersey State Cancer Registry and from the New Jersey HIV/AIDS Reporting System were most recently linked in December 2004 to provide a file of cases through the end of 2002 with both HIV and cancer. Data from this linked file and from the unlinked file of HIV cases that did not receive a cancer diagnosis are the sources from which the individual case elements used in the analysis were found. In addition, a file of New Jersey resident cancer incidence rates per 100,000 person-years by cancer type, gender, and age group for each year 1991-2002 was utilized to compute expected incidence for each of the five cancer types under study.

Description of the Registries

The New Jersey State Cancer Registry

In 1992, the Congressional Cancer Registries Amendment Act mandated nationwide cancer registration (Glaser et al., 2005). By 2003, a cancer registry was operating in every state. Population-based state registries collect data on new cancer cases from hospital reports, medical records, pathology reports, hospital discharge abstracts, and death certificates (Izquierdo & Schoenbach, 2000). The legislation authorized the Centers for Disease Control and Prevention (CDC) to provide funds to states and territories to improve existing cancer registries; plan and implement new registries; develop model legislation and regulations for states to enhance the viability of registry operations; set standards for data completeness, timeliness, and quality; provide training for registry personnel; and help establish a computerized reporting system.

When the CDC created the National Program of Cancer Registries, 37 states had laws authorizing state cancer registries, 14 had all enabling regulations set, and only nine had both components. New Jersey was one of the nine states with both authorizing laws and enabling regulations. The New Jersey State Cancer Registry (NJSCR) is a population-based incidence registry that includes all reportable cases of cancer diagnosed in New Jersey residents since October 1, 1978 (New Jersey Department of Health and Senior Services, 2007a). Cancer cases are reported by hospitals, diagnosing physicians, dentists, and independent clinical laboratories.

In New Jersey, primary invasive and *in situ* neoplasms are reportable, except cervical cancer *in situ* diagnosed after 1994 and certain carcinomas of the skin. At the time of the linkage in 2004, the NJSCR was current through 2002. Case ascertainment was estimated to be 100% complete. Reporting agreements are in effect with New York, Pennsylvania, Delaware, Florida, Maryland, and North Carolina for New Jersey residents treated in hospitals in these states.

The HIV/AIDS Reporting System (HARS)

AIDS has been a reportable disease in New Jersey since 1985, and HIV has been reportable since late 1991 (New Jersey Administrative Code Title 8, 2003). The HARS was created in 1993 and incorporated data from the previously existing AIDS reporting system, and from an interim HIV reporting system. The AIDS reporting system included cases reported from 1981 to 1993; the interim HIV system included cases reported from 1991 to 1993. There were 61,277 cases of HIV/AIDS on the unmatched file from HARS who had never been diagnosed with cancer. These cases were in addition to the 3,121 HIV/AIDS cases on the matched file who had received a cancer diagnosis.

An HIV case is defined as a person diagnosed and reported to HARS with HIV infection (New Jersey Department of Health and Senior Services, 2007b). An AIDS case is defined as a person with HIV infection who has been diagnosed with an AIDS-defining opportunistic infection or whose CD4 (+) count is less than 200 cu mm or whose proportion of CD4 (+) T-lymphocytes is less than 14 percent of his/her total lymphocytes. An HIV/AIDS case is a person diagnosed and reported with either HIV or AIDS. Persons living with HIV/AIDS at the end of the reporting year are considered prevalent cases; those diagnosed during the reporting year are considered incident cases. A person, previously diagnosed

with HIV infection (but not AIDS) in HARS whose disease progresses to AIDS status, is considered an incident AIDS case for the year in which the AIDS diagnosis was made.

Data collected in HARS include gender, birth date, race/ethnicity, vital status, modes of exposure, year of diagnosis, date of report and residence for each case. The CDC-designed case report forms are completed by providers, and/or New Jersey Department of Health and Senior Services staff, based on a review of medical records. Records in HARS are updated based on laboratory reports received from testing laboratories. The HIV/AIDS surveillance system has been estimated to be 90 to 95 percent complete.

National AIDS/Cancer Match Registry

The NJSCR and HARS staffs each created a single match-compatible file from which the National Cancer Institute (NCI) and its contractors obtained all the data for the match of the two files, as well as for the National AIDS/Cancer Match Registry (NACMR) files that are returned to NCI. Prior to the match the files were standardized using commercial software. A data dictionary was developed and compiled for each of these files and an NACMR number was assigned to each entry in the files (NACMR, 2004).

The next step in the matching process involved preparing the match specifications. These included assigning each name in the database to a block based on the soundex of the last name. Matching names were then compared with social security numbers, first names, middle names, birth dates, etc. to

determine matches, depending on the degree of uncertainty allowed for each variable and the degree of reliability of the variable field.

The matching process produced a system of scores for each pair of individuals from the two files. Pre-determined scores were used to accept matching scores greater than a certain magnitude, to omit pairs whose scores fell below a certain limit, and to consider those cases falling in-between these limits as needing further review. Authorized staff made the final decisions as to which records were to be included as matches. Selected staff manually reviewed the remaining potential matches. The manual review may have involved phone calls and, in some cases, review of medical records in doctors' offices or in HMOs.

After matching, all personal identifiers were deleted from the analysis data set. The process resulted in a matched file of HIV/cancer cases which could be used for research purposes and (1) an unmatched file of HIV positive persons without cancer and (2) a separate file of individuals with cancer diagnoses, but not HIV, to be retained by the respective registry staffs. This process yields highly reliable data containing no confidential information about subjects.

New Jersey AIDS/Cancer Match File

Once the matches from the HIV/AIDS and cancer registries were identified, the software was used to write out the appropriate files of unmatched residuals and the match linkup files, eliminating all identifying information, but keeping the match score and linkage numbers between records.

De-identified unlinked data from the New Jersey AIDS/Cancer Match File (NJACM) are available from this file and can be extracted for research studies. Similar variables can be extracted from the unmatched files of (1) HIV positive persons with no cancer diagnosis to be retained by HARS and (2) individuals with cancer diagnoses but no indication of HIV infection to be retained by the Cancer Registry. Appendix A contains the Cancer Aggregate Comparison File and Appendix B the AIDS Aggregate Comparison File, a copy of each of which is retained by both registry staffs. The layout of the file of variables to be used for the current study is found in Appendix C. None of the three files (matched or unmatched) contains variables which can be used to identify or link individuals to identifiers.

Invasive cancers are coded according to the International Classification of Diseases for Oncology, second edition (ICD-O-2), for cancers diagnosed from 1979 through 2000, and according to the revised coding in the International Classification for Diseases for Oncology, third edition (ICD-O-3), for cancers diagnosed in 2001 and later. The data by cancer type were analyzed by site using the Surveillance, Epidemiology and End Results (SEER) program's "Site recode with KS and mesothelioma" (SEER, 2008). The new ICD-O-3 classification system had some changes in coding compared to the preceding classifications in ICD-O-2 (Biggar et al., 2005). However, these changes applied mostly to the reclassification of some Non-Hodgkin lymphomas in ICD-O-2 to non-specific codes in ICD-O-3. In particular, changes from ICD-02 to ICD-03 affected only the site recodes for non-melanoma skin, Non-Hodgkin lymphoma, myeloma, leukemia, and miscellaneous cancers diagnosed after 2000 (personal correspondence from New Jersey Cancer Registry staff). These changes did not impact the analysis of the five non-AIDS defining cancers under study or the total cancer cases included in the matched HIV/AIDS-Cancer file.

Matched cases in which the diagnosis of cancer precedes the diagnosis of HIV infection are excluded from analysis for this study. The file of matched HIV/AIDS and cancer cases provided to the NJSCR by the National Cancer Institute contained 5,785 cases. Forty additional cancer cases (Kaposi Sarcoma and non-Hodgkin lymphoma) were added to the matched file identified by NJSCR staff from the data linkage process. Of the 5,825 matched file cases, 2,704 were excluded, some for more than one reason for exclusion. The list of reasons for exclusion of cases from the matched file used for analysis follows:

Reason for Exclusion of Cases	**Number**
Cancer diagnosis before 1979	1
Missing year of cancer diagnosis	57
Benign tumor (not cancer)	1
Second or later primary cancer	366
Non-reportable cancer	421
Missing year of AIDS/HIV diagnosis	59
HIV diagnosis after 2002 (for HIV only cases)	21
AIDS diagnosis after 2002 (for AIDS cases)	54
HIV diagnosis after cancer diagnosis, HIV and cancer diagnosis occurred in the same month, or unknown if HIV after cancer diagnosis (for HIV only cases)	275
AIDS diagnosis after cancer diagnosis, AIDS and cancer diagnosis occurred during the same month, or unknown if AIDS after cancer diagnosis (for AIDS cases)*	1,946
Duplicate or erroneous match	22

*Cases that were diagnosed with AIDS after the cancer diagnosis date were included on the analysis file if they were diagnosed with HIV before the cancer diagnosis date (n=322)

In the final matched file used for analysis, non-reportable cancers included certain non-melanoma skin cancers and *in situ* cervical cancers; these cases were excluded from analysis. Cases reported to HARS with AIDS after 2002 were excluded, but cases reported with an HIV diagnosis before 12/31/02 and diagnosed with AIDS after 2002 were not excluded from the analysis file.

The NJSCR records race and Hispanic ethnicity as two separate characteristics, while the HARS converted its coding of race/ethnicity from a single characteristic to separate reporting of race and Hispanic ethnicity in recent years. The data analysis by race/ethnicity in this report utilized the reporting of white non-Hispanic, black non-Hispanic, Hispanic (of any race), and other/unknown race/ethnicity from the HARS file. The use of the HIV/AIDS race/ethnicity data was necessary to achieve consistency in the definition of race/ethnicity in data from the unmatched HIV/AIDS file and the matched file of HIV/AIDS and cancer cases, the two files used to form a concatenated file for selected portions of the analysis.

Children under the age of 15 on either the unmatched HIV/AIDS file or the file of matched cases of HIV/AIDS and cancer were excluded from analysis. A total of 1,169 HIV/AIDS cases of the 61,277 unmatched cases were omitted from the analysis because of the age restriction. The file of HIV/AIDS cases over the age of 14 without a cancer diagnosis encompassed 60,108 individuals. Additionally, 9 cancer cases of the total 3,121 cases with both HIV/AIDS and cancer had an age at cancer diagnosis under the age of 15 and 13 HIV/AIDS cases on the matched file were diagnosed under the age of 15. Of these, 8 had both age at cancer diagnosis and age at HIV infection under the age of 15, leaving 14 individuals excluded for young age from the matched file. The matched file of cases used for analysis consisted of 3,107 persons aged 15 and over.

Chapter IV. Statistical Analysis

Imputation of critical missing data

Year of reported HIV seroconversion and age at HIV infection were critical variables for the assessment of the study hypotheses. During the early years of the epidemic when only AIDS cases were reportable in New Jersey, a number of individuals were reported to HARS with an AIDS diagnosis, but no year or age at HIV diagnosis. Of the 63,213 cases in the study population, 10,639 (16.8%) had no reported year of HIV infection and 11,020 (17.4%) did not have a stated age at HIV seroconversion. These two variables are critical for estimation of person-years of observation as well as for analysis of each of the remaining three objectives. Estimation or imputation of missing values for these two variables was needed to allow inferences to be made.

Current methods of estimation of missing data involve the use of information available on the matched and unmatched files through imputation of the missing values using information from other variables available on the file. For purposes of imputation, missingness is characterized by the degree of randomness. According to Garson (2008), values are considered to be missing completely at random when the missing values are randomly distributed across all observations. This state can be confirmed by use of t-tests of mean differences comparing cases with missing and non-missing data on other pertinent variables. If the missing values meet the test of missing completely at random, then it is acceptable practice to delete the cases with missing data. Otherwise, the missing values should be imputed.

When the missing values are not distributed completely at random across all observations, but are randomly distributed within one or more subsamples, the missingness is considered to be systemically missing. There are a number of strategies which have been developed to impute values when this condition exists (Yuan, 2002 & Barzi & Woodward, 2004).

Among the methods that can be used to estimate missing data when the data can be assumed to be systematically missing are unconditional mean, conditional mean, multiple hot deck, expectation maximization, and a variety of multiple imputation strategies. In unconditional mean imputation, the study-specific mean for all reported values is substituted for the missing value. The conditional mean strategy (also called "cold deck"), is a similar method, however the population data are cross-classified according to values of the predictor variables and the missing cases are assigned the mean value of their cross-classification. Hot deck imputation is akin to cold deck imputation, but instead of receiving the mean value of the subject's cross-classification, the case receives a randomly drawn value from a random sample of the same size as the group that has reported data for the variable to be imputed. Expectation maximization provides maximum likelihood estimates in the presence of missing data. This requires specifying a joint probability distribution for the variable to be imputed and the predictor variables. Multiple imputation is a refinement of expectation maximization which introduces a random component into the process, through the use of a Bayesian technique to replace the deterministic expectation and maximization steps with stochastic equivalents. Given a sufficient number of iterations, the algorithm converges to the Bayesian posterior distribution. Barzi and Woodward (2004), in their comparison of results of imputations from 28 cohort studies, found that studies that had less than 10% missing values had similar results from any of the methods examined, while those variables with 10% - 60% missing data had clear differences in results. From their findings and that of other research, they concluded that multiple imputation was the preferred method of

estimation of missing data in cases where the level of missingness was moderate (10%-60%).The rates of missing date and age at HIV infection on the combined file of matched and unmatched cases in this study were 16.8% and 17.4%, respectively. In addition, the missing values for year and age at HIV infection were not distributed at random, but were concentrated in the early years of the epidemic when only AIDS was reportable to the registry, but not HIV infection alone (data not shown). This lack of randomness meant that omitting the cases with missing values would lead to biased results (Garson, 2008). The rates of missing values led to the decision to use multiple imputation to estimate the year and age for those cases where it was unreported. This method generates multiple simulated values for each missing case and the dataset is iteratively analyzed with each simulated value in turn.

There are several available methods of multiple imputation. IVEware (Raghunathan, Solenberger & Poles, 2002) is a SAS® callable statistical software package that can perform multiple imputations of missing values as well as generate a number of analyses for both descriptive and model-based survey statistics. IVEware creates imputed values through a multivariate sequential regression approach and can fit linear, logistic, polytomous, Poisson and proportional hazard regression models using repeated replication to estimate sampling variances and multiple imputation when there are missing values.

The SAS® software package also contains multiple imputation procedures, PROC MI and PROC MIANALYZE. Inference from multiple imputation in these procedures requires three steps: missing data are imputed the number of times selected (usually 5 to 10); the data sets created from these imputations are analyzed in standard ways (such as univariate analysis); and the results from the data sets are combined for the inference (Yuan, 2002). There are three available MI procedures in SAS®. The method of choice for an arbitrary missing data pattern is a Markov Chain Monte Carlo (MCMC) method which assumes

multivariate normality. In a Markov chain a number of random variables exist such that each element is dependent on the preceding one, resulting in a chain long enough to stabilize to a common distribution. The MCMC method creates multiple imputations by using simulations from a Bayesian prediction distribution for normal data, which simulate the entire joint posterior distribution of the unknown quantities. This enables the process of obtaining the simulation-based estimates of posterior parameters that are needed.

Robustness of the process was assessed through use of a sensitivity analysis with comparison of results from the calculation of lung cancer SIRs using IVEware and MCMC imputation. Table 1 provides results of the approximation of relative risk through the SIR ratios from IVEware and MCMC imputations.

Table 1. Comparison of the ratio of lung cancer SIRs pre- and post-HAART by two methods: IVEware and MCMC imputation of missing data

Period	Observed	IVEware			MCMC		
		Expected	SIR	SIR Ratio	Expected	SIR	SIR Ratio
Pre-HAART	**81**	**28.1**	**2.89**	**0.24 (0.18,0.32)**	**31.5**	**2.57**	**0.25 (0.18,0.32)**
Post-HAART	**185**	**264.6**	**0.70**		**292.0**	**0.63**	

Both year of HIV infection and age at infection have a major effect on the estimated distribution of individuals contributing to the person-years of observation. The two multiple imputation methods tested gave the results for lung cancer shown in Table 1. Although the methods yielded different numbers of expected cases, there was no significant difference in the relative risks (ratio of SIRs) and confidence intervals. These findings were consistent with those of Yu, Burton, and Rivero-Arias (2007) who concluded from their study of multiple imputation of semi-continuous data that when the degree of missingness was

less than 20 percent, all of the methods they tested including SAS MI and IVEware gave results that were quite consistent, resulting in less than 1 percent bias. A major advantage of IVEware is that it imputes missing values for all variables; however the database for this study could not make use of this feature due to the lack of variables that applied to all or most of the cases included in the database. A large percentage of the cases had not yet been diagnosed with AIDS and only a small percentage of HIV-infected persons had received a cancer diagnosis, so that imputation of missing values of variables relating to these conditions would be not be useful.

Barzi and Woodward (2004) concluded from their study of imputation methods that, although MI imputation in SAS assumes normally distributed data, only for studies with many missing values was there a difference between the results obtained from models specifying a multivariate normal model and those more suitable to mixed data types. They presented results of simulation studies that show that MI is generally robust to departures from normality and model misspecifications when the amounts of missing data are not large. In view of the needs of the current study to impute values primarily for year of HIV infection and age at which the seroconversion occurred, the SAS procedure MCMC was used to generate multiple imputations.

In order to estimate person-days of observation, it is necessary to have month, day, and year of HIV seroconversion. The day of HIV diagnosis is not available on the study files. To compute duration of observation, the day of HIV diagnosis was assumed to be the 15^{th}. Missing months were assumed in every case to be "06".

Analysis of Competing Risks

Competing risks are defined as the occurrence of more than one type of outcome, where experiencing one outcome precludes experiencing any other (Shiels et al., 2008). The theory of competing risks allows assessment of the effects of prognostic factors on the AIDS-specific hazard. The competing risk model instead of the usual survival analysis model is preferable in assessing the risk of developing cancer in subjects who are at risk of dying from competing causes, in this case, from AIDS. The usual survival analysis model assumes that subjects who died from other competing causes such as AIDS (censored) would develop cancer if they had not died from other competing causes. This assumption is likely to result in overestimation of the cancer incidence as it does not reflect the usual pattern of death in this population. Ignoring competing risks may lead to biased estimates of results.

The competing risks in these analyses arise because a death in these cohorts may have AIDS as the underlying cause or it may result from any other cause. Neither the New Jersey AIDS/Cancer Match File (NJACM) cases nor the unmatched file of HIV/AIDS cases (HARS) contain the underlying cause of death, thus the competing risks could not be taken into account.

Shiels et al. (2009) reported on three informal methods of accounting for competing risks when estimating the effect of HAART. These authors make the assumption that the competing risks were independent, although conceding that it is not possible to determine whether the risks are independent or dependent from observed data. Each of these methods assumes that competing cause data are available.

The first ad hoc method proposed is to censor each competing risk (in this case, death from a specific cause) as it occurs. This is the most commonly used method to handle competing risks, although in many cases the method is used without acknowledging that this is a competing risk method. As it entails having cause of death information for censoring, it cannot not be performed for this study. This method produces a cause-specific hazard ratio.

The second method put forth by Shiels and colleagues is to exclude cases with competing risks. This analysis produces a conditional hazard ratio, which only applies, in the current study, to those cases that have not died from the competing risk. As this requires knowledge of the cause of death, it also cannot be carried out with the available data. This method produces an estimate of a conditional hazard ratio.

The third method extends all competing risks to the study cut-off date and censors them at that point. This method produces an estimate of the subdistribution hazard ratio, but requires the inverse probability of censoring weights to provide a consistent estimator.

When competing risks are dependent, the first method, censoring competing risks (such as death by cause), produces estimates of the marginal hazard that were less biased than those produced by excluding competing risks or censoring to the end of the study. The authors recommend investigating the use of more formal methods, as the informal methods may induce bias.

A potentially more useful method proposed by the CASCADE Collaboration. (2002) suggests that additional information on the file could serve to project whether a death was caused by AIDS or not. That is, if an individual was reported to have progressed to AIDS, the underlying cause of death was assumed to

be AIDS and deaths to HIV-infected persons without a reported AIDS diagnosis would be censored. Although cause of death information is not required for this competing risks analysis, the authors state that there are limitations and sources of bias that may influence results from this method. First, the progression to AIDS may not have been reported to HARS. In addition, the hazard of developing AIDS (or dying without an AIDS diagnosis) is not constant over time following seroconversion and the incidence of specific AIDS events (and of death without AIDS) is confounded with time. Finally, study is needed of cohorts with underlying cause of death information to determine the rates of deaths in persons with AIDS who die of other likely causes including homicide, suicide, and unintentional injuries.

In addition to the validity problems noted by the report provided by the CASCADE Collaboration, an assumption is made in that study that the progression to AIDS is an available data element. However, data files made available for this dissertation do not include data that allow assessment of progression to AIDS. The CD4 counts that are included on the available data files are taken at one time, and may not provide current status. Additionally, reporting of CD4 counts is incomplete, available on only 56% of records of HIV-infected persons without cancer and not included on the file of HIV-infected individuals with a diagnosed cancer. Viral load, another variable that could be used to assess progression to AIDS, is not available on research data files made available for this study. Thus no attempt to assess the progression to AIDS can be made with the existing data.

Due to the lack of data on cause of death in the files available for study, the data will not be assessed for competing risks, which means that the hazard ratios resulting from the analyses may be overestimated. The CASCADE Collaboration (2002) assessed the differences between a competing risks and a cause specific proportional hazards model on a pooled data set from 20 cohorts and found that the differences were small

when percent risk reduction in 1997-2001 was compared with 1994-1996. The competing risks model yielded 17% risk reduction (9% increase-37% reduction) while the cause specific model resulted in 20% risk reduction (5% increase-39% reduction).

Statistical Analysis for Objectives 1 through 4

For analysis of data for the first objective, which tests the hypothesis of change in incidence of five selected cancers pre- and post-HAART, SIR ratios of observed and expected cases were required. The observed cases and expected cases for the two periods of study, 1991 through 1995 and 1996 through 2002, were summed for each of the five cancer types to form the SIR ratios during each of the two HAART periods.

The calculation of person-years of observation among the HIV-infected population in New Jersey who were alive during any part of the pre-HAART period, defined as 1991 through 1995, or for any time during post-HAART years defined for this study as the years 1996-2002, was accomplished through computation of the days of observation of all HIV-positive persons present in the HARS registry for each year of the entire study period, 1991 through 2002. These data were required for estimating expected cancer incidence and for forming the denominators for calculation of cumulative incidence rates. The data were available from two separate sources: the file of unmatched HIV/AIDS (HARS) cases with no reported cancer diagnoses through the end of 2002 and the NJACM file of matched cases of cancer and HIV infection.

Persons who were diagnosed with cancer before 1991 and/or who died before the beginning of 1991 were excluded from the calculation of person-years of observation. The numbers of observed person-years by age group/gender in each year during the period 1991 through 2002 were contributed by individuals who

fell into any one of three categories: (1) HIV-infected persons diagnosed with cancer between 1991 and 2002: (2) persons infected with HIV who died sometime between 1991 and 2002 without a cancer diagnosis; and (3) persons positive for HIV infection at any point during the study period who were not diagnosed with cancer and did not die before the end of 2002.

Person-years for individuals who were infected with HIV prior to 2002, did not die before 1991, and were diagnosed with cancer between 1991 and 2002 were calculated from data available on the matched file of HIV/AIDS/cancer cases. Current age was estimated by adding age at HIV infection (reported or imputed) to current year of estimation and subtracting year of HIV infection. Observation person-days for these persons consisted of the number of person-days by current age in five-year age groups and gender by year, for each year (1991 through 2002) from the matched file. The person-days of observation were estimated by subtracting the month, day, and year of HIV infection (reported or imputed) from the month, day, and year of cancer diagnosis for each HIV-infected person. The person-days were then distributed by gender and five-year age group for each of the years 1991-2002.

Information on persons who were infected with HIV prior to 2002, did not die before 1991, but did die before the end of 2002 was available on the unmatched file of HIV/AIDS cases without a cancer diagnosis. Person-days for this group were calculated by summing, for cases that died after 1990 and before 2003, the number of person-days for each year in which the individual was alive, by five-year age group and gender by year until death for each of the cases from the unmatched file (cases from the matched file would have been eliminated from this sum because of receiving a cancer diagnosis prior to death). The person-days of observation were estimated by subtracting the month, day, and year of HIV infection (reported or imputed) from the month, day, and year of death for each HIV-infected person. For all persons on the unmatched file

who had a year of HIV infection less than or equal to the current year (each of the years 1991 through 2002, separately), the number of years between the current year and the year of HIV infection was added to the age at HIV infection, to estimate age in the current year. The person-days were then distributed by gender and five-year age group for each of the years 1991-2002.

The remaining group of HIV-infected persons who contributed person-years of observation was composed of those who did not die before 1991, remained alive through the end of 2002 and were not diagnosed through 2002 with cancer. Person-days for this group came from the unmatched file and included the days of observation for those on the unmatched file who were indicated as alive on HARS from date of infection to the end of 2002. For all persons who were infected with HIV infection in a year prior to or in the current year (each of the years 1991 through 2002, separately), the number of years between the current year and the year of HIV infection was added to the reported or imputed age at HIV infection to estimate age in the current year. The person-days of observation for each of the study years were estimated by subtracting the month, day, and year of HIV infection from the last day of the current year (12/31/1991, 12/31/1992, etc.) for HIV-infected persons. The person-days were then distributed by gender and five-year age group for each of the years 1991-2002.

Person-days of observation for each of the cancer types for each year consisted of the sum of person-days from the three sources defined above. Person-years by gender and five-year age group were obtained through division of persons in each category by 365.25, the number of days in a year. The New Jersey incidence rates per 100,000 person-years for calculation of expected incidence were provided by the New Jersey Cancer Registry by cancer type, year of incidence (1991 through 2002), gender, and age within five-year age groups. The estimated person-years for HIV-infected persons were multiplied by New Jersey age-

and gender-standardized incidence per 100,000 person-years for each of the five study cancer types for each year. The male person-year estimates were used for calculating the age-standardized expected number of cases of prostate cancer and both the male and female estimates by age were used to compute separate estimates of expected cases for the remaining four cancer types under study. The expected numbers of cases for each of the five study cancers were summed for the years 1991 through 1995 and for 1996 through 2002 to form the cumulative expected incidence in the pre-HAART and post-HAART periods, respectively, for calculation of the SIRs.

In addition to the person-years for each of the study years, 1991-2002, total numbers of observed cases and person-years of observation were needed for the two periods of study, 1991 through 1995 and 1996 through 2002 for calculation of cumulative incidence rates. Some individuals who were HIV-infected and alive for part or all of more than one of the two periods, 1991 through 1995 and 1996 through 2002, contributed person-days to both of these two periods. These estimates differ from those calculated for each of the study years 1991 through 2002. Individuals could be included, along with their person-years, any number of years up to twelve (1991 through 2002) in the estimation of person-years, but at most would have only been counted twice in the estimation of observed persons and person-years for pre-and post-HAART periods. The numbers of observed persons and person-years of observation for those who died in either the pre- or post-HAART period, those who survived to the end of either one or both of the two periods, and HIV-infected persons who were diagnosed with cancer in either of the two periods are shown in Table 2a. Total numbers of observed persons and person-years for each of the two periods, pre- and post-HAART, are also provided. A second table (Table 2b) includes identical information for male HIV-infected persons and person-years of observation for each of the two periods.

As the means of summarizing age-specific incidence rates, cumulative rates for each of the five study cancer types were calculated for the study period (Breslow and Day, 1987). This was accomplished through division of the incidence by the person-years for ten-year age groups within the ages 30 through 69 for each of the cancer types except prostate cancer for which the ten-year age groups 30 through 79 was used. The study period encompassing the pre-HAART period (1991-1995) and the post-HAART period for which data were available (1996-2002) was divided into five time periods: 1991-1992; 1993-1995; 1996-1997; 1998-1999; and 2000-2002 for this analysis. The cumulative incidence rate for each cancer type and for each of the two periods (pre- and post-HAART) were calculated by summing the incidence for each cancer type and each ten-year age category for the grouped years within each of the two study periods, separately. The incidence for each ten-year age group for the pre-HAART period (1991-1992 and 1993-1995) was then divided by its respective person-years of observation. This process resulted in a set of incidence rates by ten-year age groups for the pre-HAART period and a similar set of incidence rates for the post-HAART group (summing the data for 1996-1997; 1998-1999; and 2000-2002). The sum of the incidence rates by age group within the two study periods resulted in the cumulative incidence rates for pre- and post-HAART periods for each of the five cancer types. Standard errors for the cumulative incidence rates were also computed. This process used the methodology presented by Breslow and Day (1987) in Table 2.8 on page 60 of *Statistical Methods in Cancer Research Volume II – The Design and Analysis of Cohort Studies*, utilizing equations 2.1 and 2.2 for calculation of cumulative rates and their standard errors, respectively.

The Standardized Incidence Ratios (SIRs) for each of the cancer types were calculated for each of the two periods, pre- and post-HAART, by dividing the observed cases for the respective period by the expected cases. The observed cases of each of the five cancer types for the two periods (1991-1995 and 1996-2002)

were obtained from the matched file of HIV/AIDS and cancer cases (NJACM) and formed the numerators of the SIRs.

Approximate 95% confidence intervals for the SIRs were computed from formulae developed by Rothman & Boice in 1979 and presented in Breslow and Day (1987). The calculation of exact 95% confidence intervals for the ratios of SIRs followed the methodology provided by Breslow and Day (1987). These authors stated that the ratios of SMRs (the methodological equivalent of the SIR for mortality) have the same interpretation as relative risk parameters estimated in case control studies. The requirement for this assertion to hold is that the ratios of age-specific rates for different observation categories must be constant over age-calendar year strata. The methodology uses exact 95% confidence limits on the associated binominal probability to estimate the confidence limits for the ratios of two SIRs.

The second objective is to assess whether the mean age at diagnosis of cancer changed from the pre-HAART to the post-HAART period, for each of the selected five cancer types. The matched file of HIV/AIDS and cancer cases was the source of demographic information, including age at diagnosis, for persons who have been reported as having both conditions. Calculation of mean age at cancer diagnosis, median age at cancer diagnosis, and standard errors was accomplished through use of the Univariate procedure in SAS® software for each of the five cancer types under study. The Student's t was used to test the difference, at a significance level of $p = 0.05$ between the mean ages of each of the five cancer types, pre- and post-HAART, assuming equal standard deviations.

The third objective is to determine whether the risk of development of AIDS in the population of HIV-infected persons, when adjusted for covariates, is different for those who have been diagnosed with cancer

than for those who have not been given a cancer diagnosis. The calculation of person-days of observation to assess the effect of a diagnosis of cancer on the development of AIDS in persons aged 15 and over infected with HIV during the pre- and post-HAART eras was accomplished through use of a concatenated file of (1) HIV-positive individuals who had not been diagnosed with cancer and (2) the matched file of HIV/AIDS cases with a cancer diagnosis. Concatenation of the two files required adjusting case numbers to eliminate duplication of these numbers used by the separate HIV/AIDS and Cancer Registry staffs and adjustment of coding differences between the two files. Elimination of duplicate case numbers was accomplished through assignment of a file type code prefix to each record specific to the matched or unmatched file. Coding differences for variables on the two files were managed through SAS® programming. The year of HIV infection and/or age at HIV infection for cases was used as reported or through multiple imputation by the Markov Chain Monte Carlo method where either or both of these were not stated.

To investigate whether the risk of progression to AIDS in HIV-infected persons is changed by a diagnosis of cancer, a regression analysis was performed. For this application, the Cox Proportional Hazards model was used to take advantage of the robustness resulting from its semiparametric form. Cox regression has another advantage that makes it an appealing choice for this type of analysis: ease of incorporation of time-dependent variables. In this analysis, diagnosis of cancer was included in the model as a time-dependent variable, equal to 1 if the patient had developed a cancer at day t or equal to 0 if not. This was assessed as whether the time to development from cancer diagnosis to AIDS was less than the time from development of HIV infection to AIDS. In addition, the availability of HAART was assessed as a time-dependent variable. If AIDS developed in 1996 or later, HAART was coded as 1. Otherwise, HAART was set equal to 0. This model provides the hazard of developing AIDS also taking into account the independent variables: reported or imputed age at diagnosis of HIV (continuous); and race/ethnicity, black non-Hispanic

vs. all other race/ethnicities. Univariate analysis was conducted to provide the probability of equality across strata for each of the potential covariates providing justification for inclusion in the model (data not shown). Variables were included in the model if the probability (p-value) in the univariate analysis was at a relatively low level (< 0.25). Variables could be retained in the model, however, if there was documentation in the literature suggesting a rationale for inclusion, even with a higher probability. The log-rank test of equality across strata was used for analyzing categorical variables and the univariate Cox proportional hazard regression p-value was used to assess continuous variables. All feasible interactions were added to the model, one at a time, to test for significance (data not shown).

The fourth objective is to evaluate if the implementation of HAART changed the incubation time from HIV infection to cancer diagnosis. In order to model survival from HIV diagnosis to cancer diagnosis for the two periods, pre-HAART and post-HAART, the covariates from the matched file of HIV/AIDS and cancer cases were first examined with univariate analysis, specifically the Kaplan-Meier method, to provide the shape of the survival function for the groups and to provide guidance on whether the groups are proportional. Time was measured by the days between month, day, and year of HIV infection diagnosis and month, day, and year of cancer diagnosis. Cases were censored if death was equal to "0" (known death), but not censored for death equal to "1", not reported as dead.

The analysis of whether risk of development of a cancer reportable to the New Jersey Cancer Registry subsequent to infection with HIV changed from the pre-HAART to the post-HAART era was accomplished through application of a Cox proportional hazard regression model to the 2,575 individuals aged 15 or over on the matched file of HIV/AIDS and cancer cases who developed cancer between 1991 and 2002 and the file of 60,108 HIV-infected persons without a cancer diagnosis. Survival to cancer was measured by the

days from HIV seroconversion to cancer diagnosis. Year of HIV infection and age at HIV infection were imputed by Markov Chain Monte Carlo multiple imputation if these were not reported. The available covariates included in the model were year of cancer diagnosis in the pre-HAART and post-HAART periods, treated as a time-dependent variable. Days from date of HIV infection to date of cancer diagnosis was coded 1 if cancers diagnosed pre-HAART had longer incubation from HIV seroconversion than post-HAART cancers and 0 if post-HAART cancers had a longer time from seroconversion than those diagnosed pre-HAART. Covariates included reported or imputed age at time of HIV infection (continuous); gender (male equal to 1 and female equal to 0); and race (black non-Hispanic = 0, all other race/ethnicities = 1).

The data analysis for this paper was generated using SAS software, Version 9.1 of the SAS® System for XP_PRO, copyright 2002-2003 SAS Institute Inc. Multiple imputation of missing data elements was performed using SAS proc MI, Markov Chain Monte Carlo method. SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc., Cary, NC, USA.

Chapter V. Results

Data Files Used for Analysis

The data files used for the analysis of the objectives of this dissertation contain cases on (1) a file of 3,121 persons who have been diagnosed with both HIV infection and/or AIDS (reported to HARS) and one of the cancers reportable to the New Jersey Cancer Registry; (2) a file of 61,277 cases with only HIV infection and/or AIDS (reported to HARS); and (3) a combined file of 64,398 cases. These files are de-identified and unlinked. At the two source agencies and during the matching process, cases were unduplicated.

This analysis was limited to cases of either HIV/AIDS alone or both HIV/AIDS and cancer in New Jersey residents 15 years of age and older. Before imputation of missing values, there were 3,105 cases with stated age of cancer diagnosis of 15 years or more on the matched file of individuals with HIV infection. In addition, there were 60,108 cases on the file of HIV-infected only individuals with the age at seroconversion greater than or equal to 15. The combined file of matched and unmatched cases older than 14 years encompassed 63,213 cases with either HIV/AIDS alone or HIV/AIDS with one or more of the reportable cancers.

There were 12,942 persons who were observed in the pre-HAART period only. These were HIV-positive individuals who were alive and positive for some portion of the years between 1991 and 1995, but either died or were diagnosed with cancer within that time span. Persons observed only during the post-HAART period numbered 21,910. These were persons who had an HIV seroconversion date of 1996 or later. There

were 21,188 individuals observed in at least some portion of both the pre- and post-HAART periods. These persons became HIV-infected before 1996 and did not die or were not diagnosed with cancer before 1996.

Table 2a. Number of HIV-infected persons observed, and number of person-years of observation by period of observation and whether diagnosed with cancer

Period of Observation	Number of persons observed	Number of person-years of observation
Persons HIV-infected before 1996 who died in the period 1991-1995	**11,774**	**26,533**
Persons HIV-infected who had not died by the end of 1995	**13,976**	**43,444**
HIV-infected persons who were diagnosed with cancer, 1991-1995	**1,157**	**2,967**
Persons HIV-infected before 2002 who died in the period 1996-2002	**8,722**	**37,081**
Persons HIV-infected who had not died by the end of 2002	**29,063**	**196,188**
HIV-infected persons who were diagnosed with cancer, 1996-2002	**1,415**	**6,512**
Total persons observed and person-years of observation, 1991-1995	**26,907**	**72,944**
Total persons observed and person-years of observation, 1996-2002	**39,200**	**239,781**

Table 2b. Number of HIV-infected males observed, and number of person-years of observation by period of observation and whether diagnosed with cancer

Period of Observation	Number of persons observed	Number of person years of observation
Males HIV infected before 1996 who died in the period 1991-1995	8,624	19,109
HIV infected males who had not died by the end of 1995	9,084	28,634
HIV-infected males who were diagnosed with cancer, 1991-1995	977	2,516
Males HIV infected before 2002 who died in the period 1996-2002	5,984	24,491
Males HIV infected who had not died by the end of 2002	18,891	127,606
HIV-infected males who were diagnosed with cancer, 1996-2002	1,041	4,655
Total males observed and person-years of observation, 1991-1995	18,685	50,258
Total males observed and person-years of observation, 1996-2002	25,916	156,752

Deaths of HIV-infected persons either with or without a cancer diagnosis totaled 11,774 in the five-year pre-HAART period. Only 8,722 deaths occurred during the post-HAART period which encompasses seven years. HIV-infected persons who died, were diagnosed with cancer or were still alive at the end of each of

the two study periods are shown in Table 2a, along with the number of person-years of observation. Identical information for males who were HIV-infected is shown in Table 2b.

After imputation of age at HIV infection, the individuals aged 15 or over in the matched file of cancer and HIV/AIDS cases totaled 3,085 and the unmatched file of HIV cases alone numbered 59,944, for a combined file of 63,029 cases aged 15 years or more; these latter two files were used for all of the data analysis performed for this dissertation. Characteristics after imputation of the cases on the three files are shown in Tables 3a and 3b.

As seen in Table 3a, a substantial majority of cases of matched HIV/AIDS and cancer occurred in males (79.3%) and more than half of individuals with HIV infection only were of the black race (57.0%). For individuals on the matched file of cases of HIV/AIDS and cancer, the majority of cancer diagnoses occurred

Table 3a. Gender and race/ethnicity for cases of HIV/AIDS or both HIV/AIDS and cancer in individuals 15 years of age and over included in the data analysis

	Matched HIV/Cancer		HIV Infection Only		Combined Cases	
Gender	**Number**	**Percent**	**Number**	**Percent**	**Number**	**Percent**
Male	**2,445**	**79.3**	**41,208**	**68.7**	**43,653**	**69.3**
Female	**640**	**20.7**	**18,736**	**31.3**	**19,376**	**30.7**
Race/Ethnicity						
Hispanic	**496**	**16.1**	**10,729**	**17.9**	**11,225**	**17.7**
Black	**1,434**	**46.5**	**34,173**	**57.0**	**35,607**	**56.5**
White	**1,128**	**36.6**	**14,320**	**23.9**	**15,448**	**24.5**
Multi Race	**15**	**0.5**	**276**	**0.5**	**291**	**0.5**
Other/Unknown	**12**	**0.4**	**446**	**0.7**	**458**	**0.7**

Note: The data file consists of cases 15 years and over after multiple imputation of age at HIV infection

at age 40 and older (54.8%) while only 38.3% of the ages at HIV seroconversions in those who also had a cancer diagnosis occurred at age 40 or above (Table 3b) . HIV positive persons who were diagnosed with cancer prior to seroconversion to HIV infection were not included on the file. Although these demographic differences may not be significant, it would be instructive to investigate the distribution of cancer types by age, gender, and race. In all persons reported to HARS with HIV/AIDS, the modal age group at HIV infection was 30 though 39 years (44.7%).

Table 3b. Age at diagnosis of HIV infection and age at diagnosis of cancer in individuals aged 15 and over included in the data analysis

Age at HIV Diagnosis	Matched HIV/Cancer		HIV Infection Only		Combined Cases	
	Number	Percent	Number	Percent	Number	Percent
15-19	15	0.5	715	1.2	730	1.2
20-29	429	13.9	11,403	19.0	11,832	18.8
30-39	1,459	47.3	26,703	44.5	28,162	44.7
40-49	827	26.8	15,465	25.8	16,292	25.8
50-59	252	8.2	4,239	7.1	4,491	7.1
60+	103	3.3	1,419	2.4	1,522	2.4
Age at Cancer Diagnosis						
15-19	5	0.2	NA		NA	NA
20-29	251	8.1	NA		NA	NA
30-39	1,135	36.8	NA		NA	NA
40-49	1,044	33.8	NA		NA	NA
50-59	467	15.1	NA		NA	NA
60+	183	5.9	NA		NA	NA

Note: The data file consists of cases 15 years and over after multiple imputation of age at HIV infection

A number of persons on both the HIV/AIDS only file and the matched file of HIV/AIDS and cancer cases were missing the age at HIV infection. The dates of HIV infection and age at HIV seroconversion have been imputed on both files for all missing records through use of MCMC multiple imputation and all records include either a reported or imputed date and age at HIV seroconversion.

Table 4a. Number of HIV-infected persons who were on the HIV/AIDS Reporting System, number of person-years of observation and person-years per case, 1991-2002

Year of Observation	HIV-Infected Persons*	Number of Person-years	Person-years Per Person
1991	**5,273**	**10,460.5**	**2.0**
1992	**8,487**	**16,761.1**	**2.0**
1993	**11,564**	**26,824.6**	**2.3**
1994	**14,483**	**38,576.6**	**2.7**
1995	**17,101**	**52,357.7**	**3.1**
1996	**18,849**	**66,115.6**	**3.5**
1997	**20,308**	**81,693.8**	**4.0**
1998	**22,071**	**101,562.2**	**4.6**
1999	**24,052**	**124,003.4**	**5.2**
2000	**26,346**	**148,100.6**	**5.6**
2001	**28,289**	**174,831.1**	**6.2**
2002	**30,181**	**203,076.0**	**6.7**

***HIV-infected persons, alive and without a cancer diagnosis, for any portion or all of the year**

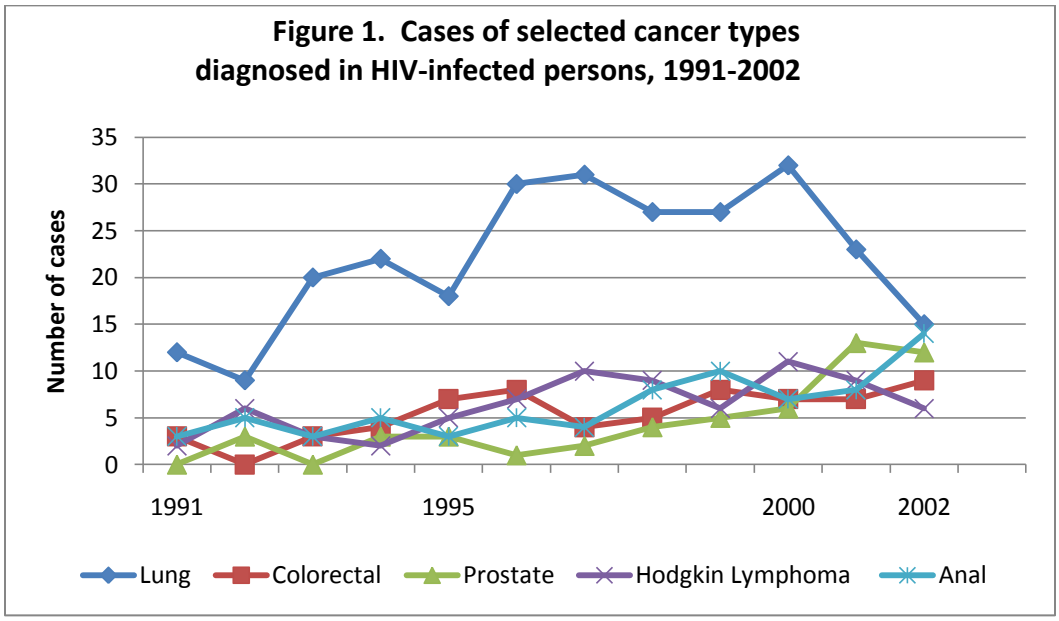
Tables 4a and 4b provide the numbers of HIV-infected persons who were on the HIV/AIDS Reporting System with or without a cancer diagnosis for each year, 1991 through 2002, for the total population and for males, respectively. The tables also show the number of person-days and person-years of observation and person-years per case for these two groups by year.

Table 4b. Number of HIV-infected males who were on the HIV/AIDS Reporting System, person-years of observation and person-years per case, 1991-2002

Year of Observation	HIV-Infected Persons*	Number of Person-years	Person-years Per Person
1991	**3,661**	**7,348.7**	**2.0**
1992	**5,812**	**11,675.4**	**2.0**
1993	**7, 777**	**18,361.8**	**2.4**
1994	**9,686**	**26,055.7**	**2.7**
1995	**11,345**	**35,083.7**	**3.1**
1996	**12,390**	**43,770.2**	**3.5**
1997	**13,216**	**53,623.3**	**4.1**
1998	**14,343**	**66,549.1**	**4.6**
1999	**15,570**	**80,840.5**	**5.2**
2000	**17,126**	**96,435.8**	**5.6**
2001	**18,409**	**113,684.5**	**6.2**
2002	**19,622**	**132,040.0**	**6.7**

***HIV-infected persons, alive and without a cancer diagnosis, for any portion or all of the year**

The average number of person-years of observation in both tables increased steadily over the twelve year period, rising from 2.0 years per observed HIV-infected individual in 1991 to 6.7 years in 2002 for both total and male HIV-infected populations.



Trends in New Jersey from 1991 through 2002 in the incidence of the five cancers under study in individuals also reported as infected with HIV are shown in Figure 1. Except for 1992, more of the individuals diagnosed each year with one of the selected cancers were reported with lung cancer than any other type. Although the numbers are too small to define a trend, the decline in the number of cases of lung cancer in individuals also infected with HIV after 2000 and the concomitant increase in anal cancer incidence made these cancers almost equal in incidence in this population by 2002.

Objective 1.

Cumulative incidence rates calculated for each of the five cancer types under study for the two study periods (pre-HAART and post-HAART) indicated a decrease in the cumulative rate from pre- to post-HAART for each of the cancer types other than prostate cancer (Table 5). There were major differences in the trends over the two periods in incidence rates by age group in the five cancers under study. Although not unexpected, the highest cancer rates for each of the cancer types was found in the oldest age group. As a result of the improving survival of the HIV-infected population, the age group with the highest age-specific rate increased from pre- to post-HAART in three of the cancers under study. Figures 2a through 2e provide the cumulative incidence rates by age group in the both the pre- and post-HAART periods.

Table 5. Cumulative incidence rates, five selected cancers, pre- and post-HAART

	Pre-HAART			Post-HAART		
Cancer	Cases	Cumulative Rate*	Standard Error	Cases	Cumulative Rate*	Standard Error
Lung	78	9.154	2.055	178	3.488	0.446
Colorectal	16	0.953	0.339	45	0.818	0.210
Prostate	9	6.281	2.429	42	7.776	2.050
Hodgkin lymphoma	16	0.911	0.336	55	0.804	0.192
Anal	15	1.644	0.851	55	0.642	0.164

Note: Cumulative rates were calculated using incidence occurring only in age-groups 30 through 69 years, except for prostate cancer for which ages 30 through 79 years were used.
*Cumulative rates are per 1,000 person-years of observation.

The cumulative rates by cancer type and ten-year age group varied by cancer type and, in particular, by whether the time span covered was part of the pre-HAART or post-HAART period. The incidence of each of the cancer types, the person years of risk and the incidence rates for grouped spans of years covering the

Table 5a. Incidence of lung cancer (i) 1991-2002, person-years at risk (n, in thousands) and incidence rates (λ, per 1000 person-years) in HIV-infected persons. Calculation of the cumulative rate and its standard error

Age range	1991-1992	1993-1995	1996-1997	1998-1999	2000-2002
30-39	i 9	9	13	5	3
	n 9.681	30.339	38.344	49.627	62.255
	λ 0.930	0.297	0.598	0.101	0.048
40-49	i 9	31	30	28	24
	n 5.281	22.958	32.992	53.509	97.765
	λ 1.704	1.350	0.909	0.523	0.245
50-59	i 3	11	11	16	25
	n 0.898	4.509	7.311	14.196	39.060
	λ 3.341	2.440	1.505	1.127	0.640
60-69	i 0	6	5	5	13
	n 0.251	1.026	1.510	2.669	7.075
	λ 0.0	5.848	3.311	1.873	1.837
Total	i 21	57	59	54	65
	n 16.111	58.832	80.157	120.001	206.155
	λ 1.303	0.969	0.815	0.450	0.315
Cumulative rate (%)	5.975	9.935	6.323	3.624	2.770
Standard error (%)	2.035	2.387	1.561	0.891	0.528

Table 5b. Incidence of colorectal cancer (i) 1991-2002, person-years at risk (n, in thousands) and incidence rates (λ, per 1000 person-years) in HIV-infected persons. Calculation of the cumulative rate and its standard error

Age range	1991-1992	1993-1995	1996-1997	1998-1999	2000-2002
30-39	i 1	5	3	2	5
	n 9.681	30.339	38.344	49.627	62.255
	λ 0.103	0.165	0.078	0.040	0.080
40-49	i 1	6	5	7	5
	n 5.281	22.958	32.992	53.509	97.765
	λ 0.189	0.261	0.152	0.131	0.051
50-59	i 1	2	3	2	8
	n 0.898	4.509	7.311	14.196	39.060
	λ 1.114	0.444	0.410	0.141	0.205
60-69	i 0	0	0	2	3
	n 0.251	1.026	1.510	2.669	7.075
	λ 0.0	0.0	0.0	0.749	0.424
Total	i 3	13	11	13	21
	n 16.111	58.832	80.157	120.001	206.155
	λ 0.186	0.221	0.137	0.108	0.102
Cumulative rate (%)	1.406	0.870	0.640	1.061	0.760
Standard error (%)	1.135	0.339	0.251	0.542	0.280

periods before and during HAART availability are shown in Tables 5a through 5e. The tables also provide the cumulative rate (as a %) and the standard error also as a percent.

Table 5c. Incidence of prostate cancer (i) 1991-2002, male person-years at risk (n, in thousands) and incidence rates (λ, per 1000 male person-years) in HIV-infected males. Calculation of the cumulative rate and its standard error

Age range		1991-1992	1993-1995	1996-1997	1998-1999	2000-2002
30-39	i	0	0	1	0	0
	n	6.777	19.704	23.433	29.931	36.425
	λ	0.0	0.0	0.043	0.0	0.0
40-49	i	0	0	0	0	3
	n	4.086	17.282	24.097	37.486	65.000
	λ	0.0	0.0	0.0	0.0	0.046
50-59	i	1	3	0	2	14
	n	0.704	3.478	5.560	10.703	29.113
	λ	1.420	0.863	0.0	0.187	0.481
60-69	i	2	3	2	4	8
	n	0.182	0.757	1.063	1.870	5.086
	λ	10.989	3.963	1.881	2.139	1.573
70-79	i	0	0	0	3	5
	n	0.032	0.113	0.160	0.357	0.901
	λ	0.0	0.0	0.0	8.403	5.549
Total	i	3	6	3	9	30
	n	11.781	41.334	54.313	80.347	136.525
	λ	0.255	0.145	0.055	0.112	0.220
Cumulative rate (%)		12.409	4.826	1.924	10.729	7.649
Standard error (%)		7.913	2.342	1.331	1.915	2.546

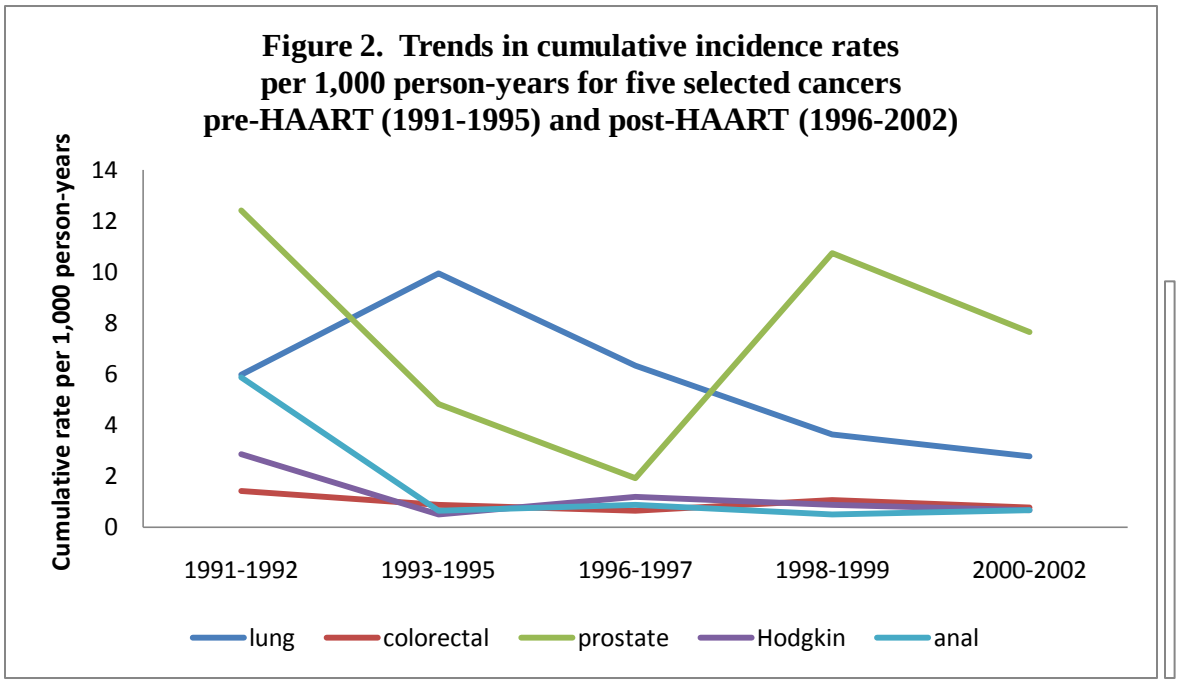
Table 5d. Incidence of Hodgkin lymphoma (i) 1991-2002, person-years at risk (n, in thousands) and incidence rates (λ, per 1000 person-years) in HIV-infected persons. Calculation of the cumulative rate and its standard error

Age range		1991-1992	1993-1995	1996-1997	1998-1999	2000-2002
30-39	i	6	4	7	3	4
	n	9.681	30.339	38.344	49.627	62.255
	λ	0.620	0.132	0.183	0.060	0.064
40-49	i	0	3	7	5	11
	n	5.281	22.958	32.992	53.509	97.765
	λ	0.189	0.261	0.152	0.131	0.051
50-59	i	2	1	1	5	8
	n	0.898	4.509	7.311	14.196	39.060
	λ	2.227	0.222	0.137	0.352	0.205
60-69	i	0	0	1	1	2
	n	0.251	1.026	1.510	2.669	7.075
	λ	0.0	0.0	0.662	0.375	0.283
Total	i	8	8	16	14	25
	n	16.111	58.832	80.157	120.001	206.155
	λ	0.497	0.136	0.200	0.117	0.121
Cumulative rate (%)		2.847	0.485	1.194	0.880	0.665
Standard error (%)		1.595	0.243	0.684	0.410	0.218

Table 5e. Incidence of anal cancer (i) 1991-2002, person-years at risk (n, in thousands) and incidence rates (λ, per 1000 person-years) in HIV-infected persons. Calculation of the cumulative rate and its standard error

Age range		1991-1992	1993-1995	1996-1997	1998-1999	2000-2002
30-39	i	2	6	3	8	11
	n	9.681	30.339	38.344	49.627	62.255
	λ	0.207	0.198	0.078	0.161	0.177
40-49	i	3	0	4	7	13
	n	5.281	22.958	32.992	53.509	97.765
	λ	0.568	0.0	0.121	0.131	0.133
50-59	i	1	2	0	3	3
	n	0.898	4.509	7.311	14.196	39.060
	λ	1.114	0.444	0.0	0.211	0.077
60-69	i	1	0	1	0	2
	n	0.251	1.026	1.510	2.669	7.075
	λ	3.984	0.0	0.662	0.0	0.283
Total	i	7	8	8	18	29
	n	16.111	58.832	80.157	120.001	206.155
	λ	0.434	0.136	0.100	0.150	0.141
Cumulative rate (%)		5.873	0.642	0.861	0.503	0.670
Standard error (%)		4.246	0.358	0.662	0.143	0.215

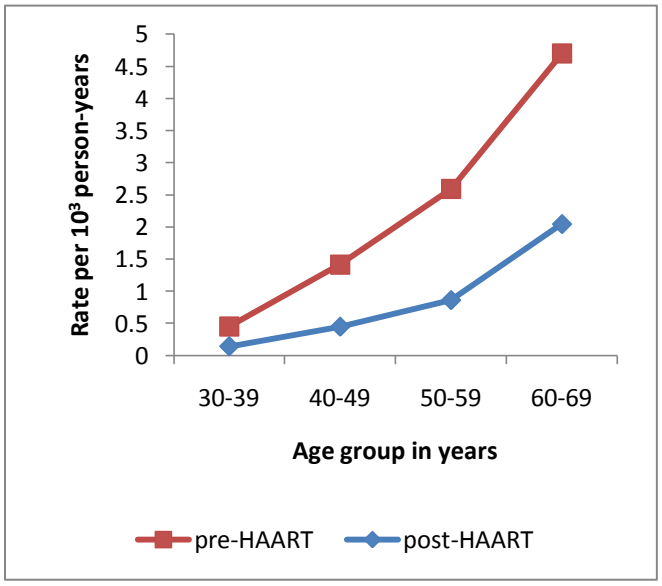
The trends in cumulative incidence rates for each of the five cancer types studied over the pre- and post-HAART periods, are shown in Figure 2, grouped into two- and three-year intervals.



Despite the study results indicating substantially increased cumulative incidence for each of the five cancer types under study, the trends in the cumulative incidence rates per 1,000 person-years pre-HAART and post-HAART indicate that prostate cancer is the only one of the five cancer types under study with an increasing trend in cumulative incidence rate between the end of the pre-HAART in 1995 and the last date of data availability in 2002. Hodgkin lymphoma had a higher cumulative incidence rate in 2002 than in 1995, but the trend during the years of HAART availability was a declining one. The remaining cancer type with an increase in cumulative incidence rate was anal cancer, with a slightly increased rate at the end of the post-HAART period, but with a decreasing cumulative incidence rate throughout the post-HAART period. Both lung and colorectal cancer had cumulative incidence rates that were lower at the end of the post-HAART period than at the beginning.

Age-specific incidence rates differed by cancer type and evidenced changes in modal group pre- and post-HAART.

Figure 2a. Age-specific lung cancer incidence rates in HIV-infected persons, pre- and post-HAART



Lung cancer age-specific rates were highest in the oldest age group, those 60 through 69 both pre- and post-HAART, but the age-specific rates were lower at every age post-HAART (Figure 2a). The lower rates post-HAART occurred despite higher incidence at every age group compared to pre-HAART incidence (data not shown). Greater survival of HIV-infected persons post-HAART resulted in higher estimated

person-years of observation, which led to lower lung cancer incidence rates after the introduction of HAART.

Figure 2b. Age-specific prostate cancer incidence rates in HIV-infected males, pre- and post-HAART

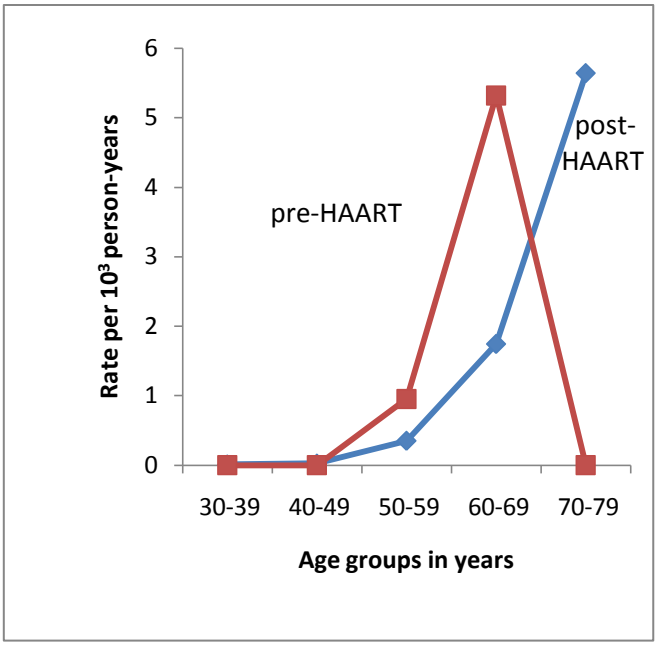
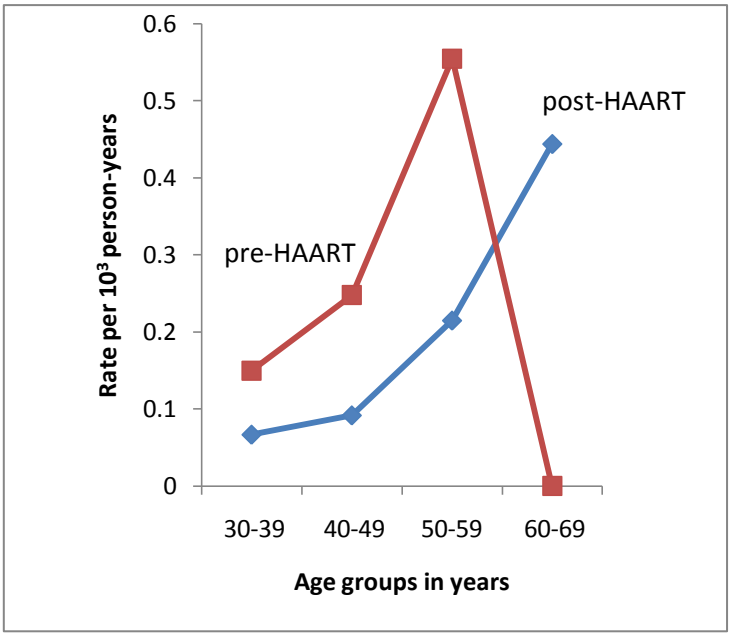


Figure 2c. Age-specific colorectal cancer incidence rates in HIV-infected persons, pre- and post-HAART



Prostate cancer was not diagnosed in the study population of HIV-infected men under the age of 30 and occurred at very low rates until the 60 through 69 age group (Figure 2b). Although there were no reported cases of prostate cancer in 70 through 79 year old HIV-infected men pre-HAART, the rate post-HAART was substantially higher than in all younger age groups. While prostate cancer occurs in men in the general population primarily at older ages, the survival to older ages of HIV-infected men in the post-HAART era is no doubt related to this high incidence rate.

The highest rate of colorectal cancer also occurred in the highest age group post-HAART, but not in pre-HAART (Figure 2c), increasing from no diagnosed cases in 60 – 69 year olds in the pre-HAART period to

the highest rate in this age group post-HAART. While each of the age-specific rates were lower post-HAART, the incidence in every age group was higher than in the pre-HAART period, more than can be accounted for by the somewhat longer time period. The higher incidence is no doubt due to the greater survival of HIV-infected persons.

Figure 2d. Age-specific Hodgkin lymphoma incidence rates in HIV-infected persons, pre- and post-HAART

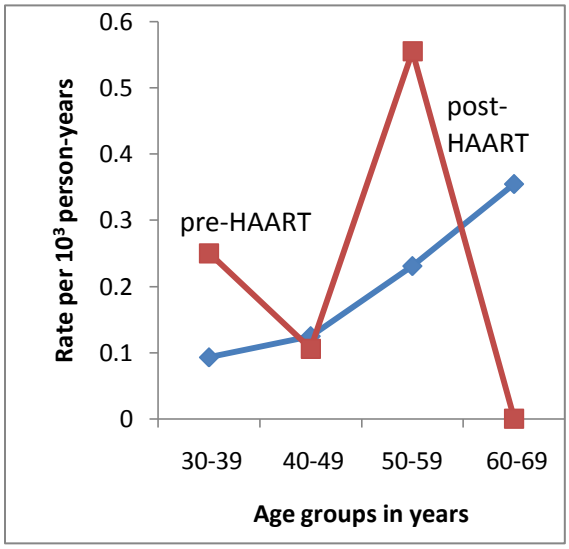
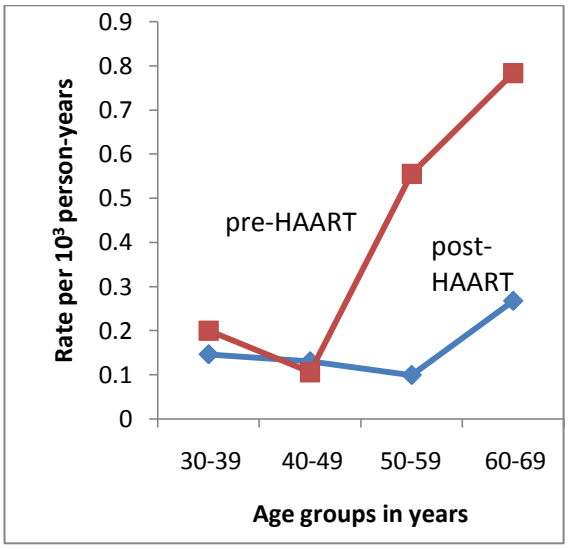


Figure 2e. Age-specific anal cancer incidence rates in HIV-infected persons, pre- and post-HAART



Greater survival of HIV-positive persons is also likely to have caused the relatively high rates of Hodgkin lymphoma in 60 through 69 year olds (Figure 2d). In addition to the high incidence rates at older ages, the 20 through 29 year old HIV-infected population had an incidence rate post-HAART that was virtually identical to that in the 50 through 59 year age group (data not shown). Further study is needed to identify factors related to this finding.

Age-specific incidence rates for anal cancer (Figure 2e) were also highest both pre- and post-HAART for the older population 60 through 69. The incidence post-HAART of anal cancer at younger ages increased

substantially from pre- to post-HAART, from 11 pre-HAART aged 30 through 49, to 46 cases post-HAART. This finding may be related to an increase in papillomavirus in the younger HIV-infected population. Anal cancer in this study population is too small to draw definitive conclusions and needs further study.

When cumulative incidence restricted to individuals who contributed person-years of observation solely in the pre-HAART period is compared to cumulative incidence for persons with an HIV seroconversion year of 1996 or later, all of the cancer types under study had a higher cumulative incidence rate when seroconversion occurred after HAART availability. The increases in cumulative incidence rate were dramatic for colorectal cancer, prostate cancer, and Hodgkin lymphoma (Table 6a). This is in contrast to the cumulative survival for all persons who seroconverted in one or the other period. Only prostate cancer evidenced an increase in cumulative survival when all cases with seroconversion in either period and observation is one or both periods are considered. Prostate cancer had the second-highest incidence rate (after lung cancer) for persons treated only pre-HAART, but had the highest incidence rate for persons treated only post-HAART. These increases are no doubt due at least in part to the extended years of observation of persons infected and living with HIV after the introduction of HAART.

Table 6a. Cumulative incidence of five selected cancers in HIV-infected persons who contributed person-years of observation in either the pre-HAART or post-HAART period, but not both

Cancer type	Pre-HAART Only		Post-HAART Only	
	Cases	Cumulative Rate/ 1,000 Person-Years	Cases	Cumulative Rate/ 1,000 Person-Years
Lung	**78**	**8.3**	**63**	**10.5**
Colorectal	**16**	**0.8**	**20**	**2.8**
Prostate	**9**	**7.3**	**17**	**18.3**
Hodgkin lymphoma	**16**	**0.8**	**22**	**2.4**
Anal	**15**	**1.5**	**23**	**1.7**

Standardized incidence rates (SIRs) were computed to measure the observed number of cases of each cancer type divided by the expected number from application of the New Jersey state rates standardized by age and gender to the person-years of the HIV-infected population. Observed cases were higher than expected for anal cancer and Hodgkin disease in both pre-HAART and post-HAART eras, and for lung cancer in the pre-HAART period. Observed incidence of lung cancer in the post-HAART period and for colorectal and prostate cancer in both periods was less than expected. All of the SIRs were significant except for

Table 6b. Observed and expected incidence of five selected cancers, standardized incidence ratios (SIRs), and SIR ratios among HIV-infected individuals pre- and post-HAART

Cancer type	Observed Cases	Expected Cases	SIR (95% CI)	Ratio of SIRs Post/Pre-HAART (95% CI)
Lung cancer				**0.25 (0.18,0.32)**
Pre-HAART	**81**	**31.5**	**2.57 (2.07,3.24)**	
Post-HAART	**185**	**292.0**	**0.63 (0.55,0.73)**	
Colorectal cancer				**0.26 (0.14,0.50)**
Pre-HAART	**17**	**22.5**	**0.76 (0.44,1.21)**	
Post-HAART	**48**	**243.2**	**0.20 (0.15,0.26)**	
Prostate cancer				**0.24 (0.12,0.60)**
Pre-HAART	**9**	**22.1**	**0.41 (0.19,0.77)**	
Post-HAART	**43**	**423.4**	**0.10 (0.07,0.14)**	
Hodgkin disease				**0.60 (0.35,1.06)**
Pre-HAART	**18**	**6.7**	**2.69 (1.58, 4.23)**	
Post-HAART	**58**	**35.0**	**1.62 (1.23,2.09)**	
Anal cancer				**0.26 (0.15, 0.46)**
Pre-HAART	**19**	**0.8**	**23.75 (14.36,37.00)**	
Post-HAART	**56**	**9.1**	**6.15 (4.65,8.00)**	

colorectal cancer in the pre-HAART period. The SIRs for anal cancer and Hodgkin disease were significantly high in both periods, although in each case the SIR declined from pre- to post-HAART.

The ratio of the SIRs, post-HAART to pre-HAART, was less than 1 for each of the five cancer types under study and, except for Hodgkin disease, each of these SIR ratios was significantly less than 1 at $p = 0.05$ (Table 6b). Hodgkin disease, with a significantly high SIR in both the pre- and post-HAART periods, was

the only cancer type studied that did not exhibit a reduced relative risk, post-HAART compared to pre-HAART.

Objective 2. The second objective is to assess whether the mean age at diagnosis of cancer changed from the pre-HAART to the post-HAART period, for each of the selected five cancer types.

Calculation of the mean age at cancer diagnosis for each of the five cancer types under study during the pre-HAART and post-HAART periods was affected through use of the matched file of HIV/AIDS and cancer cases. The age, gender, and race/ethnic distributions of HIV-infected persons also diagnosed with cancer for each of the periods, 1991 through 1995 and 1996 through 2002, are shown in Table 7.

There were 1,157 individuals aged 15 or over when diagnosed with cancer in the pre-HAART period and 1,418 in the post-HAART period. The population with both cancer and HIV/AIDS was older at HIV sero-conversion in the post-HAART period and consisted of larger percentages of females and blacks. In order to estimate the year of HIV infection for persons who had been diagnosed with cancer for proper allocation into pre- and post-HAART periods, MCMC multiple imputation was performed on year of HIV seroconversion when the year was not reported. Age at HIV infection was also imputed for cases with missing age values.

Table 7. Age at HIV seroconversion, gender, and race/ethnicity of individuals with cancer diagnosed pre- and post-HAART

Age At HIV/AIDS Infection	Pre-HAART		Post-HAART	
	Cases	Percent	Cases	Percent
15-19	3	0.3	2	0.1
20-29	80	6.9	66	4.6
30-39	510	44.2	386	27.3
40-49	441	38.2	572	40.4
50-59	94	8.2	269	19.0
60+	26	2.3	120	8.5
Gender				
Male	977	84.7	1,041	73.6
Female	177	15.3	374	26.4
Race/Ethnicity				
Hispanic	199	17.2	216	15.3
Black	430	37.3	801	56.6
White	519	45.0	383	27.1
Other/ Unknown	6	0.5	15	1.1

The mean, median, and standard error for the age distribution at diagnosis for each of the cancer types under study are shown in Table 8. There were only 20 cases with missing date of cancer diagnosis. The statistical significance of the difference in mean age at diagnosis for each of the five selected cancer types pre- and post-HAART was assessed through use of a Student's t test for the difference between the means of two populations. A two-sided test of the hypothesis was used in very case. Table 9 provides the mean age at diagnosis for each of the cancers in the two eras and the 95% confidence interval for the difference in the mean ages in the two periods. In addition, the Student's t score resulting from the test of the means and the probability of obtaining a difference that large or greater are given.

Table 8. Median, mean, and standard error of the age at diagnosis for five selected cancer types, pre-HAART and post-HAART

Cancer Type	1991-1995				1996-2002			
	No.	Age at Diagnosis			No.	Age at Diagnosis		
		Median	Mean	Standard Error		Median	Mean	Standard Error
Lung	**81**	**44.0**	**44.3**	**1.0**	**185**	**48.0**	**49.0**	**0.7**
Colorectal	**17**	**40.0**	**41.3**	**1.8**	**48**	**46.5**	**47.3**	**1.5**
Prostate	**9**	**60.0**	**61.0**	**1.8**	**43**	**61.0**	**60.7**	**1.4**
Hodgkin Lymphoma	**18**	**35.0**	**37.7**	**1.9**	**58**	**43.0**	**44.1**	**1.2**
Anal	**19**	**35.0**	**38.6**	**2.5**	**56**	**41.5**	**42.2**	**1.1**

Table 9. Mean age at diagnosis of cancer, pre- and post-HAART, in the HIV- infected population for each of five cancer sites, and significance of change

Cancer Type	Number Of Cases	Mean Age at Diagnosis of Cancer	Post-Pre HAART Mean (95% CI)	Student's t of Difference in Means	P > t score
Lung			**4.7 (2.31,7.09)**	**3.853**	**0.0002**
Pre-HAART	**81**	**44.3**			
Post-HAART	**185**	**49.0**			
Colorectal			**6.0 (1.41,10.59)**	**2.564**	**0.0138**
Pre-HAART	**17**	**41.3**			
Post-HAART	**48**	**47.3**			
Prostate			**-0.3 (-4.77,4.17)**	**-0.131**	**0.8949**
Pre-HAART	**9**	**61.0**			
Post-HAART	**43**	**60.7**			
Hodgkin Lymphoma			**6.4 (1.99,10.81)**	**2.842**	**0.0073**
Pre-HAART	**18**	**37.7**			
Post-HAART	**58**	**44.1**			
Anal			**3.6 (-1.75,8.95)**	**1.319**	**0.1927**
Pre-HAART	**19**	**38.6**			
Post-HAART	**56**	**42.2**			

Except for prostate cancer, the mean ages for each of the five cancer types increased in the post-HAART period compared to pre-HAART years, however only three of the increases reached statistical significance as measured by Student's t test: lung and colorectal cancer and Hodgkin lymphoma. The mean age at diagnosis of anal cancer increased by 3.6 years from the pre- to post-HAART periods, but the resulting t test score failed to reach statistical significance. The mean age for the diagnosis of prostate cancer declined very slightly in the post-HAART period, by 0.3 of a year, however the mean age at prostate cancer diagnosis was substantially higher in this population than that for any other of the selected cancers under study, being above 60 in both time periods.

Objective 3. The third objective is to determine whether the risk of development of AIDS in the population of HIV-infected persons, when adjusted for covariates, is different for those diagnosed with cancer than for those who have not been given a cancer diagnosis.

The hazard of developing AIDS by HIV-infected persons with and without a cancer diagnosis, pre- and post-HAART, was assessed through use of a Cox Proportional Hazards model. The matched file of HIV/AIDS and cancer cases was concatenated with the unmatched file of HIV/AIDS-only cases to form the population of HIV-infected persons at risk of developing AIDS. The combined file of cases formed from HIV/AIDS cases without cancer (59,944 cases) and the file of matched HIV/AIDS-cancer cases (3,085 cases) yielded 63,029 cases.

The dependent variable in the regression analysis was the survival time from HIV infection to development of AIDS, censored by death and the end of follow-up (December 31, 2002). The covariates used in the proportional hazards model $\lambda(t|z(t))=\lambda_o(t)\beta^T z(t)$ where $\lambda_0(t)$ denotes the baseline hazard function at time t, β

is the regression parameter vector and *z(t)* represents a numerically coded vector of possibly time-dependent covariates. In particular, *z(t)* included age at seroconversion to HIV infection, reported to the registry and analyzed as a continuous variable; race, coded 0 if non-Hispanic black and 1 otherwise; and HAART(t), a time-dependent variable equal to 0 if t is in the pre-HAART period and 1 if otherwise. Cancer diagnosis was also treated as a time-dependent covariate equal to at time t and 1 if a cancer diagnosis was received at or before time t. Multiple imputation (PROC MI and PROC MIAnalyze from SAS software, Version 9.1 of the SAS® System for XP_PRO) was applied to handle missing data and the maximum partial likelihood estimates are shown in Table 10.

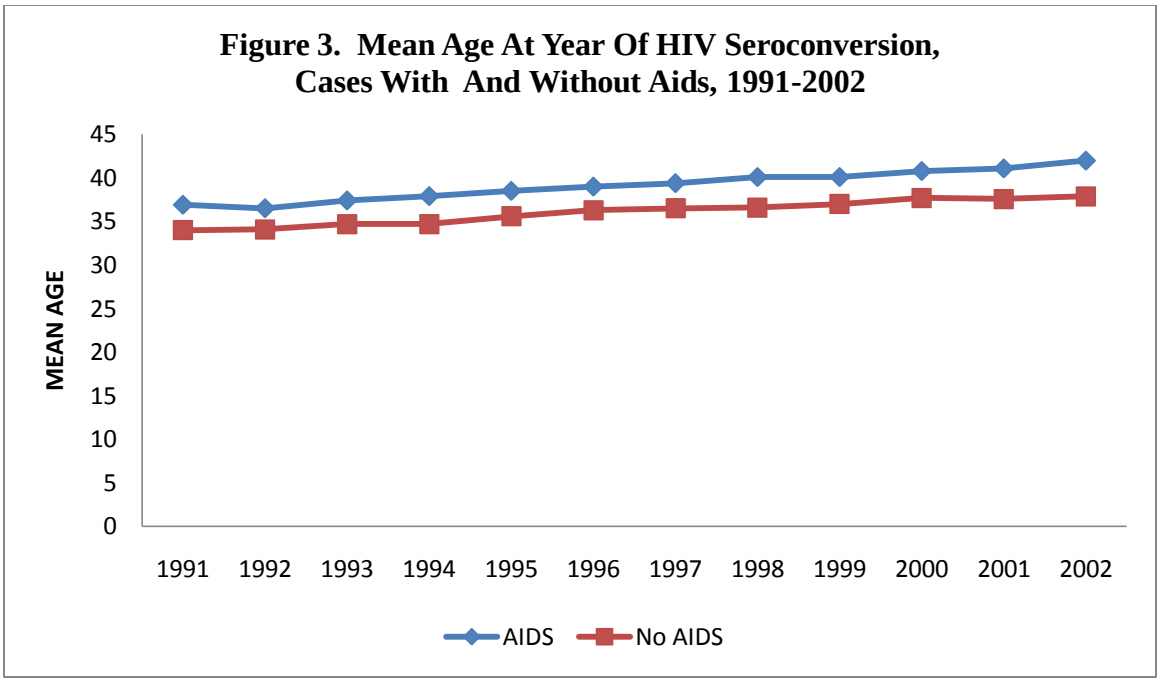
Table 10. Cox proportional hazard regression of covariates associated with development of AIDS in persons with HIV infection

Variable	Coefficient	Standard Error	P Value	Hazard Ratio	Hazard Ratio 95% CI
Cancer (Time Dependent)	**0.73999**	**0.05564**	**<.0001**	**2.096**	**1.879, 2.337**
Age at HIV infection	**-0.00246**	**0.00171**	**0.1491**	**0.998**	**0.994, 1.001**
HAART (Time Dependent)	**-0.15181**	**0.07296**	**.0375**	**0.859**	**0.745, 0.991**
Race	**0.16620**	**0.01637**	**<.0001**	**1.181**	**1.144, 1.219**
Age at HIV Infection*HAART	**0.02769**	**0.00196**	**<.0001**	**1.028**	**1.024, 1.032**

According to the results of the proportional hazards analysis, an HIV-infected person with a cancer diagnosis had a higher risk of AIDS than an HIV-positive individual without cancer, (hazard ratio 2.096 (95%CI:1.879, 2,337). Age at HIV infection had a different effect on the progression to AIDS, indicated

by the significant interaction of age and HAART ($p<0.0001$). In the pre-HAART period, age was not a significant factor in the progression to AIDS (hazards ratio associated with every year increase in age was .998 (95%CI:0.994,1.001); in the post-HAART period the HR associated with age was exp(0.02769-0.00246)≈1.026 (95%CI:0.991,1.062). The effect of HAART itself on progression to AIDS was significant and indicated a lower hazard of AIDS in the post-HAART period than in the pre-HAART period (hazard ratio 0.859, 95%CI: 0.754,0.991). Non-Hispanic black HIV-positive persons had a slightly greater hazard of progression to AIDS than other race/ethnicities (hazard ratio 1.181, 95%CI:1.144,1.219). While a greater percentage of all AIDS cases compared to HIV-infected only cases were black non-Hispanic (58.9% vs. 57.0%), the situation was reversed for Hispanic cases (19.8 % of all cases without AIDS were Hispanic compared to 17.8% of all AIDS cases). White non-Hispanic HIV-infected persons comprised 21.0% of non-AIDS HIV-infected persons and 22.6% of AIDS cases. These figures merit further study: the reporting of race/ethnicity may be problematic or, if these are valid numbers, the reasons for the differences by ethnicity should be investigated.

The interaction of age and year of HIV infection is evident in the data in the following chart. The age at seroconversion increased over both the pre- and post-HAART periods, in both cases that had progressed



to AIDS and those that had not. In every year of the study period, the mean age for those with AIDS exceeded that for the cases that had not developed AIDS by a small but significant margin.

Objective 4. The fourth objective is to evaluate whether the implementation of HAART changes the risk of progressing from HIV infection to cancer diagnosis.

An analysis of whether risk of development of a cancer reportable to the New Jersey Cancer Registry subsequent to infection with HIV changed from the pre-HAART to the post-HAART era was accomplished through application of a Cox proportional hazard model to the data on individuals on the concatenated file of persons reported to HARS, both with and without cancer in the two periods 1991 through 1995 and 1996 through 2002. Data were available for analysis on 1,157 cases of cancer diagnosed from 1991 through 1995 and 1,418 diagnosed from 1996 through 2002 for a total of 2,575 cases. Demographic characteristics of the

HIV-infected population diagnosed with cancer are found in Table 7. Year of infection and age at HIV infection were imputed by the MCMC method of imputation on the cases in the concatenated file if missing.

Survival to cancer was measured by the days from HIV seroconversion to cancer diagnosis. HAART, as in Objective 3, was treated as a time-dependent covariate in the Cox model analysis. Additional available covariates included in the model were age at time of HIV infection (continuous); gender (male equal to 1 and female equal to 0); and race (black non-Hispanic equal to 0, all other race/ethnicities coded as 1). Multiple imputation (PROC MI and PROC MIAnalyze from SAS software, Version 9.1 of the SAS® System for XP_PRO) was applied to handle missing data and the maximum partial likelihood estimates are shown in Table 11.

Although not statistically significant, the risk of development of cancer for individuals following HIV diagnosis in the post-HAART period was only 92 percent of the risk pre-HAART, taking into account the covariates. The hazard of cancer was significantly associated with year of age, increasing 2.3% for each additional year of age, while gender did not have a significant effect on hazard of development of cancer. Race/ethnicity was significantly related to the hazard of development of cancer; races other than black non-Hispanic had an almost 23% greater time to progression to cancer. Black non-Hispanic were 37.4% of the cancer cases in HIV-positive persons pre-HAART and 56.4% post-HAART. White HIV-infected individuals were 44.9% of the cancer cases diagnosed pre-HAART and 27.0% of cases post-HAART, while Hispanic cases were only slightly changed (17.2% pre-HAART and 15.2% post-HAART).

Table 11. Cox proportional hazard regression of covariates associated with time to diagnosis of cancer in HIV-infected persons, pre- and post-HAART

Variable	Coefficient	P Value	Hazard Ratio	Hazard Ratio 95% CI
HAART (Time Dependent)	-0.08389	0.3296	0.920	0.777, 1.088
Age at HIV infection (Continuous)	0.02289	<.0001	1.023	1.015, 1.031
Gender	-0.12998	0.1617	0.878	0.732, 1.053
Race/Ethnicity	0.20595	0.0128	1.229	1.045, 1.445

Chapter VI. Discussion

The data in this study are limited to residents of New Jersey, which differs from the rest of the country in a number of ways. The distribution of mode of transmission of HIV infection is different from that of a number of other states, with a relatively high proportion of IDUs. The socioeconomic and demographic distributions of the New Jersey population are somewhat different also, including a large population (approximately 8.4 million residents), a high level of personal income and a diverse population.

Data on cancer incidence and demographic characteristics in populations infected with HIV are vital to the primary and secondary prevention of these cancers, as well as to the planning of resources to treat the additional numbers of diagnosed cancers expected in this population in the years following HAART implementation (Bower et al., 2006 and Clifford et al., 2005). In a summary of three recent studies of the elevated risk of non-AIDS-related cancer in HIV-infected persons, Liz Highleyman (2009) reported that "people with HIV are more likely than HIV negative individuals to develop various non-AIDS defining cancers…". She noted that the increase is particularly pronounced for malignancies with infectious causes, such as anal cancer due to human papillomavirus and occurs even among people with relatively high CD4 cell counts. J. Cadranel (2006) reported that the sharp drop in AIDS-related mortality in industrialized countries has been accompanied by an increase in the proportion of non-AIDS defining solid tumors, in particular lung cancers. These authors thought that the risk of developing lung cancer appeared to be higher in HIV-infected subjects than in the general population of the same age, in part because HIV-positive persons tend more frequently to be smokers, especially those who are intravenous drug users.

The relationship between HIV infection and some cancers such as cervical cancer and non-Hodgkin lymphoma is well established. The increased risk of these tumors in HIV-infected persons has led to revised screening recommendations for these tumors in HIV-positive individuals. With increased length and quality of life in the HAART era, it is more important than ever to determine if current screening recommendations for other cancers that may affect HIV-infected persons need to be revised. Each of the cancer sites studied showed substantial increases in diagnosed cases in this population; however only prostate cancer exhibited an overwhelmingly increased trend in cumulative incidence rate per 1,000 person-years of observation from pre-HAART to post-HAART periods. Hodgkin lymphoma and anal cancer showed slight or no change in cumulative incidence rates from the pre- to post-HAART eras, while the remaining two cancer types under study, lung and colorectal, had substantial declines in cumulative incidence rates over the period of HAART availability. Although the calculation of cumulative incidence rates for all persons known to be HIV-positive and on the HARS registry by the end of 2002 leads to the conclusion that for the cancer types studied only prostate cancer had an increasing trend in rate, persons who had an HIV infection date after 1995 had substantially higher cumulative incidence rates for each of the cancers than persons with active HIV only in the period before HAART availability.

Implications for Lung Cancer in HIV-Infected Persons

The meta-analysis conducted by Shiels et al. found higher incidence rates of some, but not all, cancers associated with cigarette smoking in HIV-infected persons. When distinction was not made for the availability of HAART, the meta-analysis yielded a SIR for lung cancer of 2.6 (95% CL=2.1, 3.1). This was essentially identical to the pre-HAART SIR for the data available for this dissertation. While the incidence of lung cancer was higher than expected in the pre-HAART era, in the post-HAART period the reported

cases were less than expected if the rates had been identical to those in the general population. A number of researchers have postulated that the higher prevalence of cancer risk factors, in particular, smoking, among the HIV-infected population may explain the increased risk of smoking-related cancers; however some studies have shown an increase in lung cancer even after accounting for difference in smoking levels (Engels et al., 2006 and Cooley, 2003). Shiels et al. (2009) reported finding the SIR for lung cancer to be greater among HIV-infected persons who have progressed to AIDS than those without AIDS. This suggests a role for immune suppression in the development of lung cancer among those who are HIV-infected. This has led some researchers to hypothesize that HIV-associated suppression of the immune system could lead to reduced tumor surveillance, allowing lung tumors to continue to develop when they would otherwise be destroyed by the immune system.

However, according to other researchers (Cooley, 2003; Frisch et al., 2001; Sridhar et al., 1992; & Chan et al., 1993), the incidence of lung cancer does not appear to be related to level of immunosuppression. All of these studies concluded that tobacco use was at least one of the factors related to the elevated trend in lung cancer risk in HIV-infected persons.

The lung cancer incidence rate per 1,000 person-years in HIV-infected persons was substantially higher than that of any of the other cancer types studied. The number of cancers of the lung evidenced the highest cumulative incidence in both pre- and post-HAART periods, but not the highest cumulative incidence rates in the post-HAART period. The decrease in the cumulative lung cancer rate per 1,000 person-years of observation pre- to post-HAART was not due to fewer cases of lung cancer in the HIV- infected population. The number of lung cancer cases actually increased from pre- to post-HAART in the study population, but at the same time the person-years of observation increased sharply due to the increase in survival of HIV-

infected persons. The mean age at lung cancer diagnosis increased significantly from pre- to post-HAART periods at the same time that the total cumulative incidence rate decreased over the two periods. The mean age for lung cancer diagnoses increased from 44.3 to 49.0 from pre- to post-HAART. Although the mean age at lung cancer diagnosis in the HIV-infected population approached fifty years in the post-HAART era, this remains a lower mean age for lung cancer diagnosis than in the general population, as shown by summary of relevant studies by Cooley (2003).

Serraino et al. (2000) attributed the elevated risk of lung cancer in both HIV-positive and HIV-negative injection drug users (IDU) to personal behaviors unrelated to HIV infection. Brock and colleagues (2006) found that HIV-positive lung cancer patients have a shortened survival compared to HIV-negative lung cancer patients, due to advanced stage at diagnosis. As noted, Frisch, Biggar, Engels, and Goedert (2001) surmised that the excess of lung cancer in persons with HIV/AIDS found in their study resulted from heavy smoking. Informed public health and clinical practice leads to the recommendation that physicians and other health care providers take an active role in discouraging smoking in HIV-positive patients. In addition, it is recommended that nonspecific pulmonary infiltrates in these patients should be investigated aggressively.

Implications for Colorectal Cancer in HIV-Positive Persons

There were no colorectal cancers diagnosed in the 60 and over HIV-infected population pre-HAART, but the highest rate post-HAART was in the 60 through 69 year olds. Colorectal cancer in the study population for this dissertation had a non-significantly low SIR in the pre-HAART period and a significantly low SIR in the post-HAART era. The near-null SIRs for colorectal and pancreatic cancer (and a number of other

cancer types) were attributed to having no strong evidence of infectious etiology, therefore the pathogenesis is thought not to be related to HIV-induced immunosuppression.

Bini, Green, and Poles (2009) stressed that screening colonoscopy should be offered to HIV-infected persons, but left open the age of initiation and the frequency of testing. Their recent study showed a higher prevalence of colonic neoplasms developed at a younger age and detected at a more advanced stage in HIV-infected individuals.

While the observed number of colorectal cases was lower than the number expected in both the pre-HAART and post-HAART periods, the mean age at diagnosis of colorectal cancer was significantly higher in the post-HAART period, increasing from 41.3 to 47.3 years. As the survival rate of HIV-infected persons continues to increase, it may become important to suggest that health care providers refer HIV-infected patients with a family history of colorectal cancer or with a history of polyposis coli for colonoscopy at a younger age than that recommended for the general population.

Implications for Prostate Cancer in HIV-Infected Persons

The observed incidence of prostate cancer was less than the expected number of cases in both pre- and post-HAART periods, however the cumulative incidence rate per 1,000 person-years of observation increased pre- to post-HAART. All of the increase in the cumulative incidence rate occurred in the 70 through 79 year age group. Of the five cancer sites under study, only prostate cancer showed no change in mean age at diagnosis over the two periods with the average age at which prostate cancer was diagnosed basically remaining unchanged from the pre-HAART to post-HAART periods. The mean age at diagnosis of prostate

cancer in both the pre- and post-HAART periods was substantially higher than that for any other of the selected cancers under study, being above 60 in both time periods.

Prostate cancer had fewer observed than expected cases in both the pre- and post-HAART periods when compared to the general population of men. Shiels et al. (2009) proposed a possible protective effect of HIV on the development of prostate cancer (as well as on breast cancer), perhaps through changes in hormone levels. Another possible explanation proposed for the low SIR values for prostate cancer in HIV-infected men was the possibility of undiagnosed cases due to differential screening rates of HIV-infected persons because of poverty levels and other factors such as injecting drug use in HIV-infected men.

Clinical guidelines for use of the PSA test as a population-based screening tool vary by organization (Cooper, Merritt, Ross, John, & Jorgensen, 2004). The effectiveness as a screening tool for the entire population of men has not been demonstrated. The American Cancer Society recommends offering the PSA test annually beginning at age 50 to all men who have a life expectancy of 10 years. The American College of Preventive Medicine recommends against routine PSA population screening, and advises that eligible men should be given sufficient information to make informed choices, in consultation with their physicians. The American College of Physicians and the American Urological Association have said that the decision should be individualized, with the provision of information to the patient on the potential benefits and harm of the procedure. The US Preventive Services Task Force states "…the evidence is insufficient to recommend for or against routine screening for prostate cancer using prostate-specific antigen (PSA) testing…" Amid calls for expanded PSA testing in the elderly (Manfredi et al., 2006 & di Gennaro, et al., 2005), other studies recommend the use of American Cancer Society guidelines for men in general (Catalona et al., 2002) or extension to age 40 for men at high risk (Levinson et al., 2005). At the same time,

Crum et al. (2004) predicted that "as the life expectancy of men with HIV infection increases, prostate carcinoma screening will become increasingly important in this population."

Implications for Hodgkin Lymphoma in HIV-Infected Persons

Although the total cumulative incidence of Hodgkin lymphoma in HIV-infected persons changed only slightly from pre- to post-HAART periods, the highest cumulative rate by age moved from the 50 through 59 year age group in the pre-HAART period to the next higher group, 60 through 69 years in the post-HAART group. There were no reported diagnoses of Hodgkin lymphoma in HIV-positive persons older than 59 years in the pre-HAART era or above 69 years in the post-HAART period.

The observed incidence of Hodgkin lymphoma was greater than the number of cases expected at the general population rates in both the pre- and post-HAART periods. The SIRs for both periods were significant and the ratio of SIRs was relatively high but not significant. This indicates that the observed incidence, which was higher than expected pre-HAART, was also greater than expected post-HAART. At the same time, the age at diagnosis was significantly higher in the post-HAART period, compared to the earlier era, with the average age at diagnosis in the post-HAART period increasing from 37.7 years to 44.1 years ($p=0.0073$). Hodgkin disease is one of the malignancies that Goedert et al. (1998) suggested might be affected by immunological failure to control herpes (EBV), in addition to the contribution from behavioral risk factors. Frisch et al. (2001) found that Hodgkin disease, most notably of the mixed cellularity and lymphocytic depletion subtypes, along with a few other malignancies, "…may be genuinely influenced by immunosuppression.", adding that Hodgkin lymphoma occurs most often at only moderate levels of immunosuppression. This finding may be related to the increase in diagnosis of Hodgkin lymphoma in the

time period since HAART has been available. A number of experts have called for designating Hodgkin lymphoma as an AIDS-defining disease (Serraino et al., 1997; Grulich, Wan, Law, Coates, & Kaldor, 1999; & Clarke & Glaser, 2001).

Implications for Anal Cancer in HIV-Infected Persons

Cumulative incidence rates for anal cancer were higher in the pre-HAART era than in the post-HAART period for all ages except those 40 through 49 years. There were no reported anal cancer diagnoses in the 70 through 79 age group in either period. Despite this, anal cancer had significantly higher observed than expected incidence for each of the two periods, pre- and post-HAART. In addition, the change in mean age was not significant over the time span (from 38.6 to 42.2 years). Goldie et al. (1999) studied the clinical effectiveness and cost-effectiveness of screening for anal squamous intraepithelial lesions in homosexual and bisexual HIV-positive men. The authors' results led them to conclude that "screening homosexual and bisexual HIV-positive men for anal squamous intraepithelial lesions (ASIL), at all stages of HIV, would prolong quality-adjusted life expectancy". Pap screening every two years early in HIV disease was found to be robust over a wide range of sensitivity analyses, while annual screening was recommended when initiation of screening occurred with later HIV disease. The authors concede that their estimates are preliminary due to the fact that the long-term impact of HAART is still unknown. Despite the unknown factors, the authors conclude that depending on the length of time of HIV disease, the cost-effectiveness of either a yearly or every-2-year screening schedule is comparable with other accepted preventive measures in clinical medicine and therefore is a recommended preventative strategy for homosexual and bisexual HIV-infected persons.

Relative Risk

When relative risk was approximated through calculation of ratios of the Standardized Incidence Rates for the post-HAART period compared to pre-HAART SIRs, the results were remarkably consistent. All of the five cancers other than Hodgkin disease had SIR ratios that were significantly less than 1. This is an indication that for the population of HIV-infected persons encompassed by this study, the relative risk for these four cancers was significantly lower, post-HAART compared to pre-HAART. The unchanged relative risk for Hodgkin disease from the post- to pre-HAART periods results from a continued high level of observed to expected cases.

Implications for Progression to AIDS

When the hazard of HIV-infected persons developing AIDS pre- and post-HAART was analyzed through use of a proportional hazards model incorporating covariates, the results indicated that an HIV-infected person with a cancer diagnosis had more than twice the hazard of development of AIDS as a similarly infected person without cancer. The results indicated a lower hazard of AIDS in the post-HAART period than in the pre-HAART period. Race and the interaction of age at HIV infection and HAART also were significantly associated with higher risk of progressing to AIDS with positive HIV status and older age. These results lead to recommendations to aggressively monitor opportunistic infections and provide routine assessments of CD4 counts and viral loads as an integral part of therapy for the treatment of HIV-infected persons.

Implications for the Diagnosis of Cancer Pre- and Post-HAART

For this study population, the risk of development of cancer for HIV-infected persons did not change significantly from the pre-HAART to post-HAART period. However, the hazard ratio was significantly associated with age while races other than black non-Hispanic had a greater hazard of progression to cancer than all other race/ethnicities.

Study Limitations

There are limitations to the completeness and coverage of this study. In addition to the estimation and imputation methodologies covered in the Statistical Analysis section above which were necessitated by missing data, there are additional factors inherent in the data and the subsequent analysis that could have an impact on the findings. While imputation is a commonly used technique for estimation of missing data in epidemiological studies, issues of bias resulting from its use should not be ignored. In predictive models, in particular, using imputed values on relatively large amounts of missing data has been shown to result in considerable bias either toward or away from the null, depending on the missingness pattern (Gorelick, 2006). However Newgard (2006) and others have shown that multiple imputation can increase precision and reduce bias compared to complete case analysis, particularly when the percentage of missingness remains moderate, as in this study.

The two important variables with relatively large missing percentages which were critical to these analyses were year of HIV seroconversion and age at HIV seroconversion. These values were missing for the most part when the registry recorded only AIDS cases. Since its expansion to an HIV/AIDS registry, these two

variables are essentially fully reported. Thus missingness for these two variables will affect a smaller percentage of the total cases in the future and therefore be less of a problem to plans for future analysis.

A major lack in the data available for analysis was that cause of death was not provided in the study files. In order to analyze competing risks, the cause of failure (death, in these analyses) could not be properly allocated to the event of interest (death from AIDS or cancer) in the presence of competing risks. The analysis of competing risks allows assessment of the effects of prognostic factors on the AIDS-specific hazard. The competing risks in these analyses arise because a death in these cohorts may have AIDS as the underlying cause or it may result from any other cause. Neither the New Jersey AIDS/Cancer Match File (NJACM) cases nor the unmatched file of HIV/AIDS cases (HARS) contains the underlying cause of death, thus the competing risks could not be taken into account. This may have resulted in biased results and invalid inferences, although data from the CASCADE Collaboration (2002) indicate the percentage bias in this type of study may not be large. Multiple cause of death information is collected for each person registered on HARS who dies. This information would have enabled analysis of competing risks for this study. However, cause of death information for an individual is considered confidential, and permission to use the data would have had to be requested of the Institutional Review Boards of both the Department of Health and Senior Services and the University of Medicine and Dentistry of New Jersey. As noted, these data were not present on the files available for research. Time requirements would have made this process difficult and staff to abstract and provide the additional data were not available. Future projects should consider incorporating the analysis of competing risks through utilization of existing cause of death data.

The individual effects of HAART could not be ascertained from the files, and could only be approximated by allocation of cases into pre-HAART (1991-1995) and post-HAART (1996-2002) periods in which the

diagnoses occurred. There is no indication of which individuals had been exposed to HAART or had received other therapy for HIV infection.

CD4 counts were not sufficiently complete to allow an analysis of the effect of immunosuppression on the findings. Only 56% of the records of HIV-infected persons who had not been diagnosed with cancer had a figure recorded in the file as the most recent CD4 reading and CD4 readings were completely absent from the matched file of HIV/AIDS and cancer cases.

The HARS registry file provides no information on current smoking or smoking history, which would have been helpful for analysis of effect of behavior on cancer incidence in HIV-infected persons.

The racial/ethnic codes reported on HARS compared to those on the New Jersey Cancer Registry may differ, however resources were not available to quantify any differences.

Chapter VII. Conclusions

These data show that among the five cancer types under study, total cumulative incidence rates restricted to HIV-infected persons alive and without a cancer diagnosis for any part of the pre-HAART period were lower than for persons with an HIV infection date in the post-HAART period. This finding may have important future implications for the effects of HAART and the increasing age of HIV-infected populations on projected trends in the incidence of these five cancer types in HIV-infected persons.

The change in cumulative incidence by age group from pre-HAART to post-HAART for all persons with person-years of observation in this study (those with seroconversion dates before 2002 and alive and without cancer by 1991) varied by cancer type and age. The most consistent change from the pre-HAART to post-HAART eras was in a shift toward the highest age-specific rate in the oldest age group. This trend was evident in colorectal cancer, prostate cancer, and Hodgkin lymphoma, although the Hodgkin lymphoma incidence rate for 40-49 year olds also increased pre- to post-HAART. At the same time, lung cancer age-specific incidence rate declined in every age group and anal cancer rates increased only for those in the 40 through 49 year age group. Despite the differences in trend by age group, the mean age at diagnosis for each of the five cancer types was higher in the post-HAART period, except for prostate cancer with a virtually unchanged mean age.

The etiology of different cancers is known to vary. For example, some cancers are associated with infections (e.g., anal cancer, Hodgkin lymphoma, cervical cancer, liver cancer), while other cancers are related to environmental exposures (i.e., lung cancer and mesothelioma). The differences observed in the development of the five cancers studied may be related to the etiology of the cancer. The variations may also be related to differences in pathophysiology and organs of origin of the different tumors. Therefore, in some instances HIV disease may increase the risk of cancer development or increase the risk of development at an earlier age whereas for other cancers HIV status may be unrelated to tumor development. The five cancers under study provide striking examples of the effects of the differences in etiology of cancer types. The increase in prostate cancer cumulative incidence rates post-HAART and the high age-specific rates in older men is undoubtedly related to the increased survival of HIV-infected men in the period after 1996. Anal cancer and Hodgkin lymphoma are associated with infectious agents which may explain the increases in incidence rates at younger ages despite the large increase in person-years of observation. Lung

and colorectal cancers indicated a decrease in cumulative incidence rates, which was not unexpected due to the current lack of known links to infectious causes.

For all cancers in the HIV-infected population, persons with a cancer diagnosis had a greatly increased risk of progressing to AIDS, although the risk decreased in the post-HAART period. The hazard ratio was not significantly different for the development of cancer pre- and post-HAART, although this analysis should be repeated when data from additional years of HAART treatment are available. A number of cancer types may have more extended incubation periods than available through use of data only from the seven-year study period of HAART availability, 1996 through 2002.

The major advantage of this study of the available files of New Jersey residents is that the records of all HIV-infected persons ever reported to HARS were available, encompassing both those with and without a cancer diagnosis. This broadened the research objectives that could be accomplished beyond only persons with an AIDS diagnosis to encompass reporting of individuals who have tested positive for HIV infection and, in addition, allowed the investigation of factors associated with the diagnosis of cancer in all HIV-infected persons known to the state registry.

Due to the access to records of HIV-infected persons who had not been diagnosed with any reportable cancer, the person-years of observation are much larger than those reported by other studies using only the matched cases of HIV/AIDS/Cancer. The findings from this study may differ from other research utilizing only matched files of HIV/AIDS and cancer cases. The critical difference is the population to which the findings relate. In this case, generalizations can be made to the entire state's population of HIV-infected persons aged 15 years and over and not just those who have had a cancer diagnosis or progressed to AIDS.

References

Allison, P.D.,(2005). Survival Analysis Using SAS A Practical Guide. SAS Press. SAS Institute Inc., Cary, NC, USA.

American Cancer Society. (2007a). Global Cancer Facts & Figures [Electronic Version], fromhttp://www.cancer.org/downloads/STT/Global_Cancer_Facts_and_Figures_2007_rev.pdf

American Cancer Society. (2007b). Cancer Facts & Figures 2007 [Electronic Version], from http://www.cancer.org/downloads/STT/CAFF2007PWSecured.pdf

American Cancer Society. (2007c). Lung Cancer [Electronic Version], from http://www.cancer.org/downloads/PRO/LungCancer.pdf

American Cancer Society. (2007d). Overview: Colon and Rectum Cancer [Electronic Version], from http://www.cancer.org/docroot/CRI/content/CRI_2_2_1X_How_Many_People_Get_Colorectal_Cancer.asp?sitearea=

American Cancer Society. (2007e). Cancer Reference Information Detailed Guide: Prostate Cancer [Electronic Version], from http://www.cancer.org/docroot/CRI/content/CRI_2_4_1X_What_are_the_key_statistics_for_prostate_cancer_36.asp?sitearea

American Cancer Society. (2007f). Cancer Reference Information Detailed Guide: Hodgkin Disease [Electronic Version], from http://www.cancer.org/docroot/CRI/content/CRI_2_4_1x_What_Is_Hodgkin_Disease.asp?rnav=cri

Barbaro, G., & Barbarini, G. (2007). HIV infection and cancer in the era of highly active antiretroviral therapy (Review).

Barzi, F., & Woodward, M. (2004). Imputations of missing values in practice: Results from imputations of serum cholesterol in 28 cohort studies. *Am J Epidemiol, 160*, 34-45.

Bedimo, R., Chen, R. Y., Accortt, N. A., Raper, J. L., Linn, C., Allison, J. J., et al. (2004). Trends in AIDS-Defining and Non-AIDS-Defining Malignancies among HIV-infected patients:1989-2002. *CID, 39*, 1380-1384.

Biggar, R. J., Burnett, W., Miki, J., & Nasca, P. (1989). Cancer among New York men at risk of acquired immunodeficiency syndrome. *International Journal of Cancer, 43*(6), 979-985.

Biggar, R. J., Kirby, K. A., Atkinson, J., McNeel, T. S., & Engels, E. A. (2004). Cancer risk in elderly persons with HIV/AIDS. *J Acquir Immune Defic Syndr, 36*, 861-868.

Biggar, R. J., Engels, E. A., Ly, S., Kahn, A., Schymura, M. J., Sackoff, J., et al. (2005). Survival after cancer diagnosis in persons with AIDS. *J Acquir Immune Defic Syndr, 39*(3), 293-299.

Bini, E.J., Green, B., & Poles, M.A. (2009). Screening colonoscopy for the detection of neoplastic lesions in asymptomatic HIV-infected subjects. *Gut, 58*, 1129-1134.

Bonnet, F., Lewden, C., May, T., Heripret, L., Jougla, E., Bevilacqua, S., et al. (2004). Malignancy-related causes of death in human immunodeficiency virus-infected patients in the era of highly active antiretroviral therapy *Cancer, 101*, 317-324.

Bower, M., Powles, T., Newsom-Davis, T., Thirlwell, C., Stebbing, J., Mandalia, S., et al. (2004). HIV-associated anal cancer Has Highly Active Antiretroviral Therapy Reduced the Incidence or Improved the Outcome? *J Acquir Immune Defic Syndr, 37*(5), 1563-1565.

Bower, M., Palmieri, C., & Dhillon, T. (2006). AIDS-related malignancies: changing epidemiology and the impact of highly active antiretroviral therapy. *Curr Opin Infect Dis, 19*(1), 14-19.

Breslow, N. E., & Day, N. E. (1987). Statistical Methods in Cancer Research (Vol. II - The Design and Analysis of Cohort Studies). Lyon: International Agency for Research on Cancer.

Brock, M. V., Hooker, C. M., Engels, E. A., Moore, R. D., Gillison, M., Alberg, A. J., et al. (2006). Delayed diagnosis and elevated mortality in an urban population with HIV and lung cancer: implications for patient care. *J Acquir Immune Defic Syndr, 43*(1), 47-55.
Brookmeyer, R. (1990). Statistical problems in epidemiologic studies of the natural history of disease. *Environmental Health Perspectives, 87*, 43-49.

Burgi, A., Brodine, S., Wegner, S., Mailazzo, M., Wallace, M. R., Spooner, K., et al. (2005). Incidence and risk factors for the occurrence of non-AIDS-defining cancers among human immunodeficiency virus-infected individuals. *Cancer, 104*, 1505-1511.

Cadranel, J., Garfield, D., Lavole, A., Wislez, M., Milleron,B., Mayaud,C. (2006). Lung cancer in HIV infected patients: facts, questions, and challenges. *Thorax*,61:1000-1008.

Cameron, J. E., & Hagensee, M. E. (2007). Human papillomavirus infection and disease in the HIV+ individual. *Cancer Treat Res, 133*, 185-213.

Cancer Epidemiology Services, New Jersey Department of Health and Senior Services (2005). *Trends in Cancer Incidence and Mortality in New Jersey 1979-2002* from http://www.nj.gov/health/ces/documents/cancer_incidence79_02.pdf.

CASCADE Collaboration (2002). Changes over calendar time in the risk of specific first AIDS-defining events following HIV seroconversion, adjusting for competing risks. *International Journal of Epidemiology*, 31:951-958.

Catalona, W. J., Antenor, J. A. V., & Roehl, K. A. (2002). Screening for prostate cancer in high risk populations. *The Journal of Urology, 168*, 1980-1984.

Centers for Disease Control and Prevention. (1992). *1993 Revised Classification System for HIV Infection and Expanded Surveillance Case Definition for AIDS among Adolescents and Adults*, from http://www.cdc.gov/mmwr/preview/mmwrhtml/00018871.htm.

Centers for Disease Control and Prevention MMWR Weekly. (2005). Trends in HIV/AIDS Diagnosis – 33 States, from http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5445a1.htm

Centers for Disease Control and Prevention. (2009). HIV Incidence| Statistics and Surveillance| Topics| CDC HIV/AIDS, from http://www.cdc.gov/hiv/topics/surveillance/incidence.htm

Chan, T. K., Aranda, C. P., & Rom, W. N. (1993). Bronchogenic carcinoma in young patients at risk for Acquired Immunodeficiency Syndrome. *Chest, 103*(3), 862-864.

Chiao, E. Y., & Krown, S. E. (2003). Update on non-acquired immunodeficiency syndrome-defining malignancies. *Curr Opin Oncol, 15*, 389-397.

Chiao, E. Y., Krown, S. E., Stier, E. A., & Schrag, D. (2005). A Population-Based Analysis of Temporal Trends in the Incidence of Squamous Anal Canal Cancer in Relation to the HIV Epidemic *Journal of Acquired Immune Deficiency Syndrome, 40*(4), 451-455.

Clarke, C. A., & Glaser, S. L. (2001). Epidemiologic trends in HIV-associated lymphomas. *Current Opinion in Oncology, 13*, 354-359.

Clifford, G. M., Polesel, J., Rickenbach, M., Del Maso, L., Keiser, O., Kofler, A., et al. (2005). Cancer risk in the Swiss HIV cohort study: associations with immunodeficiency, smoking, and highly active antiretroviral therapy. *J Natl Cancer Inst 97*, 425-432.

Clifford, G. M., & Franceschi Silvia. (2007). Immunity, infection, and cancer. *The Lancet, 370*(9581), 6-7.

Coffield, A. B., Maciosek, M.V., McGinnis, J.M., Harris, J.R., Caldwell, M.B., Teutsch, S.M., Atkins, D., Richland, J.H., & Haddix, A.,. (2001). Priorities among recommended clinical preventive services. *Am J Prev Med, 21*, 1-9.

Cooley, T. P. (2003). Non-AIDS-defining cancer in HIV-infected people. *Hematol Oncol Clin N Am, 17*, 889-899.

Cooper, C. P., Merritt, T. L., Ross, L. E., John, L. V., & Jorgensen, C. M. (2004). To screen or not to screen, when clinical guidelines disagree: primary care physicians' use of the PSA test. *Preventive Medicine, 38*, 182-191.

Crum, N. F., Spencer, C. R., & Amling, C. L. (2004). Prostate carcinoma among men with human immunodeficiency virus infection. *Cancer, 101*(2), 294-299.

Crum-Cianflone, N., Hullsiek, K. H., Marconi, V., Weintrob, A., Ganesan, A., Barthel, R. V., et al. (2009). Trends in the incidence of cancers among HIV-infected persons and the impact of antiretroviral therapy; a 20-year cohort study. *AIDS, 23*(1), 41-50.

di Gennaro, G., Cinelli, R., Vaccher, E., Spina, M., & Tirelli, U. (2005). Cancer prevention and early diagnosis in HIV-positive individuals. *J Acquir Immune Defic Syndr, 38*(5), 628-629.

Diamond, C., Taylor, T. H., Aboumrad, T., Bringman, D., & Anton-Culver, H. (2005). Increased incidence of squamous cell anal cancer among men with AIDS in the era of highly active antiretroviral therapy. *Sexually Transmitted Diseases, 32*(5), 314-320.
Division of Cancer Prevention and Control. (2004). Cancer Mortality Surveillance---United States, 1990-2000 [Electronic Version]. *MMWR Surveillance Summaries, 53*, 1-108, from http://www.cdc.gov/mmwr/preview/mmwrhtml/ss5303a1.htm

Engels, E. A. (2001). Human immunodeficiency virus infection, aging, and cancer. *Journal of Clinical Epidemiology, 54*, S29-S34.

Engels, E. A., & Goedert, J. J. (2005). Human immunodeficiency virus/acquired immunodeficiency syndrome and cancer: past, present, and future. *J Natl Cancer Inst, 97*, 407-409.
Engels, E. A., Brock, M. V., Chen, J., Hooker, C. M., Gillison, M., & Moore, R. D. (2006a). Elevated incidence of lung cancer among HIV-infected individuals. *Journal of Clinical Oncology, 24*(9), 1383-1388.

Engels, E. A., Pfeiffer, R. M., Goedert, J. J., Virgo, P., McNeel, T. S., Scoppa, S. M., et al. (2006b). Trends in cancer risk among people with AIDS in the United States 1980-2002. *AIDS, 20*(12), 1645-1654.

Engels, E. A., Biggar, R. J., Hall, H. I., Cross, H., Crutchfield, A., Finch, J. L., et al. (2008). Cancer risk in people infected with human immunodeficiency virus in the United States. *Int J Cancer, 123*, 187-194.

Frisch, M., Olsen, J. H., Bautz, A., & Melbye, M. (1994). Benign Anal Lesions and the Risk of Anal Cancer. *The New England Journal of Medicine, 331*(5), 300-306.

Frisch, M., Biggar, R. J., & Goedert, J. J. (2000a). Human papillomavirus-associated cancers in patients with human immunodeficiency virus infection and acquired immunodeficiency syndrome. *J Natl Cancer Inst, 92*, 1500-1510.

Frisch, M., & Goodman, M. T. (2000b). Human papillomavirus-associated carcinomas in Hawaii and the mainland U.S. *Cancer, 88*(6), 1464-1469.

Frisch, M., Biggar, R. J., Engels, E. A., & Goedert, J. J. (2001). Association of Cancer with AIDS-Related Immunosuppression in Adults. *JAMA, 285*(13), 1736-1745.

Gallagher B., Wang Z., Schymura M. J., Kahn A., & Fordyce E. J. (2001). Cancer incidence in New York State Acquired Immunodeficiency Syndrome patients. *American Journal of Epidemiology, 154*(6), 544-556.

Garson, G. D. (2008). Data Imputation for Missing Values [Electronic Version]. *Statnotes: Topics in Multivariate Analysis*, from http://faculty.chass.ncsu.edu/garson/pa765/statnote.htm

Ghadirian P., Howe G.R., Hislop T.G., & Maisonneuve, P. (1997). Family history of prostate cancer: A multi-center-case-control study in Canada *Int J Cancer, 70*, 679-681.

Glaser, S. L., Clarke, C. A., Gulley, M. L., Craig, F. E., DiGiuseppe, J. A., Dorfman, R. F., et al. (2003). Population-based patterns of Human Immunodeficiency Virus-related Hodgkin lymphoma in the Greater San Francisco Bay Area, 1988-1998. *Cancer, 98*(2), 300-309.

Glaser, S. L., Clarke, C. A., Gomez, S. L., O'Malley, C. D., Purdie, D. M., & West, D. W. (2005). Cancer surveillance research: a vital subdiscipline of cancer epidemiology. Cancer Causes and Control, 16, 1009-1019.

Goedert, J. J., Cote, T. R., Virgo, P., Scoppa, S. M., Kingma, D. W., Gail, M. H., et al. (1998). Spectrum of AIDS-associated malignant disorders. *The Lancet, 351*(9119), 1833-1839.

Goldie, S. J., Kuntz, K. M., Weinstein, M. C., Freedberg, K. A., Welton, M. L., & Palefsky, J. M. (1999). The Clinical Effectiveness and Cost-effectiveness of Screening for Anal Squamous Intraepithelial Lesions in Homosexual and Bisexual HIV-Positive Men. *Journal of the American Medical Association, 281*(19), 1822-1829.

Gorelick, M.H. (2006). Bias arising from missing data in predictive models. *J Clin Epidemiol*. 59(10); 1115-1123.

Grulich, A., Wan, X., Law, M., Coates, M., & Kaldor, J. (1999). Risk of cancer in people with AIDS. *AIDS, 13*, 839-843.

Grulich, A. E., Li, Y., McDonald, A., Correll, P. K. L., Law, M. G., & Kaldor, J. M. (2002). Rates of non-AIDS-defining cancers in people with HIV infection before and after AIDS diagnosis. *AIDS, 16*, 1155-1161.

Hakimian, R., Fang, H., Thomas, L., & Edelman, M. J. (2007). Lung cancer in HIV-infected patients in the era of highly active antiretroviral therapy. *Journal of Thoracic Oncology, 2*(4), 268-272.

Hendriks, J., Satten, G., van Ameijden, E., van Druten, H., Coutinho, R., & van Griensven, G. (1998). The incubation period to AIDS in injecting drug users estimated from prevalent cohort data, accounting for death prior to an AIDS diagnosis *AIDS, 12*(12), 1537-1544.
Herida, M., Mary-Krause, M., Kaphan, R., Cadranel, J., Poizot-Martin, I., Rabaud, C., et al. (2003). Incidence of non-AIDS-defining cancers before and during the highly active antiretroviral therapy era in a cohort of human immunodeficiency virus-infected patients. *J Clin Oncol, 18*, 3447-3453.

Hessol, N. A., & Palacio, H. (1996). Gender, ethnicity and transmission category variations in HIV disease progression. *AIDS, 10 (suppl A)*, S69-S74.

Hessol, N. A., Pipkin, S., Schwarcz, S., Cress, R. D., Bacchetti, P., & Scheer, S. (2007). The impact of highly active antiretroviral therapy on non-AIDS-defining cancers among adults with AIDS. *Am J Epidemiol, 165*, 1143-1153.
Highleyman,L. (2009). Three studies look at elevated risk of non-AIDS-related cancer in people with HIV. HIV and Hepatitis.com. from http://www.hivandhepatitis.com/recent/2009/120809_g.html

Hoffmann, C., Chow, K. U., Wolf, E., Faetkenheuer, G., Stellbrink, H.-J., Lunzen, J. v., et al. (2004). Strong impact of highly active antiretroviral therapy on survival in patients with human immunodeficiency virus-associated Hodgkin's disease. *British Journal of Haematology, 125*, 455-462.

Hogg, R. S., O'Shaughnessy, M. V., Gataric, N., Yip, B., Craib, K., Schecter, M. T., et al. (1997). Decline in deaths from AIDS due to new antiretrovirals. *The Lancet, 349*, 1294.

Hogg, R. S., Heath, K. V., Yip, B., Craib, K. J. P., O'Shaughnessy, M. V., Schecter, M. T., et al. (1998). Improved survival among HIV-infected individuals following initiation of antiretroviral therapy. *JAMA, 279*(6), 450-454.

Holly, E. A., Whittemore, A. S., Aston, D. A., Ahn, D. K., Nickoloff, B. J., & Kristiansen, J. J. (1989). Anal Cancer Incidence: Genital Warts, Anal Fissure or Fistula, Hemorrhoids, and Smoking *Journal of the National Cancer Institute, 81*(22), 1726-1731.

Huynh, D., Vincan, E., Mantamadiotis, T., Purcell, D., Chee-Kai, C., & Ramsay, R. (2007). Oncogenic Properties of HIV-Tat in colorectal cancer cells. *Current HIV Research, 5*(4), 403-409.

International Collaboration on HIV and Cancer. (2000). Highly active antiretroviral therapy and incidence of cancer in human immunodeficiency virus-infected adults. *J Natl Cancer Inst 92*, 1823-1830.

Jemal, A., Ward, E., Hao, Y., & Thun, M. (2005). Trends in the leading causes of death in the United States, 1970-2002. *JAMA, 294*(10), 1255-1259.

Johnson, L. G., Madeleine, M. M., Newcomer, L. M., Schwartz, S. M., & Daling, J. R. (2004). Anal Cancer Incidence and Survival: the Surveillance, Epidemiology, and End Results Experience, 1973-2000. *Cancer, 101*(2), 281-288.

Kinlen, L. (2004). Infections and immune factors in cancer: the role of epidemiology. *Oncogene, 23*(38), 6341-6348.

Kirk, G. D., Merlo, C., O'Driscoll P., Mehta, S. H., Galai, N., Vishov, D., et al. (2007). HIV infection is associated with an increased risk for lung cancer, independent of smoking. *Clin Infect Dis, 45*(1), 103-110.

Lai, D., & Hardy, R. J. (2004). An update on the impact of HIV/AIDS on life expectancy in the United States. *AIDS, 18*(12), 1732-1734.

Lederman, M. M., & Valdez, H. (2000). Immune restoration with antiretroviral therapies Implications for clinical management. *JAMA, 284*(2), 223-227.

Levinson, A., Nagler, E. A., & Lowe, F. C. (2005). Approach to management of clinically localized prostate cancer in patients with human immunodeficiency virus. *Urology, 65*, 91-94.

Lichtenstein, P., Holm, N. V., Verkasalo, P. K., Iuadou, A., Kaprio, J., Koskenvuo, M., et al. (2000). Environmental and heritable factors in the causation of cancer. *The New England Journal of Medicine, 343*(2), 78-85.

Lim, S. T., & Levine, A. M. (2005). Non-AIDS-defining cancers and HIV infection. *Curr HIV/AIDS Rep., 2*(3), 146-153.

Lloyd-Smith, E., Brodkin, E., Wood, E., Kerr, T., Tyndall, M. W., Montaner, J. S. G., et al. (2006). Impact of HAART and injection drug use on life expectancy of two HIV-positive cohorts in British Columbia. *AIDS, 20*(3), 445-450.

Lu, K. & Tsiatis, A.A. (2005). Comparison between two partial likelihood approaches for the competing risks model with missing cause of failure. *Lifetime Data Analysis*. 11, 29-40.

Manfredi, R., Fulgaro, C., Sabbatani, S., Dentale, N., & Legnani, G. (2006). Disseminated, lethal prostate cancer during human immunodeficiency virus infection presenting with non-specific features. Open questions for urologists, oncologists, and infectious disease specialists. *Cancer Detection and Prevention, 30*, 20-23.

Mocroft, A., Vella, S., Benfield, T. L., Chiesi, A., Miller, V., Gargalianos, P., et al. (1998). Changing patterns of mortality across Europe in patients infected with HIV-1. *The Lancet, 352*, 1725-1730.
Munoz, A., & Xu, J. (1996). Models for the incubation of AIDS and variations according to age and period. *Statistics in Medicine, 15*, 2459-2473.
National AIDS/Cancer Match Registry (NACMR) Protocol (2004).

Nelson, W. G., De Marzo, A. M., & Isaacs, W. B. (2003). Mechanisms of disease: Prostate cancer. *The New England Journal of Medicine, 349*, 366-381.

New Jersey Administrative Code Title 8 Department of Health and Senior Services Chapter 57. (2003). Communicable Diseases Proposed: September 25, 2003 Expiration Date: September 25, 2008. Subchapter 2. Reporting of Acquired Immunodeficiency Syndrome and Infection with Human Immunodeficiency Virus [Electronic Version], from www.state.nj.us/health/cd/chapter57.pdf

New Jersey Department of Health and Senior Services. (2007a). New Jersey Cancer Registry [Electronic Version], from http://www.state.nj.us/health/ces/njscr.shtml

New Jersey Department of Health and Senior Services State Cancer Registry. (2004). Cancer Incidence and Mortality in New Jersey, 1998-2002 [Electronic Version]. Retrieved June 29, 2009, from www.state.nj.us/health/cancer/documents/reports98-02.pdf

New Jersey Department of Health and Senior Services Division of HIV/AIDS Services. (2007b). HIV/AIDS Epidemiology Profile for the State of New Jersey 2005 [Electronic Version], from http://www.state.nj.us/health/aids/documents/epi2005.pdf

Newgard, C. D. (2006). The validity of using multiple imputation for missing out-of-hospital data in a state trauma registry. *Acad Emerg Med*.. 13(3);314-324.

Palefsky, J. M., Holly, E. A., Hogeboom, C., Ralston, M. L., DaCosta, M. M., Botts, R., et al. (1998). Virologic, Immunologic, and Clinical Parameters in the Incidence and Progression of Anal Squamous Intraepithelial Lesions in HIV-Positive and HIV-Negative Homosexual Men. *Journal of Acquired Immune Deficiency Syndrome, 17*(4), 314-319.

Palella, F. J., Delaney, K. M., Moorman, A. C., Loveless, M. O., Fuhrer, J., Satten, G. A., et al. (1998). Declining Morbidity and Mortality among Patients with Advanced Human Immunodeficiency Virus Infection. *The New England Journal of Medicine, 338*(13), 853-860.

Patel, P., Hanson, D. L., Sullivan, P. S., Novak, R. M., Moorman, A. C., Tong, T. C., et al. (2008). Incidence of types of cancer among HIV-infected persons compared with the general population in the United States, 1992-2003. *Annals of Internal Medicine, 148*(10), 728-736.

Porter, K., Johnson, A. M., Phillips, A. N., & Darbyshire, J. H. (1999). The practical significance of potential biases in estimates of the AIDS incubation period distribution in the UK register of HIV seroconverters. *AIDS, 13*(14), 1943-1951.

Raghunathan, T.E., Solenberger, P.W. & Hoewyk, J.V. (2002). IVEware: Imputation and Variance Estimation Software User Guide. [Electronic Version], from http://www.isr.umich.edu/src/smp/ive.

Rosner, B. (2006). *Fundamentals of Biostatistics* (5th ed.): Cengage Learning.

Rothman, K. J., & Boice, J. D., Jr. (1979). Epidemiologic Analysis with a Programmable Calculator. US Government Printing Office, NIH Publication 79-1649.Ryan, D. P., Compton, C. C., & Mayer, R. J. (2000). Carcinoma of the Anal Canal. *The New England Journal of Medicine, 342*(11), 792-800.

Ryan, D. P., Compton, C. C., & Mayer, R. J. (2000). Carcinoma of the Anal Canal. *The New England Journal of Medicine, 342*(11), 792-800.

State of New Jersey, Department of Health and Senior Services. (2009). *New Jersey HIV/AIDS Semi-annual Newsletter*, from http://www.state.nj.us/health/aids/aidsqtr.shtml

Surveillance Epidemiology and End Results Program. (2008). Seer Limited Use File Documentation [Electronic Version], from http://seer.cancer.gov/manuals/CD2.SEERDic.pdf

Surveillance Epidemiology and End Results. (2009). Browse the SEER Cancer Statistics Review 1975-2006, from http://www.seer.cancer.gov/csr/1975_2006/browse_csr.php?section=10&page=sect_10_table.04.html

Serraino, D., Pezzotti, P., Dorrucci, M., Alliegro, M. B., Sinicco, A., Rezza, G., et al. (1997). Cancer incidence in a cohort of Human Immunodeficiency Virus seroconverters. *Cancer, 79*(5), 1004-1008.

Serraino, D., Boschini, A., Carrieri, P., Pradier, C., Dorrucci, M., Dal Maso, L., et al. (2000). Cancer risk among men with, or at risk of, HIV infection in southern Europe. AIDS, 14, 553-559.

Shiels, M.S., Cole, S.R., Wegner, S., Armenian, H., Chmiel, J.S., Ganesan, A., et al. (2008). Effect of HAART on incident cancer and noncancer events among male HIV seroconverters. *J Acquir Immune Defic Syndr*, 48, (4) 485-490.
Shiels, M.S., Cole, S.R., Chmiel, J.S., Margolick, J., Martinson, J., Zhang, Z.-F., et al. (2009). A comparison of ad hoc methods to account for non-cancer AIDS and deaths as competing risks when estimating the effect of HAART on incident cancer AIDS among HIV-infected men. *Journal of Clinical Epidemiology*, Article in Press.

Shiels, M.S., Cole, S.R., Kirk, G.D., & Poole, C. (2009). A meta-analysis of the incidence of non-AIDS cancers in HIV-infected individuals. *J Acquir Immune Defic Syndr*, 52:611-622.

Sobhani, I., Vuagnat, A., Walker, F., Vissuzaine, C., Mirin, B., Hervatin, F., et al. (2001). Prevalence of high-grade dysplasia and cancer in the anal canal in human papillomavirus-infected individuals. *Gastroenterology, 120*, 857-866.

Spano, J. P., Massiani, M. A., Bentata, M., Rixe, O., Friard, S., Bossi, P., et al. (2004). Lung cancer in patients with HIV infection and review of the literature. *Med Oncol*, *21*(2), 109-115.
Sridhar, K. S., Flores, M. R., Raub, W. A., & Saldana, M. (1992). Lung cancer in patients with Human Immunodeficiency Virus infection compared with historic control subjects. *Chest, 102*, 1704-1708.

Tassie, J.-M., Grabar, S., Lancar, R., Deloumeaux, J., Bentata, M., Costagliola, D., et al. (2002). Time to AIDS from 1992 to 1999 in HIV-1-infected subjects with known date of infection. *JAIDS, 30*, 81-87.

Tirelli, U., Spina, M., Sandri, S., Serraino, D., Gobitt, I. C., Fasan, M., et al. (2000). Lung carcinoma in 36 patients with Human Immunodeficiency Virus infection. *Cancer 88*(3), 563-569.

Uronis, H. E., & Bendell, J. C. (2007). Anal Cancer: An Overview. *The Oncologist, 12*, 524-534.

U.S. Department of Health and Human Services Centers for Disease Control and Prevention Division of Cancer Prevention and Control. Cancer (2007) Colorectal Cancer Trends. [Electronic Version], from www.cdc.gov/cancer/colorectal/statistics/trends.htm

Vianna L.E., Lo Y., & Klein R.S. (2006). Serum prostate-specific antigen levels in older men with or at risk of HIV infection. *HIV Medicine, 7*, 471-476.

Winawer, S., Fletcher, R., Rex, D., Bond, J., Burt, R., Ferrucci, J., et al. (2003). Colorectal cancer screening and surveillance: clinical guidelines and rationale—update based on new evidence. *Gastroenterology, 124*, 544-560.

Yu, L-M, Burton, A., & Rivero-Arias, B. (2007). Evaluation of software for multiple imputation of semi-continuous data. *Statistical Methods in Medical Research*, 16:243-258.

Yuan, Y. C. (2002). *Multiple Imputation for Missing Data: Concepts and New Development*. Rockville, MD: SAS Institute Inc.

APPENDIX A. CANCER AGGREGATE COMPARISON FILE

NCI Name	Length	Description	Michigan Variable Names	Michigan Completeness & Comments
AGE_YRS	3	Age in years at diagnosis of this tumor		
BEHAVIOR	1	ICD-O-2 behavior		
BIRTHMM	2	Month of patient's birth		
BIRTHYR	4	Year of patient's birth		
DEATHMM	2	Month of patient's death		
DEATHYR	4	Year of patient's death		
DXConfirm	1	Method of Diagnosis 1=Histology 2=Cytology 3=Radiographic 4=Micro NOS 5=Marker 6=Surgical View 7=Imaging 8=Clinical 9=Unknown		
DXMM	2	Month of Diagnosis		
DXYR	4	Year of Diagnosis		
GRADE	1	ICD-O Grade 1=Well differentiated 2=Moderately differentiated 3=Poorly differentiated 4=Undifferentiated 5=T-cell 6=B-cell 7=Null cell 9=Unknown		
HISP	1	Ethnicity		
ICDOSITE	4	ICD-O-2 site		
MORPH	4	ICD-O-2 histology		
NCI ID	8	NCI generated Subject ID	N/A	Not on the Michigan Registry
NCI SEQNO	2	NCI Tumor Sequence Number	N/A	Not on the Michigan Registry
RACE	2	Cancer Registry Race code 01=White 02=Black 03=American Indian 04-32,96,97=Asian 90=Multi racial 98=Other 99=Unknown		
RSEQNO	2	Registry Tumor Sequence Number		
Sex	1	Sex of patient 1=Male 2=Female		
Stage	1	1=Local 2=Regional		

		3=Distal 9=unknown		
Treatment	1	Was the patient treated for this tumor? 1=Yes 2=No 9=Unknown		
VitalStatMM	2	Month of ascertainment of Vital Status		
VitalStatYR	4	Year of ascertainment of Vital Status		
Vital Status	1	0=Dead 1=Alive 9=Unknown		
Cigarette	1	Cigarette Smoker 0=Never Smoked 1=Ex-smoker 2=Current smoker 9=Unknown		
Alcohol	1	Drinks Alcohol 0=Never 1=Occasional 2=Heavy drinker 9=Unknown		
HIV/AIDS	1	AIDS/HIV 0=No 1=AIDS 2=HIV only 3=AIDS or HIV 9=Unknown		

APPENDIX B. AIDS AGGREGATE COMPARISON FILE

Variable Name	Length	Description
AB_MOYR	5	Date of Overall other HIV antib
AGE_MOS	2	Age in months at AIDS/HIV diagnoses
AGE_YRS	2	Age in yr at dx of AIDS/HIV
ANTIRETV	1	Rec d antiretroviral Rx
ASYMMOYR	5	RxDate diagnosed symptomatic
ARV_DATE	8	Date received anti-retroviral Drugs (***need format)
BACT	1	Bacterial infections
BACTMOYR	5	Dx date for bacteria inf
BIRTHMO	2	Month of Birth
BIRTHYR	4	Year of Birth
BLDPRD	1	Rec d blood prod(clotting fac)
BURKL	1	Burkitt s Lymphoma
BRCNTRY	3	Birth Country (FIPS CODES) (***what is the difference between this and orig_oth?)
BURKMOYR	5	Dx date for Burkitt's lymphoma
CANDESOP	1	Candidiasis esophageal
CANDLUNG	1	Candidiasis lungs
CATEG	1	Category of patient
CCMOYR	5	Dx date for coccidioidomycosis
CDISMOYR	5	Dx date for carcinoma cervical
CERVDIS	1	Carcinoma cervical
CESOMOYR	5	Dx date for esophageal candida
CLNGMOYR	5	Dx date for candida lungs trac
CLASS1	1	HIV clinical case classification C = AIDS B=Symptomatic, not AIDS A=HIV infection, not symptomatic X=Unknown
CLASS2	1	HIV laboratory case classification (***CD4?) 1=(>=500/mm3) or (>29%) 2=(200-499/mm3) or (***14% to 29%) 3=(<200/mm3) or (<14%) 9=unknown
CMBRECNT	1	Result of most recent combi
CMRCMOYR	5	Date of most recent combi test
CMV	1	Cytomegalovirus disease
CMVMOYR	5	Dx date for cytomegalovirus
CMVRET	1	Cytomegalovirus retinitis
CMVRMOYR	5	Dx date for cmv retinitis
CNEGMOYR	5	Date of last NEG combination
COCCI	1	Coccidioidomycosis
COMBIRES	1	Result of Overall combi test
CPOSMOYR	5	Date of 1st pos combi test
CRYPMOYR	5	Dx date for cryptosporidiosis
CRYPTOSP	1	Cryptosporidiosis
CRYPTOCCO	1	Cryptococcosis *** not in HARS???
CTCCMOYR	5	Dx date for cryptococcosis
CTG1DATE	8	Date case met 1985 AIDS definition (***need format)
CTG2DATE	8	Date case met 1987 AIDS definitive
CTG3DATE	8	Date case met 1987 AIDS presumptive

Variable Name	Length	Description
CTG4DATE	8	Date case met 1993 AIDS definitive
CTG5DATE	8	Date case met 1993 AIDS presumptive
CTG6DATE	8	Date case met 1993 AIDS immunologic
CTG7DATE	8	Date case met 1985 AIDS definition
CURCNTRY	3	Patient's current country (FIPS)
Curr_zip	9	Patient's current Zipcode
DEATH	8	Date of patient death
DEMENTIA	1	HIV encephalopathy
DEMMOYR	5	Dx date for HIV encephalopathy
DET_MOYR	5	Date of Overall HIV detection
DXMOYR	5	Date of AIDS dx
DXX_MOYR	5	Date of applicable AIDS dx
EIARECNT	1	Result of most recent EIA
ELISA	1	Result of Overall elisa test
ENEGMOYR	5	Date of last negative EIA
EPOSMOYR	5	Date of first positive EIA
HAGE_MOS	2	Age in months when HIV+
HAGE_YRS	2	Age in years when HIV+
HCATEG	1	Category for HIV case
HDNGMOYR	5	Date of last NEG HIVdetec
HDPOMOYR	5	Date of first POS HIV detec
HDRCMOYR	5	Date of most recent HIV detec
HISP	1	Ethnicity (new in version 6)
HISTMOYR	5	Dx date for histoplasmosis
HISTO	1	Histoplasmosis
HV1LOAD	8	Viral load - Copies per ml
HV2LOAD	8	Viral load - Copies per ml
HV3LOAD	8	Viral load - Copies per ml
HV4LOAD	8	Viral load - Copies per ml
HV5LOAD	8	Viral load - Copies per ml
HV6LOAD	8	Viral load - Copies per ml
HV7LOAD	8	Viral load - Copies per ml
HV8LOAD	8	Viral load - Copies per ml
HV9LOAD	8	Viral load - Copies per ml
HVD1MOYR	5	Month and Year of HV1LOAD (mm/yy)
HVD2MOYR	5	Month and Year of HV2LOAD (mm/yy)
HVD3MOYR	5	Month and Year of HV3LOAD (mm/yy)
HVD4MOYR	5	Month and Year of HV4LOAD (mm/yy)
HVD5MOYR	5	Month and Year of HV5LOAD (mm/yy)
HVD6MOYR	5	Month and Year of HV6LOAD (mm/yy)
HVD7MOYR	5	Month and Year of HV7LOAD (mm/yy)
HVD8MOYR	5	Month and Year of HV8LOAD (mm/yy)
HVD9MOYR	5	Month and Year of HV9LOAD (mm/yy)
HIV2EIA	1	Result of HIV 2 EIA test
HIV2MOYR	5	Date of HIV 2 EIA test
HIV2WBLT	1	HIV 2 Western Blot test
HIVNMOYR	5	Date of last neg of all HIV
HIVPMOYR	5	Date of 1st pos of all HIV
HIV_AB	1	Result Overall oth HIV antib
HIV_DET	1	Result Overall HIVtec
HNEGMOYR	5	Date of last neg oth HIV antib
HPOSMOYR	5	Date of first + other HIVti

Variable Name	Length	Description
HS	1	Chronic mucocutaneous herpes
HSMOYR	5	Dx date for chronic herpes
HVDRECNT	1	Result of most recent HIV det
HVRCMOYR	5	Date most recent other HIV ant
HVRECNT	1	Result of most recent other HI
IBL	1	Immunoblastic lymphoma
HZIP	9	Zipcode of residence when HIV+ (***what about zip code when AIDS +. Current ZIPcode)?
IBLMOYR	5	Dx date for immunoblastic lymp
IMMGLOB	1	Total serum immunoglob. categ.
IMMVALUE	4	Total count immunoglobulins
IMM_MOYR	5	Date of immunoglobulins test
ISO	1	Isosporiasis
ISOMOYR	5	Dx date for isosporiasis
IV	1	IV drug user
IVIGTHER	1	Rec d antiretroviral IVIG Rx
KS	1	Kaposi s sarcoma
KSMOYR	5	Dx date for Kaposi s sarcoma
LIP	1	Lymphoid interstit pneumonia
LIPMOYR	5	Dx date for lymp. inter. Pneum
LYMP_LOW	1	Lymphocyte count (<1000)
MAVIUM	1	Mycobacterium avium complex
MAVMOYR	5	Dx date for m. avium complex
MODE	2	Mode of exposure
MODEX	2	Expanded mode of exposure
MYCO	1	Atypical mycobact diagnosed
MYCOMOYR	5	Dx date for atypical mycobact.
NCI ID	.	NACMR Subject ID
NDI_AUTO	1	National Death Index Autopsy done 1=yes 0=no 8=N/A 9=Unknown
NDI_DXM1	6	NDI Multiple cause of death 1
NDI_DXM2	6	NDI Multiple cause of death 2
NDI_DXM3	6	NDI Multiple cause of death 3
NDI_DXM4	6	NDI Multiple cause of death 4
NDI_DXM5	6	NDI Multiple cause of death 5
NDI_DXM6	6	NDI Multiple cause of death 6
NDI_DXM7	6	NDI Multiple cause of death 7
NDI_DXM8	6	NDI Multiple cause of death 8
NDI_DXU	6	NDI Underlying cause of death
OIX_MOYR	5	Dx date earliest OI Pre 93 d
OI_MOYR	5	Dx date of earliest disease
OTH_HEMO	2	Other coagulation disorder
OTH_IMM	1	Other immunodeficiency
PC	1	Pneumocystis carinii pneumonia
PCMOYR	5	Dx date for pneumocystis pneu.
PCPPROPH	1	Receive PCP prophylaxis
PLB	1	Primary lymphoma of brain
PLBMOYR	5	Dx date for lymphoma of brain
PML	1	Progress multifoc leukoenceph

Variable Name	Length	Description
PMLMOYR	5	Dx date for multifoc. Leuko
PTBMOYR	5	Date of pulmonary TB diagnosis
PULM_TB	1	Pulmonary TB
RACE	1	Patient race (Old) 1=White 2=Black 3=Hispanic 4=Asian 5=Indian 9=Unknown
Race_A	1	Race Asian? 1=yes
Race_b	1	Race Black? 1=yes
Race_I	1	Race American Indian? 1=yes
Race_p	1	Race Hawiian-Pacific Islander? 1=yes
Race_u	1	Race unknown? 1=yes
Race_w	1	Race White? 1=yes
RP	1	Pneumonia recurrent
RPMOYR	5	Dx date for pneu. Recurrent
SALSMOYR	5	Dx date for salmonella sept.
SEX	1	Sex of patient
STAT	1	Current mortality status
SYMPMOYR	5	Date diagnosed Symptomatic
TB	1	M. Tuberculosis
TBMOYR	5	Dx date for m. tuberculosis
th_cn	4	First low CD4 count
th_low	1	T helper cell count (<400)
th_moyr	5	Date of first low CD4 test
th_pct	2	First low CD4 percent (<14)
th1cnt	4	First CD4 count
th1moyr	5	Date first CD4 test
th1pct	2	Percent of first CD4 count
th2cnt	4	Second CD4 count
th2moyr	5	Date 2nd CD4 test
th2pct	2	Percent of second CD4 count
th3cnt	4	Third CD4 count
th3moyr	5	Date 3rd CD4 test
th3pct	2	Percent of third CD4 count
th4cnt	4	Fourth CD4 count
th4moyr	5	Date 4th CD4 test
th4pct	2	Percent of fourth CD4 count
th5cnt	4	Fifth CD4 count
th5moyr	5	Date 5th CD4 test
th5pct	2	Percent of fifth CD4 count
th6cnt	4	Sixth CD4 count
th6moyr	5	Date 6th CD4 test
th6pct	2	Percent of sixth CD4 count
th7cnt	4	Seventh CD4 count
th7moyr	5	Date 7th CD4 test
th7pct	2	Percent of seventh CD4 count
th8cnt	4	Eighth CD4 count
th8moyr	5	Date 8th CD4 test
th8pct	2	Percent of eighth CD4 count
th9cnt	4	Ninth CD4 count

Variable Name	Length	Description
th9moyr	5	Date 9th CD4 test
th9pct	2	Percent of Ninth CD4 count
Thcrecnt	4	Most recent CD4 count
Thlwmoyr	5	Date of lowest CD4 count
Thlwprec	2	Lowest CD4 percent
Thlwrcnt	4	Lowest CD4 count
Thplmoyr	5	Date of lowest CD4 percent
Thprecnt	2	Most recent CD4 percent
Thrcmoyr	5	Date of most recent CD4 test
thts_low	1	T4 8 ratio (<1.0)
TP	1	Toxoplasmosis of brain
TPMOYR	5	Dx date for toxoplasmosis
WASTING	1	Wasting syndrome
WBRECNT	1	Result of most recent west Blo
WB_IFA	1	Result Overall Western Blot
WNEGMOYR	5	Date of last neg Western Blot
WPOSMOYR	5	Date of first pos Western Blot
XKS	1	Kaposis's Sarcoma (y/n) (Created by NCI)
XNHL	1	Non-Hodgkin's Lymphoma (y/n) (Created by NCI from burkl, ibl, & plb)
Zip_code	9	Zipcode at AIDS dx

APPENDIX C. FILE LAYOUT FOR STUDY: ANALYSIS OF TRENDS IN CANCER DIAGNOSIS AND DEVELOPMENT OF AIDS IN PERSONS WITH HIV INFECTION, BEFORE AND AFTER IMPLEMENTATION OF HIGHLY ACTIVE ANTIRETROVIRAL THERAPY

Variable Name	Explanation	Field Length
Case No.	Cases numbered in order	4
Year of Diagnosis of Cancer	4 digit year	4
Month of Diagnosis of Cancer	2 digit month	2
Day of Diagnosis of Cancer	2 digit day	2
Year of Diagnosis of HIV*	4 digit year	4
Month of Diagnosis of HIV*	2 digit month	2
Day of Diagnosis of HIV*	2 digit day	2
Year of Development of AIDS*	4 digit year	4
Month of Development of AIDS*	2 digit month	2
Day of Development of AIDS*	2 digit day	2
Age1*	Age at diagnosis of HIV	2
Age2	Age at diagnosis of Cancer	2
Age3*	Age at development of AIDS	2
Time1	Time in days between diagnosis of HIV and cancer	4
Time2	Times in days between diagnosis of HIV and AIDS	4
CD4	Most recent CD4 count	4
Race	Race of patient	1
Ethnicity	Hispanic ethnicity or not	1
Dead	Vital status of patient at time of match	1
Date of Death	Date of death	8
Tumor Grade	ICDO Grade	1
ICDO Site	ICDO Site	4
Morphology	ICDO Histology	4
SStage	Summary Tumor Stage	1
Radiation therapy**	Received radiation for cancer	1
Surgery**	Received surgery for cancer	2
Chemotherapy**	Received chemotherapy for cancer	2

Note: These variables will be provided to the researcher by staff of the Cancer Registry Program from the Cancer/HIV Matched File or additional variables from the cancer registry file on cases with cancer that are included on the matched file (**). In addition, variables will be provided by HIV/AIDS Epidemiology staff from the unmatched HIV/AIDS registry file (*). Time1 and Time2 will be calculated by the researcher. The variables radiation therapy, surgery, and chemotherapy are first-course treatment variables only. The file provided to the researcher will be a de-identified, unlinked file of matched cancer/HIV/AIDS or unmatched HIV/AIDS cases.

CURRICULUM VITAE

Rose Marie Martin

Education:

Undergraduate:	University of Kentucky B.S. in Mathematics
Graduate:	University of North Carolina M.P.H. in Biostatistics
	University of Kentucky 30 graduate hours in Statistics
	University of Medicine and Dentistry of New Jersey School of Public Health Doctoral Student in Epidemiology

Professional Experience:

April 2003 – July 2006	Research Scientist, Office of the Assistant Commissioner, Division of HIV/AIDS Services, New Jersey Department of Health and Senior Services
January 2000 – April 2003	Research Scientist, Center for Health Statistics, New Jersey Department of Health and Senior Services, Trenton, NJ
July 1991 - January 2000	Chief, Data Analysis and Evaluation Unit, Center for Health Statistics, New Jersey Department of Health and Senior Services, Trenton, NJ
December 1998 - July 1991	Research Scientist, Information and Analysis Unit, Center for Health Statistics, New Jersey Department of Health, Trenton, NJ
April 1985 - October 1988	Chief, Division of Vital Statistics, Center for Health Statistics, Maryland Department of Health and Mental Hygiene, Baltimore, MD
January 1976 - April 1985	Chief, Division of Management Information Services, Drug Abuse Administration, Maryland Department of Health and Mental Hygiene, Baltimore, MD

December 1973 - December 1975	Research Statistician, Drug Abuse Administration, Maryland Department of Health and Mental Hygiene, Baltimore, MD
January 1973 - December 1975	Statistician, Hospital Cost Analysis, Inc., Towson, MD
September 1971 - May 1972	Research Statistician, Kentucky Program Development Office, State of Kentucky, Frankfort, KY
March 1969 - May 1971	Statistician, Department of Community Medicine, University of Kentucky School of Medicine, Lexington, KY
September 1967 - February 1969	Director of Research Statistics, Park-DuValle Neighborhood Health Center, Louisville, KY
May 1960 - September 1967	Office of Biostatistics, Kentucky Department of Health, Frankfort, KY

Publications

Martin, R.M. Drug Use Among High School Students. Drug Abuse Administration, State of Maryland Department of Health and Mental Hygiene, Baltimore, MD, 1975.

Featherstone, N.L., Moss, D.K., Black, K.J., Fiorentino, N.L., Fulcomer, M.C., Martin, R.M., McDonough, D., Morgan, M., Siegel, B. Tien, S-H.M., and Wiseman, B. Healthy New Jersey 2000: A Public Health Agenda for the 1990's. State of New Jersey Department of Health, Office of Health Policy and Research, Trenton, NJ. June 1991.

Martin, R.M., Ficara, S. Regional Health Planning Data Book. State of New Jersey Department of Health, Office of Health Policy and Research, Center for Health Statistics, Trenton, NJ. September 1991.

Martin, R.M., Ficara, S. Bergenfeld, R. New Jersey Health Statistics 1989. New Jersey Department of Health, Office of Health Policy and Research, Center for Health Statistics, Trenton, NJ. May 1992.

Martin, R.M., Ficara, S. New Jersey Health Statistics 1990. New Jersey Department of Health, Division of Health Care Planning, Financing and Information Services, Trenton, NJ. December 1992.

Martin, R.M., O'Dowd, K.J., Palmer, D., Bergenfeld, R. New Jersey Health Statistics 1991. New Jersey Department of Health, Division of Health Care Planning, Financing and Information Services, Trenton, NJ, November 1994.

Holmes, L.J., Ziskin, L.Z., O'Dowd, K.J., Martin, R.M. "Medical Partnerships with Community-based Organizations in Violence Prevention". *New Jersey Medicine*. Lawrenceville, NJ. February 1995.

Martin, R.M., Baron, M. New Jersey Health Statistics 1992. New Jersey Department of Health, Division of Health Care Systems Analysis, Trenton, NJ. September 1995.

Martin, R.M., Baron, M. Cheeks, C. New Jersey Health Statistics 1993. New Jersey Department of Health and Senior Services, Division of Health Care Systems Analysis, Trenton, NJ. July 1996.

Martin, R.M., Cheeks, C. New Jersey Health Statistics 1994. New Jersey Department of Health and Senior Services, Division of Health Care Systems Analysis, Trenton, NJ. December 1996.

Martin, R.M., Baron, M., Cheeks, C. New Jersey Health Statistics 1995. New Jersey Department of Health and Senior Services, Division of Health Planning and Regulation, Trenton, NJ. February 1998.

Martin, R.M., Baron, M., Asare, L. New Jersey Health Statistics 1997. New Jersey Department of Health, Office of Policy and Research, Trenton, NJ. March 2000.

Martin, R.M. Age Standardization of Death Rates in New Jersey: Implications of a Change in the Standard Population. Topics in Health Statistics Series. Available at the New Jersey Department of Health and Senior Services Web site: http://www.state.nj.us/health/chs/topicaa.htm

Martin, R.M. Health Disparities by Race and Ethnicity in New Jersey. Topics in Health Statistics Series. Available at the New Jersey Department of Health and Senior Services Web site: http://www.state.nj.us/health/chs/hlthdisp.pdf

Martin, R.M. Estimates of Healthy Life Expectancy in New Jersey for Selected Race/Sex Subpopulations, 1996-1998. Topics in Health Statistics Series. Available at the New Jersey Department of Health and Senior Services Web site: http://www.state.nj.us/health/chs/topics0202.pdf

Martin RM, Paul, SM. A final update on achievement of the HIV-AIDS objectives in *Healthy New Jersey 2000*. *New Jersey Medicine*. 2004: Vol.101, Nos. 7-8;p.27-32.

Martin RM, Paul SM. Changes in premature mortality due to HIV infection in New Jersey during the 1990s. *New Jersey Medicine*. 2004; Vol. 101 No. 9;p. 31-35.

Sia J, Paul S, Martin RM, Cross H. HIV infection and zidovudine use in childbearing women. *Pediatrics* 2004;114:e707-e712.
URL: www.pediatrics.org/cgi/doi/10.1542/peds.2004-0414

Paul SM, Martin RM, Lu SE, Lin Y. Changing trends in human immunodeficiency virus and acquired immunodeficiency syndrome in the population aged 50 and older. *J Am Geriatr Soc.* 2007 Sep.55(9):1393-7.