EXAMINING THE INCREASING INCIDENCE OF COLORECTAL CANCER IN YOUNGER ADULTS

Caitlin C. Murphy

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Gillings School of Global Public Health (Epidemiology).

Chapel Hill
2016

Approved by:

Robert S. Sandler

John A. Baron

Jennifer L. Lund

Hanna K. Sanoff

Y. Claire Yang

ABSTRACT

Caitlin C. Murphy: Examining the increasing incidence of colorectal cancer in younger adults
(Under the direction of Robert S. Sandler)

The overall incidence of colorectal cancer (CRC) has declined in the U.S., but the incidence of CRC in younger adults (age <50) is rapidly increasing. The underlying mechanisms that have contributed to this increase are poorly understood. This dissertation project sought to: 1) describe the demographic, clinicopathologic, and socioeconomic characteristics and treatment patterns of younger stages II and II CRC patients; 2) estimate the contribution of age, time period, and birth cohort to the increasing incidence of young-onset CRC; and 3) determine patterns of colonoscopy use in younger adults.

Several population-based data sources were leveraged to examine reasons for the increasing incidence of young-onset CRC. The National Cancer Institute's Patterns of Care studies were used to describe differences in the characteristics of stages II and III CRC patients by age at diagnosis. Hierarchical Poisson models were used to estimate the independent contribution of age, time period, and birth cohort on increases in the incidence of young-onset CRC. Incidence rates were derived from the Surveillance, Epidemiology, and End Results program of cancer registries. The prevalence of CRC risk factors, including obesity, physical inactivity, and smoking, were also described across time period and birth cohort using data from national surveys. Lastly, MarketScan Commercial Claims and Encounters Data, an employer-based claims database, was used to characterize patterns of colonoscopy use in younger adults.

There were differences in the distribution of young-onset CRC by race/ethnicity. A higher proportion of black and Hispanic patients were diagnosed at a younger age compared to whites. Results of the age-period-cohort analysis showed a significant age and birth cohort effect in both younger and older populations, but the effect of time period was only observed among older ages. The prevalence of obesity generally increased across both time period and birth cohort, while smoking prevalence declined. Trends in physical inactivity remained relatively constant. Colonoscopy use among younger adults increased across sex, age, and geographic region from 2001 to 2009 and decreased through 2013.

This study provides strong support for different mechanisms involved in the development of CRC across the life course. The factors responsible for increases in young-onset CRC, albeit small on in absolute magnitude, remain unknown.

ACKNOWLEDGEMENTS

I would like to thank Dr. Robert Sandler for his support and trust in my work (and timeline). Like Dr. Sandler, the best mentors are those who teach by their example. I would also like to thank my committee members for their thoughtful feedback and care, Dr. Linda Harlan, Dr. Joan Warren, and Dr. Ann Geiger from the Healthcare Delivery Research Program at the National Cancer Institute for their endless encouragement and allowing me to use the Patterns of Care data, and Dr. Sally Vernon at the University of Texas School of Public Health for her inspiration and guidance since the very beginning.

I have received funding support from the Digestive Disease Epidemiology Training Program (T32 DK07634) in the Division of Gastroenterology and Hepatology at UNC School of Medicine.

And to my family and friends, for never questioning my resolve or slowing me down; in particular, Heather, for her impeccable editorial skills, patience, and iced beverages, Mary, for a Bethesda home, Allison, for exciting travel and much-needed distraction, the Gimbel-Welles family, for summer writing retreats on Martha's Vineyard, and Wayne, for Tuesdays.

In the years since my mom died from breast cancer, I have been fortunate to know countless others on a similar mission against cancer. Each has shared their kindness and wisdom with me, a gift that gives deeper meaning to this work. As for my mom, I would be amiss if I did not acknowledge the hours she spent editing my high school essays and her faith that one day I could brave it on my own. She once stayed awake until the early hours of the morning to help me finish a project on literary tone; we scoured our household books for passages that captured the essence of each tone word—playful, disgruntled, resilient. On other occasions we spent inordinate amounts of time together dissecting letters in words in sentences in essays, a tedious process that has cursed me with a tendency toward perfectionism. I never anticipated, during the late nights on our Compaq Presario, that I would author numerous scientific publications, much less write a lengthy doctoral dissertation. Somehow, though, I think she always knew.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AHRF	Area Health Resource File
APC	age-period-cohort
BMI	body mass index
BRFSS	Behavioral Risk Factor Surveillance System
CCREM	cross-classified random effects model
CDC	Centers for Disease Control and Prevention
CPT	Current Procedural Terminology
CRC	colorectal cancer
CS	Collaborative Staging
EOD	Extent of Disease (staging)
FOBT	fecal occult blood test
HAPC	hierarchical age-period-cohort
ICD-9	International Classification of Diseases, Ninth Edition
ICD-O-3	International Classification of Disease, Oncology, Third Edition
ML	maximum likelihood
MMR	mismatch repair
MSI	microsatellite instability
NCI	National Cancer Institute
NHANES	National Health and Nutrition Examination Survey
NHIS	National Health Interview Survey
PCR	polymerase chain reaction
POC	Patterns of Care
REML	restricted maximum likelihood
SEER	Surveillance, Epidemiology and End Results

USPSTF U.S. Preventive Services Task Force

CHAPTER 1
STATEMENT OF SPECIFIC AIMS

The overall incidence and mortality of colorectal cancer (CRC) have declined over the past two decades in the U.S., but recent research suggests CRC in younger adults (age <50 years) is rapidly increasing. Incidence rates have increased annually by up to 4% for rectal cancer and 2% for colon cancer among younger patients since the 1970s, with the steepest increases in the 40-49 year age group. By 2030, approximately 11% and 23% of all colon and rectal cancers, respectively, will be diagnosed in patients younger than the current screening age. Because guidelines recommend that screening (with hemoccult testing, colonoscopy, or sigmoidoscopy) begin at age 50 for average risk individuals, early recognition of CRC in young patients without a known family history or genetic predisposition (i.e., familial adenomatous polyposis, Lynch syndrome) is challenging. This is especially concerning because more than half of young patients report no family history of CRC, and only a small minority have hereditary cancer syndromes.

Although the magnitude of the increase in young-onset CRC incidence has been previously described, the underlying mechanisms that have contributed to this increase are poorly understood. Reasons for the increase in incidence in younger patients remain largely unknown. The current literature is limited to single institution settings, small sample sizes, and inconsistent inclusion criteria across studies (e.g., age cutoff of young-onset CRC). No study has examined the potential influence of modifiable risk factors or health services use on changes in incidence. Research is needed to better describe the determinants of CRC in

younger adults. This study used a multidimensional approach to address the characteristics of and factors contributing to young-onset CRC. The specific aims were as follows:

Specific Aim 1a*:* Describe the demographic, clinicopathologic, and socioeconomic characteristics of younger and older patients newly diagnosed with stages II and III CRC during the period 1990-2010. *Hypothesis – Specific Aim 1a:* Younger and older stages II and III CRC patients differ by demographic, clinicopathologic, and socioeconomic characteristics.

Specific Aim 1b: Examine differences in treatment patterns, including type of surgery, chemotherapy receipt, and radiation receipt, by age at diagnosis. *Hypothesis – Specific Aim 1b*: Younger CRC patients more frequently receive chemotherapy and radiation therapy compared to older patients.

Patients newly diagnosed with stages II and III CRC were identified from the National Cancer Institute's Patterns of Care (POC) studies in 1990-2010. Detailed tumor information (e.g., histologic grade, tumor site) was abstracted from patient medical records and verified by treating physicians. POC data were linked with the Area Health Resource File to describe socioeconomic indicators and healthcare access in this patient population. As a secondary aim, treatment patterns (e.g., chemotherapy agents and regimens received) were described. Characteristics of younger patients (age <50 years) were compared to older CRC patients (age ≥50 years) during the same period to determine differences in disease patterns.

Specific Aim 2a*:* Estimate the contribution of age, time period, and birth cohort to the increasing incidence of young-onset CRC. *Hypothesis – Specific Aim 2a*: There are differential effects of time period and birth cohort in the incidence of CRC by age.

Specific Aim 2b: Describe changes in the prevalence of lifestyle-related modifiable risk factors, including obesity, physical inactivity, and smoking, by time period and birth cohort.

Hypothesis – Specific Aim 2b: Changes in the prevalence of modifiable risk factors parallel increases in the incidence of young-onset CRC.

Hierarchical Poisson models were used to determine the extent to which age, time period, and birth cohort account for changes in incidence over time. CRC incidence rate data was derived from the Surveillance, Epidemiology, and End Results 9 registries (1973-2012). Prevalence of obesity (% body mass index ≥30 kg/m^2), physical inactivity (% no physical activity in last 30 days), and current smoking was estimated using the National Health and Nutrition Examination Survey, Behavioral Risk Factor Surveillance System, and National Health Interview Survey, respectively.

Specific Aim 3: Determine patterns of colonoscopy use in individuals younger than age 50 years. *Hypothesis – Specific Aim 3*: Colonoscopy use in younger adults has increased over time but does not fully account for trends in the incidence of young-onset CRC.

An estimated 14 million colonoscopies are performed in the U.S. annually, but colonoscopy use in younger populations not yet eligible for screening has not been extensively studied. Individual-level healthcare claims from MarketScan (Truven Health Analytics, Ann Arbor, MI), a large employer-based claims database, were used to determine changes in the number of colonoscopies performed in persons age younger than 50 years over the period 2001-2013. Understanding increases in colonoscopy use provides additional information on CRC trends in younger adults that expand knowledge gained in Aims 1 and 2.

This dissertation study used an innovative approach to leverage data from three large, well-defined sources to examine the increasing incidence as a function of patient and tumor characteristics, modifiable risk factors (obesity, physical inactivity, current smoking), and health services use. By using complementary methods and techniques from the fields of epidemiology, demography, and health services, the study provides insight into the reasons for and distribution

of CRC in younger populations. Results may inform clinical guidelines and recommendations regarding the prevention, diagnosis, and management of CRC in younger adults.

CHAPTER 2
REVIEW OF LITERATURE

A. BACKGROUND

Increasing Incidence of Colorectal Cancer in Younger Adults

The incidence of colorectal cancer (CRC) in younger adults is rapidly increasing in the U.S. Data from the National Cancer Institute's (NCI) Surveillance, Epidemiology, and End Results (SEER) Program show the incidence of CRC has increased by up to 5% per year among younger patients (age <50 years) since the 1970s.[1-6] The steepest increases have occurred in the 40-49 year age group. For example, CRC incidence was 11.9 per 100,000 among 40-44 year olds in 1987 but rose to 17.9 per 100,000 in 2006, an increase of 67%.[1] If current trends persist, in the next 20 years the incidence of colon and rectal cancer may increase by 90% and 124%, respectively, for patients 20-34 years, and by 28% and 46%, respectively, for patients 35-49 years (Figure 2.1).[7] Despite the overall population trends in aging, by 2030, approximately 11% of all colon and 23% of all rectal cancers are expected to be diagnosed in patients younger than age 50.[7]

The rise of CRC in younger adults (i.e., young-onset CRC) stands in sharp contrast to CRC incidence patterns in older populations. Beginning in the 1990s, CRC incidence and mortality rates have consistently decreased among adults age 50 years and older, with the largest absolute declines among adults 65 years and older.[8] Overall 5-year relative survival rates have simultaneously increased, from 50% in 1975-1977 to 66% in 2003-2009.[8] Much of this improvement has been attributed to screening.[9] CRC screening facilitates earlier

detection of CRC and lowers mortality and incidence by removing premalignant polyps (e.g., adenomatous, sessile serrated, traditional serrated). Current guidelines recommend screening with colonoscopy, sigmoidoscopy, or fecal occult blood test for those at average risk begin at age 50.[10] Other than cervical cancer, CRC is the only cancer for which both incidence and mortality can be reduced through population-based screening.[11]

Challenges of Diagnosing CRC in Younger Adults

Early recognition of CRC in younger patients is challenging. Because guidelines recommend that screening begin at age 50 for individuals at average risk of CRC, it is difficult to diagnose CRC in young patients without a known family history or genetic predisposition (i.e., familial adenomatous polyposis, Lynch syndrome). This is especially concerning because more than half of young patients report no family history of CRC, and a small minority have hereditary cancer syndromes.[12] A recent study[13] found only 5% of individuals in a random sample of patients who were diagnosed with CRC before age 50 had Lynch syndrome; this finding is only slightly larger than the prevalence of Lynch syndrome (1-3%) in all patients with CRC.[14] The use of screening in younger populations is limited (i.e., not considered appropriate care), and symptoms often go unrecognized. Younger patients who are symptomatic often do not seek medical attention for several months after the onset of symptoms, [15,16] and, because CRC is not commonly seen in this population, physicians may not attribute symptoms such as abdominal pain and rectal bleeding to cancer.

B. CHARACTERISTICS OF YOUNGER CRC PATIENTS

Sociodemographic and Clinicopathologic Features of Young-Onset CRC

Characteristics of younger CRC patients have not been extensively studied. Studies of young-onset CRC typically use demographic variables to describe incidence patterns in population subgroups (e.g., by race/ethnicity) and clinicopathologic and molecular variables to

examine biological features of the tumor that may explain the earlier onset of CRC. Data from SEER registries and the National Cancer Database show the incidence of young-onset CRC is similar among males and females,[2,5,6] but African Americans have higher incidence rates compared to non-Hispanic whites.[2,5,17] Other factors with the most consistent association with young-onset CRC are tumor site (rectum or right-colon),[15,18-22] mucinous or signet ring cell histology,[18,19,21,23-29] poorly differentiated or undifferentiated grade tumors,[15,18,21,23,24,26-28,30] and a higher stage of disease at diagnosis.[15,18,21,23,24,26-28,31,32] Lymphovascular invasion,[12,20,21,23-25,28,33,34] perineural invasion,[19,21,23,33,34] and synchronous tumors[23,28,33] are less consistently associated with young-onset CRC, with some studies reporting a positive relationship and others suggesting no association. Tumor growth patterns (e.g., infiltrative vs. expanding, gross type) have not been frequently studied.[25,28,34] Appendix A provides a summary of the empiric findings regarding the relationship between young-onset CRC and select patient and tumor characteristics in studies with a comparison group; Appendix B provides a summary of those findings in studies without a comparison group.

Molecular Pathways of CRC Carcinogenesis in Younger Patients

Established pathways of CRC carcinogenesis only account for a small subset of young patients diagnosed with CRC. Carcinogenic pathways are often defined by their molecular features: chromosomal instability, microsatellite instability (MSI), and CpG island methylation (CIMP), with mutations in specific genes often associated with each pathway.[35] The majority of CRCs are thought to arise through the classic adenoma-carcinoma sequence of pathogenesis, which involves alterations to *APC*, *KRAS*, *SMAD2*, *SMAD4*, and *TP53* genes. *APC* and *KRAS* mutations are found in 80% and 45% of these carcinomas, respectively.[36] The second major pathway of carcinogenesis is characterized by the presence of MSI due to defective mismatch repair (MMR) genes. In non-inherited CRC, this is almost always due to the acquired promoter hypermethylation of the *MLH1* gene that silences its expression. Finally, the CIMP pathway is

characterized by widespread aberrant DNA hypermethylation at select CpG islands (i.e., DNA regions with a high frequency of CpG sites, where a cytosine nucleotide occurs next to a guanine nucleotide) and are preferentially located in the proximal colon and associated with *BRAF* mutation.[37] Tumors are often both CIMP and MSI-high because methylation-associated MSI generally does not occur among sporadic (i.e., not hereditary) cases outside the context of CIMP.[38] A few investigators have recently explored the molecular features of young-onset CRC and found no difference in the overall mutational rate (as measured by the number of single nucleotide variations per patients) among younger and older patients.[39,40] However, the specific mutations involved in young-onset CRC appear to be distinct. *SMAD4*, MMR genes (*MSH6*, *MLH1*, *MLH2*), *ARID1A*, *IGF1R*, and *KIT* have been shown to be more frequently mutated among younger patients. Compared to older patients, the majority of tumors from young patients do not show *KRAS*, *APC*, *TP53*, and *BRAF* mutations, [20, 28, 39] which are common features of carcinogenic pathways in CRC. These findings suggest the biology of young-onset CRC may be different from what is currently known about the biology of CRC in the overall population.

Treatment of Young-Onset CRC

Younger CRC patients often receive more aggressive treatment regimens. Studies consistently show that younger CRC patients more frequently receive chemotherapy and/or radiation therapy compared to older patients.[19,41] Even among stage II colon cancer patients, for whom guidelines recommend against adjuvant chemotherapy,[42] a large proportion of younger patients receive therapy.[33,43-45] The higher proportion of younger CRC patients treated with chemotherapy and/or radiation therapy may be an indication of both over-treatment of normal-risk and appropriate treatment of high-risk (e.g., poorly differentiated histology, lymphovascular invasion) patients. Most studies, however, do not describe patterns of care in subgroups of patients by age and high-risk features (i.e., do not make the distinction between therapy given

because of younger age and therapy given for aggressive tumor biology). Further, differences in the type of chemotherapy agents given to older and younger patients have not been well studied. Despite more aggressive therapies in younger patients, many studies report that the overall survival is no better than that observed in older CRC patients.[18,21,23,24,26,33,34,46,47]

The literature on characteristics of younger CRC patients is limited in several ways. Most studies are of patients treated in a single institution. Few of the same variables have been examined across studies, and even when the same characteristics were studied, they were defined or measured differently. Inclusion criteria are not consistent (e.g., some studies define young-onset CRC as <40 years while others use <50 years), and there are differences in stage at diagnosis among younger and older patients (i.e., younger patients tend to be diagnosed at later stages), making comparisons difficult. It is not possible to draw conclusions from the few exploratory studies of molecular differences in young- and older-onset CRC. As a consequence, we know very little about the relative importance of clinicopathologic and/or molecular features in the development of young-onset CRC. Findings of previous studies warrant further investigation in larger, population-based samples.

C. REASONS FOR THE INCREASE IN YOUNG-ONSET CRC

Modifiable Risk Factors of CRC

Reasons for the increase in CRC incidence in younger patients are poorly understood. Although the magnitude of the increase has been previously described, little is known about the underlying mechanisms that have contributed to this increase. Several established modifiable risk factors for CRC, including unhealthy diet, obesity, physical inactivity, and smoking, have been proposed as the major drivers of the increase in adults age <50 years.[48,49] Meta-analyses demonstrate significant associations between CRC risk and red and processed meat consumption,[50,51] adiposity (including measures of body mass index, waist circumference, and

waist-to-hip ratio),[52-54] sedentary behavior and physical inactivity,[55] and cigarette smoking.[56] Although some studies report only a modest association between diet, adiposity, and CRC, the relationship between CRC and physical activity is consistent across multiple studies and in different settings and populations.[57] Increases in the prevalence of many of these risk factors have paralleled the rise in young-onset CRC. Consumption of fast food and/or food prepared away from home increased 5-fold among children and 3-fold among adults between the late 1970s and mid-1990s.[58] Obesity prevalence has also risen dramatically among adults, from 13% in 1960-1962[59] to 35% in 2011-2012.[60] Central/abdominal obesity, in particular, which is associated with a higher relative risk of CRC[54], has increased by nearly 50% in men and women over the last 25 years.[61] As poor dietary behavior and obesity has risen, physical activity rates in children and young adults have also declined. More than half of U.S. adults and adolescents do not meet recommended physical activity guidelines.[62] The prevalence of current smoking has varied over time, with an overall decreasing trend; however, younger adults (age 18-35) reported the highest prevalence of current use of tobacco products in recent national surveys compared to other age groups.[63] The extent to which changes in the prevalence of risk factors explain growth in the incidence of young-onset CRC has not been studied.

D. ROLE OF HEALTH SERVICES USE

Prevalence of Endoscopy Use

Changes in the use of endoscopy procedures, such as colonoscopy, may also facilitate further understanding of the incidence of CRC in younger patients. Approximately 11-14 million colonoscopies are performed each year in the U.S.,[64,65] but little is known about the prevalence of colonoscopy use in populations that are not yet eligible for screening recommendations. Across all age groups, the overall use of colonoscopy has dramatically increased in the last 15 years. The National Health Interview Survey, which is considered the gold standard for estimating the prevalence of cancer screening behaviors, shows an increase in the use of

colonoscopy for CRC screening, from 19% in 2000 to 55% in 2010.[66,67] Other studies of administrative billing claims estimate there were 3,800 lower endoscopies (for any indication) performed per 100,000 insurance enrollees in 2009, an increase of 17% since 2000.[65]

Colonoscopy is a key component of the CRC care continuum. Although adults age <50 years generally do not undergo routine colonoscopy, patterns of colonoscopy use in younger populations may explain some of the increase in young-onset CRC. For example, if colonoscopy use has increased by 10%, while the incidence of CRC has simultaneously doubled, it would suggest that endoscopy does not fully account for the increase in young-onset CRC. Understanding the use of colonoscopy in persons age <50 years is important in determining the extent to which endoscopy independently accounts for changes in the incidence of young-onset CRC.

E. SUMMARY

Multidimensional Framework

The burden of young-onset CRC is rapidly increasing. Reasons for this increase are complex. Most studies have failed to consider the multiple influences that have likely contributed to CRC in younger adults, instead only focusing on a single dimension (e.g., clinicopathologic characteristics) of CRC. Evidence is lacking regarding the impact of modifiable risk factors and health services use on young-onset CRC. Guided by a conceptual model (Figure 2.2), this dissertation project used complementary methods to better describe and understand the underlying mechanisms that have led to the increase of CRC in adults age <50 years.

Innovation of the Study

This study was the first to examine reasons for the increase in young-onset CRC. The rapid growth in CRC among younger adults has been well described, but several important gaps in understanding remain. Sociodemographic characteristics and clinicopathologic features of

young CRC patients are largely unknown. Many studies are limited by small sample sizes, confounding by stage at diagnosis, single institution settings, and lack of a comparison group. In addition, the etiology of young CRC is not well understood. Trends in unhealthy dietary choices, obesity, and physical inactivity, and endoscopy use among younger adults may contribute to increases in CRC. However, risk factors have not been systematically evaluated in this setting, and no study has examined patterns of colonoscopy use in younger adults not eligible for screening recommendations. Population-based analyses are critical to understanding the unique disease patterns of young-onset CRC. This dissertation project used an innovative approach to address these concerns by leveraging data from three large, well-defined sources (majority population-based) to examine the increasing incidence of young-onset CRC as a function of patient and tumor characteristics, lifestyle-related modifiable risk factors (obesity, physical inactivity, current smoking), and health services use. By using complementary methods and techniques from the fields of epidemiology, demography, and health services, the study provides insight into the reasons for and distribution of CRC in younger populations. Results can be used to inform clinical guidelines and recommendations regarding the prevention, diagnosis, and management of CRC in younger adults.

Figure 2.1. Annual percentage change-based predicted incidence rates of colon (A) and rectal (B) cancers by age compared with incidence rates in 2010[68]

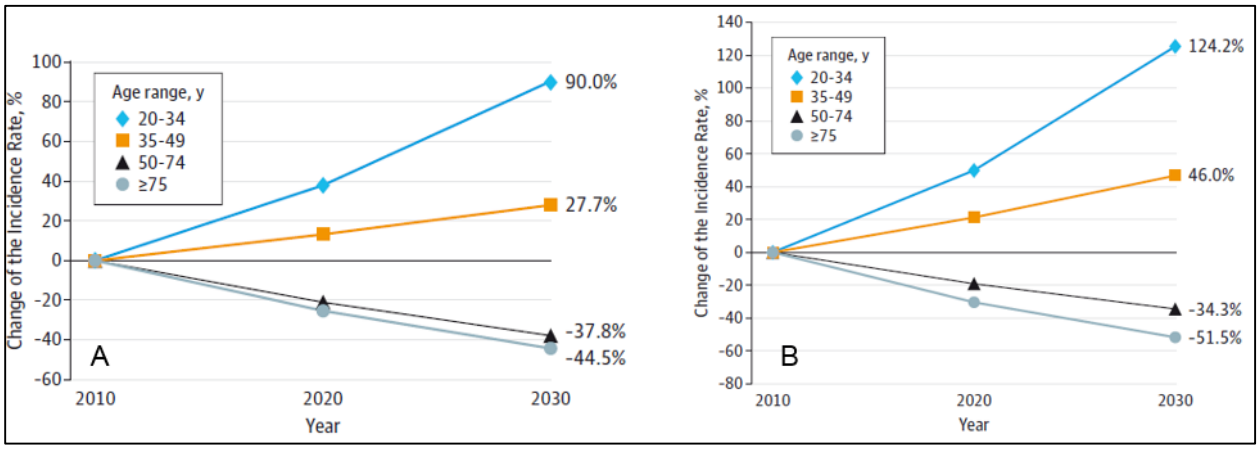
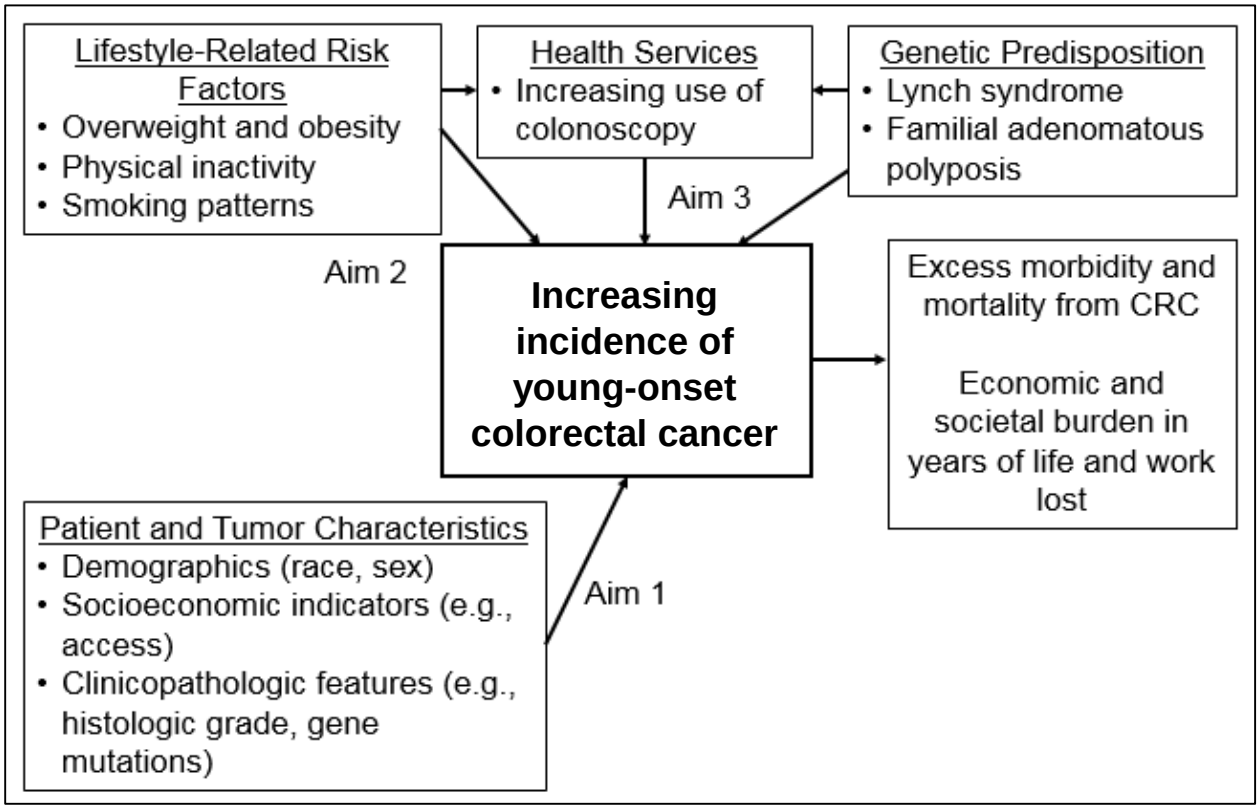


Figure 2.2. Conceptual model of increasing incidence of young-onset colorectal cancer



CHAPTER 3
METHODS

The overall goal of this dissertation study was to determine the underlying mechanisms that have contributed to the rising incidence of young-onset colorectal cancer (CRC). Three cohorts were assembled from large, well-defined data sources to gain insight into the reasons for and distribution of CRC in younger populations (Table 3.1).

A. SPECIFIC AIM 1 APPROACH

Study Design

In *Specific Aim 1*, we described the sociodemographic characteristics and clinicopathologic features of younger (age <50 years) patients diagnosed with CRC. As a secondary aim, we described treatment patterns (e.g., type of surgery and receipt of chemotherapy and radiation therapy) in this population. Stages II and III CRC patients were sampled from the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) program of cancer registries. SEER routinely collects information from hospitals, pathology laboratories, surgical centers, and radiation facilities on patient demographics, tumor characteristics, and first-course treatment for all persons diagnosed with cancer residing in SEER geographic regions. Chemotherapy and radiation therapy (neoadjuvant and adjuvant) are often underreported in SEER; therefore, the NCI annually conducts Patterns of Care (POC) studies on a sample of patients with select cancers to assess the extent to which adjuvant therapies are received in clinical practice.

Patients' medical records were abstracted to verify information on patient demographics (age, sex, race/ethnicity), clinicopathologic characteristics (tumor site, stage, histologic grade,

KRAS mutation, *BRAF* mutation, microsatellite instability [MSI]) and treatment planned or given as reported in SEER. As part of POC studies, additional treatment detail, including receipt of specific chemotherapy agents, radiation therapy, and dates of treatment following diagnosis, was collected from medical records by trained abstractors. Because adjuvant therapies are often given outside of the hospital setting, and SEER data are primarily hospital-based, the treating physician was contacted to verify therapy received or recommended. Treating physicians were also asked to provide names and addresses of other physicians who may have treated the patients, who were subsequently contacted for treatment information. Doctor verification substantially improves completeness of chemotherapy ascertainment or confirms that no chemotherapy or radiation were given. Physician responses were received on more than 85% of sampled patients.

We linked POC data with the Area Health Resource File (AHRF), an extensive county-level database comprised of socioeconomic indicators (e.g., education level, % living below poverty line). Linking AHRF with POC by patient county of residence allowed us to better describe sociodemographic characteristics of the cohort. A cohort design based on extant data was an efficient approach to obtain a large sample of CRC patients and address our research question. Such designs are often called historical because all follow-up information is available at the study outset; however, all data within POC were collected prospectively (i.e., in the year following diagnosis), and we did not take any future events into consideration when creating the cohort.

Study Population

Stages II and III CRC patients in participating SEER registries were eligible for POC studies in 1990, 1991, 1995, 2000, 2005, and 2010.[41,69,70] Participating registries included the metropolitan areas of Atlanta (1990-2010), Detroit (1990-2010), Los Angeles (1995, 2000, 2005, 2010), San Francisco (1990-2010), San Jose-Monterey (2000, 2005, 2010), and Seattle (1990-

2010), and the states of Hawaii (1990-2010), New Mexico (1990-2010), Iowa (1990-2010), Connecticut (1990-2010), Alaska (2000 only), Utah (1990-2010), Kentucky (2010 only), Greater California (2005, 2010), Louisiana (2005, 2010), and New Jersey (2005, 2010). Eligible patients were stratified within registries by tumor site (colon or rectum), and a random sample was taken from within each stratum. Beginning in 1995, there was oversampling by race/ethnicity to obtain more stable estimates for racial subgroup analyses. Patients were sampled according to the staging scheme used by SEER in each study year. In 1990, 1991, 1995, and 2000, patients were sampled based on Extent of Disease (EOD) 10 coding, and in 2005 and 2010, patients were sampled based on Collaborative Staging (CS) coding. EOD and CS coding record the farthest extent of disease based on the combined clinical and pathological assessment. Clinical information took priority when a patient was treated with preoperative therapy, otherwise pathological information took priority. TNM staging was derived by mapping T and N status from EOD and CS coding. Stage II included T3 or T4 tumors with no positive regional lymph nodes, and stage III included any T1 to T4 tumors with regional lymph node involvement. These stage definitions also correspond approximately to stage B2 and C of the Aster-Coller modification of Duke's original staging system.

Patients were ineligible for POC studies if they were younger than age 20; previously diagnosed with cancer (excluding non-melanoma skin cancer); diagnosed at autopsy or on death certificate only; or diagnosed with a synchronous cancer. We considered young CRC patients to be those age 20-49 years, and older CRC patients were age greater than 50 years.

An estimated 7,000 patients were diagnosed with stages II and III CRC during the study period; approximately 1,400 of these were patients age <50 years (Table 3.1). POC studies provided a number of unique advantages for conducting population-based epidemiologic research because each participating registry area has a defined population, and detailed tumor and treatment information was abstracted from patient medical records and verified by treating

physicians. The data available in POC studies provided a greater breadth and depth of information than that available solely from medical claims and/or SEER registries. The age and sex distributions of patients in POC reflect those of the U.S. population, and the SEER program includes registries with a high percentage of African Americans (Detroit, Atlanta, Louisiana), Asian Americans (Seattle, San Francisco, Los Angeles, San Jose-Monterey), and Hispanics (Los Angeles, Greater California). The large size and ethnic diversity of this study population were strengths that enabled us to examine CRC characteristics within population subgroups by race and sex.

Covariate Assessment

A summary of covariates that were measured in *Specific Aim 1* is provided in Table 3.2. Demographic and clinicopathologic information was abstracted from patient medical records by trained abstractors. An example of the abstraction form is shown in Appendix C. Demographic characteristics included age (years), sex (male, female), race/ethnicity (non-Hispanic white, non-Hispanic black, Hispanic, other), insurance (private, Medicare, Medicaid, other, none), and delivery site/geographic region. Clinicopathologic characteristics included tumor site (ascending, transverse, or descending colon, hepatic flexure, splenic flexure, sigmoid colon, rectosigmoid junction, rectum), stage (II, III), histology (adenocarcinoma, mucinous, signet ring cell), and histologic grade (well, moderately, poorly, undifferentiated).

TNM staging was derived by mapping T and N status from EOS and CS coding (described above), where stage II included T3 or T4 tumors with no positive regional lymph nodes, and stage III included any T1 to T4 tumors with regional lymph node involvement. Low-grade CRC was defined as well and moderately differentiated tumors, and high-grade CRC included poorly differentiated and undifferentiated tumors, according to established protocols of the College of American Pathologists and American Joint Committee on Cancer.[71,72] Mucinous

and signet ring cell tumors were defined as those with more than half of the tumor displayed as extracellular mucin or signet rings, respectively.[73,74]

KRAS mutation, *BRAF* mutation, and MSI were collected in 2010 only. The purpose of *KRAS* testing is to identify point mutations in codon 12 and 13 of the KRAS gene. There are currently no specific methodology recommendations and no FDA-approved *KRAS* mutation tests available; however, all methods are based on the polymerase chain reaction (PCR). The most common method for KRAS testing is direct sequencing of *KRAS* PCR products.[75] Although the specific method of *KRAS* genotyping may differ across POC sites, a recent study reported good agreement (90% concordance) in *KRAS* mutation status of tumors from patients across different detection methods (e.g.., sequencing, hybridization).[76] *BRAF* mutation testing, used to detect the presence of a V600E mutation on the *BRAF* gene, is most often performed by amplification and direct DNA sequence analysis, although allele-specific PCR is an acceptable alternative method. PCR-based *BRAF* testing requires small quantities of DNA and is highly sensitive.[77] Finally, microsatellite status is evaluated using tissue sections from the primary tumor with a PCR assay using a panel of five microsatellite markers (BAR-25, BAT-26, D2S123, D5S346, D17S250), as recommended by the NCI.[78] Tumors were considered MSI-high if two or more of the five microsatellite sequences were mutated; tumors with only one mutated sequence were classified as MSI-low. Microsatellite stable tumors had no mutations in the microsatellite panel.

Any surgery, radiation therapy, and chemotherapy given as first-course treatment was abstracted from patient medical records. An example of the treatment verification form is given in Appendix D. Surgery included surgery type (polypectomy, partial colectomy, subtotal colectomy/hemicolectomy, total colectomy, total proctocolectomy, or coloproctotectomy with resection of a continuous organ), date, number of lymph nodes examined, number of positive lymph nodes, pathological margins, and whether the colectomy was a laparoscopic procedure.

Radiation therapy included date of initiation, sequence with surgery (before, after, before and after, intraoperative), and sequence with chemotherapy (before, after, concurrent).

Chemotherapy included dates of initiation (dates reported for all therapies initiated), sequence with surgery (before, after, before and after), specific agents (e.g., 5-flourouracil, oxaliplatin), and combination regimens (e.g., FOLFOX, FOLFIRI).

Socioeconomic indicators were derived from POC, SEER, and the AHRF. POC contains patient-level information on clinical trial participation, hospital type (private, nonprofit, or government), hospital size (based on total bed size), and approved residency training program. We also used a composite index of socioeconomic status based on measures developed by Yost et al.,[79] including occupation, unemployment, poverty, education, income, and housing. The index was constructed to assess the relationship between socioeconomic status and cancer incidence using SEER data.[80] Data used in the index were derived from Census 2000 and American Community Survey 2005-2009 and reflect the populations and census tracts covered by the SEER 17 registries. The index (measured in quintiles) was available for study years 2000, 2005, and 2010. Lastly, we used AHRF measures of per capita income, median household income, education level (persons age ≥25 years with less than a high school diploma, high school or more, or four or more years of college), poverty (persons living below poverty line), unemployment, total number of active physicians, and total number of gastroenterologists. All socioeconomic indicators in the AHRF were measured on the county level and collected at various time points. In cases where an AHRF variable was not collected in the same year as POC, we used the next closest collection year. For example, total number of physicians was not collected in AHRF in 2010; therefore, we used data on total number of physicians collected in 2008 as a reasonable proxy.

Statistical Analysis

To test the hypothesis that there are differences in the sociodemographic characteristics and clinicopathologic features of young-onset CRC and older-onset CRC, we examined the distribution of covariates in young (age <50 years) and older (age ≥50 years) CRC patients with descriptive statistics (means, medians, minima, maxima, frequencies, cross-tabulations). As was appropriate, we also examined the distribution of covariates within subgroups of younger and older CRC patients by race and sex. As part of the secondary aim to examine differences in treatment patterns, we compared the proportions of younger and older CRC patients who received common chemotherapy agents, combination chemotherapy regimens, and radiation therapy, as well as differences in type of surgery performed and surgical outcomes (e.g., number of lymph nodes examined). For all analyses, proportions and means were calculated with stratum-specific sample weights to account for the complex survey design. Sample weights were calculated as the inverse of the sampling proportion for each sampling stratum.

Sensitivity Analysis

To account for potential heterogeneity of CRC in older patients (e.g., the CpG island methylator phenotype is most common in female CRC patients age ≥70 years), we conducted sensitivity analyses that further categorized older CRC patients into two age groups: 50-69 years and ≥70 years. Younger (age <50 years) CRC patients were compared to older CRC patients in both the 50-69 and ≥70 year age groups.

Sample Size and Power

We assumed that a minimum of 10% absolute difference was meaningful to detect. We set the level of significance at 5% and statistical power at 80% for a 2-sided test of all hypotheses. Power analyses were conducted in SAS (version 9.3). Table 3.3 provides the minimum statistically-detectable group differences across a range of prevalences among older

CRC patients (from 10% to 50%). The sample size of 1,400 younger CRC patients and 5,600 older CRC patients (see Table 3.1) provided 80% power to detect as little as a 4% difference in younger and older CRC patients when the prevalence of the covariate was 10% among older patients. If the prevalence of a covariate under study was as high as 50% among older patients, we were able to detect a difference of 7%.

B. SPECIFIC AIM 2 APPROACH

Study Design

In *Specific Aim 2*, we estimated the contribution of age, time period, and birth cohort on the increasing incidence of young-onset CRC using Poisson cross-classified random effects models (CCREM). CCREM have been widely used in sociology and demography research[81] and have more recently been applied to cancer incidence and mortality rates.[82] To complement incidence data, we also described changes in the prevalence of lifestyle-related modifiable risk factors in young-adult (age 21-49 years) populations. We used data from the National Health and Nutrition Examination Survey (NHANES), Behavioral Risk Factor Surveillance System (BRFSS), and National Health Information Survey (NHIS) to determine the extent to which life-style related modifiable risk factors (obesity, physical inactivity, and current smoking) have paralleled to the increase of young-onset CRC.

Study Population

We identified all patients diagnosed with stages I-IV CRC in 1973-2012 from SEER 9 registries. SEER 9 registries include Atlanta, Connecticut, Detroit, Hawaii, Iowa, New Mexico, San Francisco—Oakland, Seattle—Puget Sound, and Utah, approximately 9.5% of the U.S. population. They are the oldest of the SEER registries, which allowed us to examine incidence rates over the longest period of time. Populations covered by SEER are comparable to the general U.S. population with regard to race, measures of poverty, and education. We excluded

patients with a race/ethnicity other than non-Hispanic white or black from the study population in order to obtain more stable estimates for the CCREM approach. There were 378,000 patients diagnosed with stages I-IV CRC in the SEER 9 registries from 1973-2012, of which we anticipated 30,000 cases would be among those age 21-49 years (Table 3.1). We derived obesity prevalence from NHANES phases I (1971-1975), II (1976-1980), and III (1988-1994) and the continuous cycles (1999-2012); physical inactivity prevalence from BRFSS in 1984-2010; and the prevalence of current smoking from NHIS in 1974, 1976-1980, 1983, 1985, 1987-1988, 1990-1995, and 1997-2012 (Table 1). NHANES contains anthropometric measures of height and weight from population-based samples of U.S. adults. BRFSS annually collects data in all 50 states via telephone surveys regarding health-related risk behaviors, chronic health conditions, and the use of preventive services. Widely considered the gold standard in tracking health status, the NHIS collects data on a broad range of health topics through personal household interviews. Obesity, physical inactivity, and current smoking were calculated specific to four race-sex subgroups (white males, white females, black males, black females), ages 21-49 years, using survey data.

Covariate Assessment

SEER defines CRC anatomically as being located in the colon (ascending, transverse, descending, hepatic flexure, splenic flexure, sigmoid colon; International Classification of Disease for Oncology, Third Edition [ICD-O-3] codes 18.2-18.9) or rectum (rectum, rectosigmoid junction; ICD-O-3 codes 19.9, 20.9). For the purposes of this study, we considered adenocarcinoma histology (morphology codes 814_3, 84803, 84903, 82103, 82203, 82613), which represents more than 98% of all CRCs. SEER program registries routinely collect data on patient demographics, primary tumor site, tumor morphology and stage at diagnosis, and the NCI guides all registries to achieve data content and compatibility acceptable for pooling data.

The SEER Program is viewed as the standard for quality among cancer registries with > 98% case ascertainment.

Obesity

The prevalence of obesity was derived from NHANES phases I (1971-1975), II (1976-1980), and III (1988-1994) and the continuous cycles (1999-2012). NHANES is the only study that provides estimates on anthropometric measures for the U.S. population and has been used by many researchers for more than 30 years. Data collection for NHANES consists of an initial household interview, an in-home personal interview, and a standardized physical examination in a specially equipped mobile examination center. The exam includes physical measurements, a dental examination, and the collection of blood and urine samples for laboratory testing. All NHANES participants are eligible for the body measures components of the exam. The complete set of measurements includes weight, height, upper leg and arm length, mid-arm and waist circumferences, and triceps and subscapular skinfolds. Obesity is measured as body mass index (BMI) ≥ 30 kg/ m^2. A trained health technologist and recorder work together to collect the anthropometry data. Examinations are conducted in mobile exam centers, which helps to standardize their administration. All data collection procedures are overseen by the Centers for Disease Control and Prevention (CDC).

Physical Inactivity

Physical inactivity prevalence was derived from BRFSS survey data from years 1984-2010. Survey data are publicly available and have been reported in numerous publications. Assessing physical activity over time is difficult because guideline-based recommendations have changed in the U.S. Prior to 1985, recommendations called for vigorous intensity exercise for at least 3 times/week for at least 20 minutes.[83] A new recommendation[84] was issued in 1995 by the CDC and American College of Sports Medicine that allowed for 30 minutes of moderate

intensity physical activity, which could be accumulated over the course of the day, almost daily. The recommendation was meant to provide a choice for less intense kinds of activities, and it heightened awareness of what could be considered physical activity (e.g., gardening, yardwork). Although our understanding of vigorous activity has changed over time, the definition of physical *in*activity has remained relatively constant. BRFSS provides a consistent measure of physical inactivity across all survey years, defined as no leisure-time activity or exercise in the last 30 days. Data from BRFSS are preferable to other national surveys (e.g., NHIS), where physical activity and exercise have been assessed with a variety of questions and in different survey years, precluding meaningful comparisons over time. BRFSS data collection procedures are overseen by the CDC. Surveys are administrated continuously through the year by state health departments via random digit dialing. Starting in 2011, BRFSS changed weighting methodology and included cell phone only respondents in the sampling frame; data from 2011 forward are not directly comparable to previous years of BRFSS data. Therefore, we used 2010 data as the estimate for the prevalence of physical activity during the corresponding 2010-2012 period.

Current Smoking

We obtained the prevalence of current smoking from the NHIS in years 1974, 1976-1980, 1983, 1985, 1987-1988, 1990-1995, and 1997-2012. NHIS is regarded as the primary data source for tracking progress toward achieving national health objectives. Adult tobacco use was first measured on the NHIS as part of the 1965 smoking supplement and has been included annually in the adult health behaviors section of NHIS since 1997. The basic NHIS cigarette smoking questions consist of a screener question, “Have you ever smoked at least 100 cigarettes in your entire life?” and follow-up questions about current smoking practices. Although follow-up questions have varied slightly across survey years (e.g., from “Do you smoke now?” to “Do you smoke every day, some days, or not at all?”), the same smoking status variable recodes are included in the public use data files. Participants are categorized as current

(including both every day smoker and some day smoker), former, or never smokers. NHIS data are collected through personal household interviews conducted by interviewers employed and trained by the U.S. Census Bureau. The CDC oversees all data collection procedures.

Statistical Analysis

We specified Poisson CCREM using multilevel data described above. CCREM are a type of hierarchical age-period-cohort (HAPC) models that estimate independent age effects (i.e., distribution of the outcome across the life course due to aging), period effects (i.e., secular trends in the prevalence of an outcome that occur in all ages), and cohort effects (i.e., variation in the outcome among those born in or around the same year). HAPC models avoid the identification problem[85] of linear APC regression models (i.e., age, period, and cohort are perfectly collinear) because the three effects (age, period, cohort) are not assumed to be linear and additive at the same level of analysis.[86,87] By pooling SEER incidence data, we created a rectangular age by period array where columns correspond to age-specific incidence rates in each period, and rows are age-specific incidence rates across all periods (Figure 3.1). Linking the diagonal cells of the array gives incidence rates that belong to individuals born in the same calendar year and age together (i.e., birth cohort). Although only a longitudinal panel study design provides data from true birth cohorts that follow identical individuals over time, this design allowed for a synthetic cohort approach, used often in demography research.[85] The synthetic cohort approach has the advantage of simultaneously testing age and period effects because it is based on nationally representative data (e.g., SEER) collected regularly from one period to the next and covers several decades.

HAPC-CCREM is a member of the class of linear mixed models, and is the most widely used form of hierarchical or multivariable linear models.[88] Linear mixed models consist of two components: the level 1 component is a regression of an individual-level outcome variable on a set of individual level covariates (i.e., age) with an intercept term, fixed regression slope

coefficients, and an individual-level random error term. The level 2 component uses level 1 regression coefficients as outcomes and contains intercepts and specification of random effect coefficients for the effects of each cohort and time period distinguished in the model.[81]

Level 1 component: $\log(y_{ijk}) = \beta_{0jk} + \beta_1 AGE_{ijk} + \beta_2 AGE^{2i}_{jk} + \beta_3 SEX_{ijk} + \beta_4 RACE_{ijk} + e_{ijk}$

with $e_{ijk} \sim N(0, \sigma^2)$

Level 2 component: $\beta_{0jk} = \Upsilon_0 + c_{0j} + p_{0k}$

with $c_{0j} \sim N(0, T_c)$, $p_{0k} \sim N(0, T_k)$

Combined model: $\log(y_{ijk}) = \Upsilon_0 + \beta_1 AGE_{ijk} + \beta_2 AGE^{2i}_{jk} + \beta_3 SEX_{ijk} + \beta_4 RACE_{ijk} + c_{0j} + p_{0k} + e_{ijk}$

where: $i = 1,2,\ldots,n_{jk}$ incidence rates within cohort j and period k;

$j = 1,\ldots,13$ birth cohorts;

$k = 1,\ldots,9$ time periods

The combined model is defined by the statistical parameters: the regression coefficients, Υ_0, β_1, β_2, β_3, and β_4, and the variance components, σ^2, T_c, and T_k. Regression coefficients are interpreted similarly as in Poisson regression, where a one unit change in the predictor variable corresponds to the expected change in the difference in the logs of expected counts by the respective regression coefficient, given all other variables in the model are held constant. Variance components are interpreted as the period or cohort effect, or the contribution of cohort *j* and period *k* averaged over all periods and cohorts, respectively. We used restricted maximum likelihood (REML) to estimate parameters. REML is preferable to other methods (e.g., maximum likelihood (ML)), particularly for estimating variance, because it takes into account the loss of degrees of freedom that result from first estimating regression parameters. ML estimators for variance components have a downward bias, but REML estimators do not.

To test the hypothesis that lifestyle-related modifiable risk factors have paralleled the increasing incidence of young-onset CRC, we pooled period-obesity, -physical inactivity, and –

current smoking prevalence across all ages for four sex-race groups (white males, white females, black males, black females) for corresponding NHANES, BRFSS, and NHIS periods of data collection, respectively. Prevalence by cohort was derived by averaging age-specific prevalence for the same four sex-race groups across the relevant time periods. Because no NHANES data were available in certain years, we interpolated obesity prevalence by averaging the prevalence for the previous and subsequent periods and/or cohorts. Physical activity was not measured in BRFSS prior to 1984; therefore, the analysis of physical inactivity was limited to years 1984-2012. This approach allowed us to quantify the extent to which these modifiable risk factors parallel the increasing incidence of young-onset CRC.

Sensitivity Analysis

To account for the possibility that incidence of young-onset CRC increases linearly with BMI (vs. threshold effect of BMI ≥30), we also calculated period- and cohort-specific mean and median BMI.

Sample Size and Power

The primary purpose of *Specific Aim 2* was to determine the contribution of age, time period, and birth cohort on the increasing incidence of young-onset CRC using CCREM specification of HAPC models. CCREM have only recently emerged as a type of APC analysis, and there is no established or formal method for calculating statistical power. However, the CCREM approach is an improvement over linear APC models because it allows researchers to identify key explanatory factors in addition to age, period, and cohort indicators. We estimated there would be 378,000 total patients diagnosed with stages I-IV CRC in SEER 9 registries from 1973-2012, with approximately 30,000 cases in patients age <50 years (Table 3.1). Therefore, we felt the sample size was sufficiently large to detect meaningful associations.

C. SPECIFIC AIM 3 APPROACH

Study Design

In *Specific Aim 3*, we characterized the patterns of colonoscopy use in younger adults by analyzing administrative claims of employer-insured adults in the MarketScan Commercial Claims and Encounters data (MarketScan; Truven Health Analytics, Ann Arbor, MI). Using a cohort design based on extant data allowed us to efficiently obtain a large sample size and address our research question. Because current guidelines do not recommend routine colonoscopy for individuals age <50 years, the prevalence of colonoscopy use in a younger population was likely to be low (~5%). A claims-based analysis yielded a large enough cohort to determine use of a relatively uncommon procedure. Further, colonoscopy is accurately reported in claims data because it is an expensive procedure. Although follow-up information was available at the study outset, MarketScan collected all claims prospectively

Study Population

The study population was derived from individuals with health care claims in MarketScan, a large employer-based claims database. MarketScan contains more than 500 million claim records on person-specific clinical utilization, expenditures, and enrollment across inpatient and outpatient services from approximately 100 payors (Table 3.1). We identified individuals aged 18-49 years during the period 2001-2013 (i.e., the most recent release of MarketScan data) using electronic administrative records. There is no diagnosis code or administrative algorithm to determine the presence of inherited cancer syndromes (e.g., Lynch syndrome). As a result, we could not reasonably exclude younger adults who are recommended to undergo colonoscopy (i.e., those at higher risk or with a family history), but we anticipated the proportion of these patients to be minimal and constant across study years.[13,14] Patients represented in MarketScan data are all insured through their employers; therefore, findings

cannot be generalized to uninsured patients or those without access to endoscopy services. However, MarketScan is preferred to other claims data because data are fully integrated (inpatient, outpatient, enrollment) at the patient-level, represent all ages (i.e., compared to Medicare claims which only cover age ≥ 65), and are not limited to a single payer or geographic region.

Covariate Assessment

Any colonoscopy delivered in the outpatient setting was identified from administrative claims in the outpatient services table in MarketScan. Colonoscopy use was identified with International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) procedure codes and Common Procedural Terminology (CPT) codes. Several validation studies of administrative claims data support the use ICD-9-CM and CPT codes for estimating colonoscopy utilization, with specificity of 72-76% and sensitivity of 70-83%. Higher sensitivity is considered the most important measure of validity when describing ascertainment of a procedure in administrative data.[89] El-Serag and colleagues[90] developed an algorithm and applied it to several decades of national Veterans Health Administration (VHA) administrative data to evaluate trends in colonoscopy utilization, which is widely used in health services research of endoscopy procedures. The algorithm was later modified to include parameters (e.g., other administrative data such as ICD-9 diagnosis codes) to determine why a colonoscopy was performed, with 83% sensitivity and 76% specificity for colonoscopy indication.[91] ICD-9-CM and CPT codes used to identify colonoscopy are listed in Table 3.4.

Statistical Analysis

To test the hypothesis that colonoscopy use has increased over time, we estimated the rate of colonoscopy in each calendar year (2001-2013). We summed the total number of months that individuals aged 18-49 years were enrolled in their insurance plan. For example,

patients who were enrolled in their insurance plan for 8 months in the year 2001 (January 1, 2001 to December 31, 2001) contributed 8 months of “enrollee-time” during that year. Patient enrollee-time was counted in each calendar year without regard to enrollment in previous or subsequent years (i.e., patients could contribute enrollee-time across all calendar years.) Using standardized denominators of enrollee-time is a reasonable approach because our primary goal was to estimate change in colonoscopy use by year.

We also examined differences in the rate of colonoscopy use (overall and in calendar study year) by sex (male, female), age (18-29 years, 30-39 years, 40-49 years), and geographic region (northeast, north central, west, south). We did not examine the indication for colonoscopy because our primary interest was in understanding CRC incidence trends (i.e., a colonoscopy would detect cancer regardless of why it was performed).

Sample Size and Power

The primary purpose of *Specific Aim 3* was to determine increases in the use of colonoscopy across study years (2001-2013). Given the expected low prevalence of colonoscopy in patients age <50 years, we assumed that a minimum of 1% absolute difference per year was meaningful to detect. We set the level of significance at 5% for a 2-sided test of all hypotheses. The estimated sample size of 7,000,000 continuously-enrolled adults ages 18-49 years (see Table 3.1) provided >99% power to detect a1% increase in colonoscopy use per year.

Table 3.1: Data source, description, and preliminary data by study aim

Aim	Data Source, years	Description	Preliminary Data
Aim 1	Patterns of Care (POC), 1990-2010	Detailed tumor and treatment information from a random sample of patients residing in SEER registry areas	7,000 total patients diagnosed with stages II-III CRC; approx. 1,400 cases among patients <50 years
Aim 2	SEER 9, 1973-2012	Cancer incidence and survival from regional registries; routinely collects data on patient demographics, primary tumor, and tumor type and stage	378,000 total patients with stages I-IV CRC; approx. 30,000 cases among patients <50 years
	NHANES, phases I (1971-1975), II (1976-1980), and III (1988-1994) and continuous cycles (1999-2012)	Anthropometric measures of height and weight from population-based samples of adults	Approx. 5,000 survey participants in each year
	BRFSS, 1984-2010	Responses to telephone survey questions on recent physical activity and exercise from population-based samples of adults	Approx. 100,000 survey participants in each year
	NHIS, 1974, 1976-1980, 1983, 1985, 1987-1988, 1990-1995, 1997-2012	Responses to personal household interview questions on current smoking habits from population-based samples of adults	Approx. 87,500 survey participants in each year
Aim 3	MarketScan Commercial Claims and Encounters (Truven Health Analytics), 2001-2013	Large employer-based commercial claims database from approx. 100 payers	Approx. 7 million continuously-enrolled adults age 18-49 years in each year

Abbreviations: SEER: Surveillance, Epidemiology, and End Results; NHANES, National Health and Nutrition Examination Survey; BRFSS, Behavioral Risk Factor Surveillance System; NHIS, National Health Interview Survey

Table 3.2: Summary of covariates measured in Specific Aim 1

Type	Covariates	Source	Time Measured
Demographic	Age, sex, race/ethnicity, insurance, geographic region	SEER	Diagnosis
Clinicopathologic	Tumor site, stage, histology, grade, *BRAF* mutation, *KRAS* mutation, MSI	POC; patient medical record	Cancer-directed surgery
Treatment	Surgery type and pathology, radiation therapy, chemotherapy agents and regimens	POC; patient medical record	4 months post-treatment
Socioeconomic indicators	Hospital type, hospital size, teaching hospital, trial participation	POC	Cancer-directed surgery
	Socioeconomic status composite index (measured in quintiles)	SEER	N/A
	Per capita income, median household income, education level, poverty, unemployment, total number of physicians, total number of gastroenterologists	AHRF	N/A

Abbreviations: SEER, Surveillance, Epidemiology, and End Results; MSI, microsatellite instability; POC, Patterns of Care; AHRF, Area Health Resource File

Table 3.3: Minimum statistically-detectable absolute differences in proportions

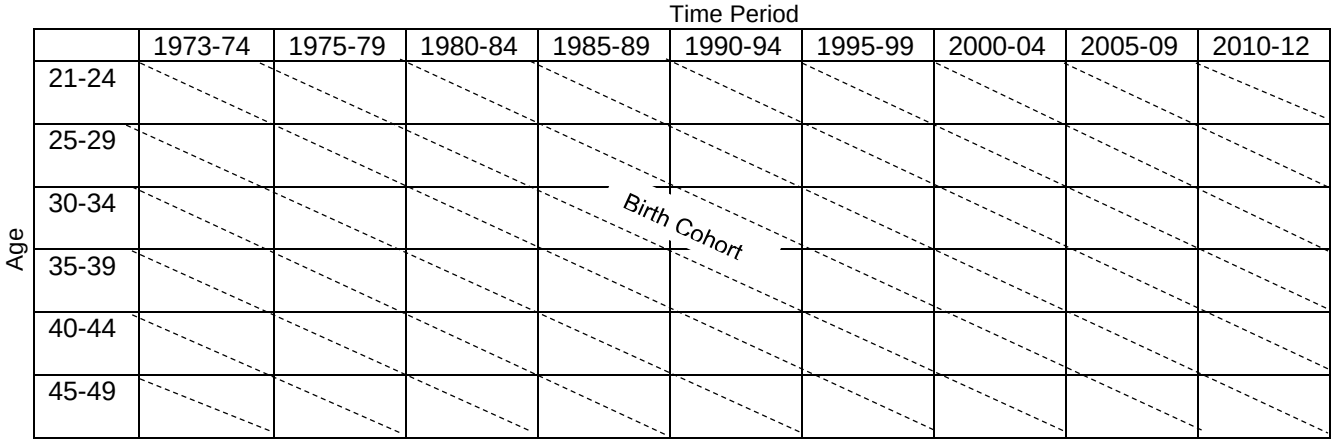
	Prevalence of Sociodemographic Characteristics and Clinicopathologic Features in Older (≥50 years) CRC Patients (n=5,600)								
	10%	15%	20%	25%	30%	35%	40%	45%	50%
Minimum Detectable Absolute Difference in Younger (<50 years) Patients (n=1,400)	4.3%	5.0%	5.5%	6.0%	6.3%	6.5%	6.7%	6.8%	6.8%

Table 3.4: ICD-9 diagnosis codes for colonoscopy indication and ICD-9 procedure and CPT codes for colonoscopy use

	ICD-9 Diagnosis Codes	ICD-9 Procedure Codes	CPT Codes
Colonoscopy Indication/Use	7873 (abdominal distention), 7890 (abdominal pain), 7893 (abdominal swelling), 2851, 2859 (anemia), 7830 (anorexia), 5609 (bowel obstruction), 7879 (change bowel habits), 5640 (constipation), 5589, 5645 (diarrhea), 7876 (fecal incontinence), 578 (GI bleed), 7921 (heme-positive stool), 5693 (hemorrhage rectum anus), 280 (iron-deficiency anemia), 7870 (nausea vomiting), 7832 (weight loss), 555 (Crohn disease), 556 (ulcerative colitis), 56212 (diverticulitis), 56213 (diverticulosis), 56985 (angiodysplasia), 5581 (colitis-radiation), 5600 (intussusception), 5601 (paralytic ileus), 5647 (megacolon), 5641 (irritable colon)	4521, 4522, 4523, 4524, 4525, 4542, 4543, 4823, 4824, 4836	44388, 44389, 44390, 44391, 44392, 44393, 44394, 44395, 44396, 44397, 45355, 45378, 45379, 45380, 45381, 45382, 45383, 45384, 45385, 45386, 45387, 45394, 45397, G0105, G0121, 0105, 0121

Abbreviations: ICD-9, International Classification of Diseases, Ninth Edition; CPT, Current Procedural Terminology

Figure 3.1: Age-by-time period data structure and synthetic birth cohorts



CHAPTER 4
RESULTS: CLINICOPATHOLOGIC AND SOCIOECONOMIC CHARACTERISTICS OF YOUNGER PATIENTS WITH STAGES II AND III COLORECTAL CANCER[1]

A. INTRODUCTION

The overall incidence and mortality of colorectal cancer (CRC) have declined over the past two decades in the U.S.,[8] but recent research suggests the incidence of CRC in younger adults (age <50 years) is rapidly increasing. Since the 1970s, incidence rates have increased annually by up to 4% for rectal cancer and 2% for colon cancer in younger populations, with the steepest increases in the 40-49 year age group.[1-6,68] If current trends persist, by 2030, approximately 11% and 23% of all colon and rectal cancers, respectively, may be diagnosed in patients younger than the current screening age.[68] Because guidelines recommend that screening (with fecal occult blood testing, colonoscopy, or sigmoidoscopy) begin at age 50 for average risk individuals,[10] early recognition of CRC at younger ages is challenging unless there is a known family history or genetic predisposition (i.e., familial adenomatous polyposis, Lynch syndrome). This is especially concerning because more than half of young patients report no family history of CRC, and only a small minority have hereditary cancer syndromes.[12-14]

As evidence of increases in incidence has mounted, researchers have also become interested in CRC in younger adults because some evidence suggests there are differences in CRC survival by age. In a pooled analysis of data from clinical trials of metastatic CRC, younger

[1] At the time of this writing, the results in this chapter are under review at *Cancer Epidemiology.* Coauthors include Drs. Robert Sandler, John Baron, Jennifer Lund, Hanna Sanoff, and Karyn Stittzenberg.

(age <50 years) patients had worse progression-free and overall survival compared to patients of middle age (approximately age 57 years), despite equivalent cancer stage and treatment.[92] This may suggest younger patients are a higher risk population with more aggressive disease. However, mechanisms contributing to reasons for increases in incidence and differential mortality rates are poorly understood.

Little is known about the relative important of sociodemographic and clinicopathologic features in the development of young-onset CRC. The literature on characteristics of younger CRC patients has been limited in several ways. Most studies are of patients treated in a single institution or rely on small sample sizes. Few of the same variables have been examined across studies, and even when the same characteristics were studied, they were defined or measured differently. Inclusion criteria are not consistent (e.g., some studies define young-onset CRC as <40 years while others use <50 years) and differences in stage at diagnosis (i.e., younger patients tend to be diagnosed at later stages) make comparisons difficult. As a consequence, we know very little about the relative importance of sociodemographic and clinicopathologic features in the development of young-onset CRC. Findings from previous research warrant further study in large, population-based samples.

Understanding differences in the characteristics of CRC patients by age may provide insight on the underlying factors that have contributed to the growing incidence of CRC in younger adults. The purpose of this study was to describe demographic, socioeconomic, and clinicopathologic characteristics of younger and older CRC in a population-based sample of stages II and III CRC patients. Given the poorer outcomes frequently observed in younger patients, we also described the receipt of treatment by age.

B. METHODS

Study Population

The study population was derived from the National Cancer Institute's (NCI) Patterns of Care (POC) studies. The NCI annually conducts POC studies on a random sample of patients with select cancers to complement data routinely collected through the Surveillance, Epidemiology, and End Results (SEER) program of cancer registries. Because chemotherapy is not reported in SEER, POC studies provide important information on the extent to which therapies are delivered in community settings. Stages II and III CRC patients in participating SEER registries were eligible for POC studies in 1990, 1991, 1995, 2000, 2005, and 2010.[93] Patients were stratified by registry, sex, age, and race/ethnicity, and a random sample was taken from within each strata. There was oversampling by race/ethnicity in 1995, 2000, 2005, and 2010 to obtain more stable estimates. Patients were ineligible for POC studies if they were younger than age 20, previously diagnosed with cancer (excluding non-melanoma skin cancer), diagnosed at autopsy or on death certificate only, or diagnosed with a synchronous cancer. For purposes of this analysis, we further excluded patients with tumors in the appendix (n=4), who did not undergo cancer-directed surgery (n=171), or with incomplete information to determine TNM staging (n=18).

Covariates

We examined the demographic, clinicopathologic, and socioeconomic characteristics of younger (age <50 years) and older (age 50-64 years, age ≥70 years) CRC patients. We also described treatment patterns (e.g., receipt of chemotherapy) in the study population.

Patient demographics included age (<50, 50-69, ≥70 years), sex, race/ethnicity (non-Hispanic white, non-Hispanic black, Hispanic, other), and insurance (private, Medicare only, any Medicaid, none).

Clinicopathologic features included tumor site (right, left, and sigmoid colon, rectum), stage at diagnosis, histologic grade (well/moderately differentiated, poorly/undifferentiated), mucinous or signet ring cell histology, and receipt of microsatellite instability (MSI) testing. Tumor site included right colon (cecum, ascending colon, hepatic flexure, transverse colon), left colon (splenic flexure, descending colon), sigmoid colon, and rectum (rectosigmoid junction, rectum) according to the International Classification of Disease for Oncology, 3rd edition (ICD-O-3). Data on MSI were collected in 2010 only.

Socioeconomic indicators were derived from POC, SEER, and the Area Health Resource File (AHRF). POC contains patient-level data on hospital type (private, government, nonprofit), an approved residency training program, total bed size, and cancer clinical trial enrollment. We also used a composite census-tract index of socioeconomic status based on measures developed by Yost et al.,[79] including occupation, unemployment, poverty, education, income, and housing. The index was constructed to assess the relationship between socioeconomic status and cancer incidence using SEER data, as described elsewhere.[80] Data used in the index were derived from Census 2000 and American Community Survey 2005-2009 and reflect the populations and census tracts covered by the SEER 17 registries. The index (measured in quintiles) was available for study years 2000, 2005, and 2010. In addition, study data were linked with the AHRF, an extensive county-level database of socioeconomic indicators maintained by the U.S. Department of Health and Human Services. We used AHRF data on per capita income, median household income, education level (% of persons age ≥25 years with less than a high school diploma, high school or more, or four or more years of college), poverty (% of persons living below poverty line), unemployment (unemployment rate), total number of active physicians, and total number of gastroenterologists. Cutpoints for all AHRF variables were based on approximate tertiles. Income measures were adjusted to 2010 dollars.

Treatment patterns included type of surgery (partial, subtotal, total colectomy), laparoscopic surgery, number of lymph nodes examined (0, 1-11, ≥12), positive margins, receipt of chemotherapy, and receipt of radiation therapy (among rectal cancer patients only). As part of POC studies, treatment information was abstracted from medical records and verified by treating physicians. Treating physicians were also asked to provide names and addresses of other physicians who may have treated the patient, who are then contacted for additional treatment details. Doctor verification substantially improves completeness of treatment ascertainment.

Statistical Analysis

Descriptive statistics (e.g., proportions, means) were used to examine the distribution of covariates by age at diagnosis (<50 years, 50-69 years, ≥70 years). To account for potential differences by race/ethnicity, we conducted a stratified analysis of select covariates in the subgroup of younger (age <50 years) non-Hispanic white (n=317), non-Hispanic black (n=200), and Hispanic (n=189) patients. We also examined the proportion of patients who received chemotherapy and radiation therapy by tumor site (colon vs. rectum), stage at diagnosis, and age. Patients who received therapy or were recommended but it was unknown whether they received therapy were considered to have received therapy (n=91); patients who refused therapy (n=221) were not considered to have received therapy.

Sensitivity analyses that considered different categorizations of age at diagnosis (e.g., <50 years, 50-64 years, ≥65 years) did not appreciably change the results; therefore, we report the results of the primary analysis only.

Proportions and means were weighted with stratum-specific sample weights to reflect the population (i.e., SEER) from which the sample was drawn. Sample weights were calculated as the inverse of the sampling proportion for each sampling stratum.

All statistical analyses were conducted using SAS version 9.3 (SAS Institute, Cary, NC).

C. RESULTS

A total of 6,862 stages II and III CRC patients were included in the analysis. Characteristics of the study population by age at diagnosis are shown in Table 4.1. Younger patients were more likely to be black (13%) and Hispanic (15%) than patients aged 50-69 years (11% and 10%, respectively) and aged ≥70 years (7% for each). More (58%) younger patients were diagnosed with stage III (vs. II) CRC. Tumor site varied considerably with age. In younger patients, 37% of tumors were located within the rectum and 22% in the right colon, whereas in patients over age 70 years, only 18% of tumors were within the rectum 48% were right-sided. A similar proportion of patients aged 50-69 years had tumors in the rectum (31%) and right colon (32%). The proportion of tumors in the left colon was similar (15%) across all strata of age.

In the stratified analysis of younger patients by race/ethnicity (Table 4.2), more whites had private insurance (86%) compared to both blacks (61%) and Hispanics (68%). There were also differences by tumor site. A larger proportion of young white (41%) and Hispanic (33%) patients had rectal tumors, whereas tumors in the right colon were the most common in young black patients (39%). Although the proportion of tumors classified as high vs. low grade was similar by race/ethnicity, a higher proportion of blacks had tumors with mucinous histology (18%) compared to whites (10%) and Hispanics (12%).

Differences in county-level socioeconomic indicators by age at diagnosis are shown in Table 4.3. Fewer (15%) young patients lived in areas with lower median household (<$50,000) compared to the two older groups of patients (age 50-69 years: 23%; age ≥60 years: 25%). A higher proportion of the oldest (age ≥70 years) patients lived in counties with lower poverty (<10% living below poverty line), lower unemployment rates (<5%), and higher education. There

was no difference in the total number of physicians or gastroenterologists per 100,000 persons by age.

The proportion of patients who received chemotherapy differed by age at diagnosis (Table 4.4). Among stage II colon cancer patients, the proportion of patients who received chemotherapy decreased with increasing age (age <50 years: 71%; age 50-69 years: 43%; age ≥70 years: 14%). A larger proportion of stage III colon cancer patients age <50 years (82%) and age 50-69 years (82%) received chemotherapy than did patients age ≥70 years (49%). A similar pattern was observed in stages II and III rectal cancer, with the vast majority of young patients receiving chemotherapy (stage II: 82%; stage III: 94%). The proportion of rectal cancer patients who received radiation therapy decreased with increasing age in both stage II (age <50 years: 72%; age 50-69 years: 61%; age ≥70 years: 43%) and III (age <50 years: 77%; age 50-69 years: 70%; age ≥70 years: 42%). More young patients received MSI testing (age <50 years: 28%; age 50-69 years: 10%; age ≥70 years: 7%) and had more lymph nodes (≥12) examined at surgery (age <50 years: 72%; age 50-69 years: 60%; age ≥70 years: 53%) (Table 4.1).

D. DISCUSSION

In contrast to the decline in CRC incidence in the overall population,[8] incidence rates in younger populations are rising quickly.[1-6,68] Using a population-based sample, we found important differences in the distribution of young-onset CRC by race/ethnicity. Younger CRC patients were considerably more likely to be black or Hispanic. Consistent with other studies,[2,5,17,30] we observed a larger proportion of black patients were diagnosed at a younger age compared to whites. SEER data from 2012 similarly shows the overall incidence of CRC (in all age groups) is higher among blacks (46.8/100,000) than whites (38.1/100,000),[94] with a disproportionate number of cases diagnosed among younger black patients.

We also found a higher proportion of Hispanic patients were diagnosed at younger ages than whites. Due to a variety of concerns, including misclassification and cultural or other differences among Hispanic and Latino groups, there has historically been limited information on cancer trends in Hispanic populations.[95] Hispanics represent the fastest-growing and youngest minority group in the U.S.,[96] and their inclusion in cancer statistics has become increasingly relevant. More recent efforts to describe cancer incidence in diverse populations show the overall incidence rates of CRC are lower in Hispanics than non-Hispanic whites,[97] although there may be some differences in incidence by country of origin (e.g., higher incidence rates are observed in Cuban Americans).[98] Another study of young-onset CRC in the California Cancer Registry[99] found that, although the absolute rate of CRC in all age groups was lowest in Hispanics, the largest relative increases in incidence during the study period (1988-2009) were for distal colon and rectal cancers among young Hispanic males and females.

Considerable evidence suggests there are distinct tumor subtypes of CRC,[100-102] often defined by combinations of deficient DNA mismatch repair or MSI, *MLH1* methylation (i.e., CpG island methylator phenotype), and mutations in *BRAF* and *KRAS* oncogenes. These subtypes are also clinically meaningful because of the prognostic and treatment implications (e.g., patients with *KRAS* mutations are unlikely to benefit from anti-EGFR therapies). As our understanding of CRC subtypes has improved, an emerging body of literature shows there are differences in tumor subtypes across racial groups. For example, in a recent study of *BRAF* and *KRAS* mutations among patients treated with FOLFOX-based chemotherapy in the Alliance N0147 trial,[103] *KRAS* mutation was more common in black patients, while the frequency of *BRAF* mutation was highest in tumors from whites. Other studies[104,105] have found that, among patients with microsatellite-stable or –low tumors, blacks have a higher frequency of *KRAS* mutations compared to whites. This difference was most pronounced in the proximal colon, with no differences in mutation frequency by race in the distal colon or rectum. In our study, right-

sided tumors predominated (39%) in younger black patients, while young white (41%) and Hispanic (33%) patients had a higher proportion of tumors located within the rectum. Combined with the growing evidence on tumor subtypes, our finding of racial differences in the age and tumor site distribution of CRC make a compelling argument for distinct mechanisms driving CRC progression in various population subgroups.

A higher proportion of younger patients (both stages II and III) in our study received "optimal" care, including better nodal counts from surgery, treatment at academic medical centers, clinical trial enrollment, chemotherapy, and radiation therapy, compared to the two older groups of patients. Even in the setting of stage II colon cancer, where the absolute benefit of chemotherapy is very small, the vast majority of young patients (71%) received adjuvant therapy. Many of the younger stage II patients treated with chemotherapy had high risk features, including T4 tumors (16% vs. 9%), low-grade histology (22% vs. 19%), and inadequately sampled (<12) lymph nodes (33% vs. 28%), compared to younger patients who did not receive therapy (data not shown). Despite more aggressive treatment, some research suggests younger CRC patients have a worse prognosis than older patients of the same stage, or overall survival is similar between the two groups.[18,21,23,24,26,33,34,46] A recent study[47] of colon cancer patients in the National Cancer Data Base found no difference in the relative survival of younger and older patients, even though younger patients more frequently received chemotherapy. Our findings and the results of others may reflect physician preferences to ambitiously treat disease simply based on younger age at presentation.

A strength of our study is the population-based sample. Data from POC studies offer a number of unique advantages for conducting population-based epidemiologic research because each participating registry area has a defined population, and detailed tumor and treatment information was abstracted from patient medical records and verified by treating physicians. Previous research has been limited to single institution studies or relied on small sample sized,

in which results generally reflect the distribution of characteristics of patients treated at that institution rather than true differences in younger and older CRC patients. POC data also provide a greater breadth and depth of information than that available solely from medical claims and/or SEER registries. This was particularly true for our assessment of receipt of chemotherapy and radiation therapy, where doctor verification substantially improved the completeness of treatment ascertainment. Further, the large size and ethnic diversity of this study population were strengths that enabled us to examine CRC characteristics within population subgroups by race/ethnicity.

Our study population was limited to stages II and III patients, and there may be different characteristics of younger CRC patients when considering early-stage or metastatic disease. For example, we observed only slight differences in county-level socioeconomic indicators among younger and older patients, but a relationship between CRC and socioeconomic status has been demonstrated most consistently in late-stage disease.[106-109] The increase in the number of younger patients diagnosed with stages II and III CRC in our study may also be a reflection of stage migration (i.e., some cases once considered stage I would now be classified as stage II); however, evidence has consistently shown meaningful increases in all stages of young-onset CRC.[68] In addition, we did not have information on genetic predisposition to CRC, either by hereditary syndrome or a first-degree relative with a history of CRC. These data may have helped explain changes in the distribution of young-onset CRC over time, but the prevalence of these syndromes in younger populations remains very low.[13] Finally, there were few patients who received MSI testing, which precluded our ability to report the prevalence of molecular markers in the study population. Although collection efforts have likely improved since 2010, and SEER now includes site-specific factors on MSI and *KRAS* mutation, our study highlights the continued need for robust sources of molecular data at the population-level.

The increasing incidence of CRC in younger populations—and the poorer prognosis of young patients—may reflect differences in risk factors for the development of CRC rather than disparities in detection and treatment. Results of this study demonstrate important differences in the distribution of young-onset CRC by race/ethnicity. These differences may be attributable to lifestyle-related risk factors, such as a higher prevalence of obesity, physical activity levels, and cultural differences in diet. Our findings support the effort to understand CRC as a heterogeneous disease based on age, race/ethnicity, and molecular markers.

Table 4.1: Characteristics of 6,862 patients diagnosed with stages II and III colorectal cancer, 1990-2010, by age at diagnosis

	Age <50 Years (n=871)		Age 50-69 Years (n=3,018)		Age ≥70 Years (n=2,973)	
	n	weighted[1] %	n	weighted[1] %	n	weighted[1] %
Demographic Characteristics						
Sex						
Male	472	54.0	1685	56.0	1377	44.3
Female	399	46.0	1333	44.0	1596	55.7
Race/ethnicity						
White, non-Hispanic	317	61.4	1490	71.5	1810	81.0
Black, non-Hispanic	200	13.3	634	11.2	449	6.7
Hispanic	189	14.7	495	9.7	386	6.7
Other	165	10.5	399	7.6	328	5.6
Insurance						
Private/HMO/VA/Other	637	78.9	2214	76.8	1848	67.2
Medicare (only)	127	1.0	351	9.1	685	23.0
Medicaid (any)	13	14.4	254	10.6	358	9.2
None	75	5.8	128	3.6	21	0.6
Clinicopathologic Features						
Tumor Site						
Right Colon	176	22.4	679	31.9	951	46.4
Left Colon	116	15.6	332	15.1	379	15.4
Sigmoid Colon	137	25.1	536	22.2	482	20.2
Rectum[2]	433	36.9	1456	30.8	1140	18.0
Stage at Diagnosis						
Stage II	360	42.2	1437	48.7	1532	54.2
Stage III	511	57.8	1581	51.3	1441	45.8
Histologic Grade						
Well/moderately differentiated	640	78.7	2352	82.2	2216	73.4
Poorly/undifferentiated	186	21.3	526	17.8	641	26.6
Mucinous Adenocarcinoma	110	11.1	293	9.5	321	11.5
Signet Ring Cell Carcinoma	12	0.8	28	0.7	29	1.2
MSI Testing Performed[3]						
No	140	72.7	540	90.3	423	92.9
Yes	68	27.3	83	9.7	44	7.1
Treatment						
Surgery Type						
Local Excision	6	0.7	16	0.5	16	0.4
Partial Colectomy	480	56.4	1619	50.5	1466	41.1
Subtotal Colectomy	240	30.7	839	38.4	1080	50.6
Total Colectomy	98	9.1	346	6.5	249	3.8
Total Proctocolectomy	38	2.4	163	3.4	139	3.5
Other	9	0.7	35	0.7	23	0.6
Laparoscopic Surgery[4]						

No	583	77.2	1763	78.0	1646	82.7
Yes	130	22.8	380	22.0	283	17.3
Lymph Nodes Examined						
0	24	1.8	89	2.0	73	1.7
1-11	258	25.8	1302	37.8	1499	45.0
≥12	560	72.4	1537	60.2	1297	53.3
Positive Margins						
No	772	94.5	2672	93.5	2558	95.0
Yes	47	5.5	133	6.5	146	5.0
Received Chemotherapy						
No	169	18.5	943	32.5	1848	66.2
Yes	683	81.5	2029	67.5	1061	33.8
Received Radiation Therapy[5]						
No	112	25.0	494	34.1	699	57.4
Yes	321	75.0	962	65.9	441	42.6
Socioeconomic Indicators (Census Tract-Level)						
SES Index[6]						
Quintile 1	122	15.8	378	18.5	322	17.2
Quintile 2	113	21.4	350	20.3	279	20.6
Quintile 3	109	19.5	322	24.7	259	17.5
Quintile 4	113	20.8	251	16.0	276	22.2
Quintile 5	104	22.6	275	20.6	238	22.6
Socioeconomic Indicators (Patient-Level)						
Hospital Type						
Private	67	11.9	238	7.2	247	10.1
Government	182	16.1	540	17.4	387	13.5
Nonprofit	621	72.0	2212	75.4	2313	76.4
Teaching Hospital						
No	370	45.2	1397	51.9	1564	53.9
Yes	498	54.8	1585	48.1	1374	46.1
Total Bed Size						
<200	194	26.1	676	25.1	710	27.2
200-400	336	39.3	1330	43.2	1335	43.6
≥400	339	34.6	982	31.6	895	29.2
Clinical Trial Enrollment						
No	728	91.7	2494	94.4	2583	98.2
Yes	67	8.3	179	5.6	74	1.8

NOTE: Table does not include patients who did not undergo cancer-directed surgery (n=167) or with incomplete staging information (n=35); no variable had more than 10% missing data, missing observations range from 35 (tumor site) to 737 (clinical trial enrollment)
Abbreviations: VA, Veterans Affairs; MSI, microsatellite instability; SES, socioeconomic status
[1] Proportions weighted by sampling fraction
[2]Rectal includes both rectum and rectosigmoid junction
[3]Microsatellite instability collected in 2010 only
[4]Laproscopic colectomy collected in 1995, 2000, 2005, and 2010
[5]Receipt of radiation therapy limited to rectal cancer patients (n=3,029); includes postoperative and preoperative radiation

[6]Socioeconomic status based on composite census-tract level indicators from Census 2000 and American Community Survey 2005-2009; limited to data collection years 2000, 2005, and 2010 and does not include Louisiana

Table 4.2: Characteristics of 706 younger (age <50 years) patients diagnosed with stages II and III colorectal cancer, 1990-2010, by race/ethnicity

	NH White (n=317)		NH Black (n=200)		Hispanic (n=189)	
	n	weighted[1] %	n	weighted[1] %	n	weighted[1] %
Sex						
Male	184	54.8	109	52.1	94	52.7
Female	133	45.2	91	47.9	95	47.3
Insurance[2]						
Private/HMO/VA/Other	275	86.1	129	60.7	114	67.7
Medicaid (any)	22	10.8	48	26.1	35	19.6
None	11	3.0	16	10.5	29	9.8
Year of Diagnosis						
1990/91	114	16.5	22	13.6	5	1.4
1995	42	8.8	38	10.1	36	6.5
2000	27	9.4	28	11.8	32	13.3
2005	75	33.5	59	30.7	50	27.6
2010	59	31.9	53	33.8	66	51.2
Tumor Site						
Right Colon	49	19.4	62	38.5	42	26.3
Left Colon	28	13.8	29	19.9	29	16.4
Sigmoid Colon	44	25.9	29	17.8	33	24.3
Rectum	193	40.9	80	23.8	83	32.9
Mucinous Histology	30	10.4	36	17.7	29	11.8
Histologic Grade						
Well/moderately differentiated	246	80.7	146	81.7	140	79.2
Poorly/undifferentiated	57	19.3	39	18.3	40	20.8
Stage at Diagnosis						
Stage II	118	42.7	86	47.4	87	38.0
Stage III	199	57.3	114	52.6	102	62.0
MSI Testing Performed[3]						
No	36	74.6	32	68.4	42	76.5
Yes	14	25.4	20	31.6	18	23.5
Received Chemotherapy[4]						
No	50	15.2	50	27.8	36	18.8
Yes	261	84.8	147	72.2	147	81.2
Received Radiation Therapy[5]						
No	61	27.9	16	16.8	17	16.7
Yes	132	72.1	64	83.2	66	83.3
SES Index[6]						
Quintile 1	11	7.7	46	37.9	45	28.0
Quintile 2	30	22.5	26	24.2	34	20.9
Quintile 3	26	16.3	18	16.3	30	27.4
Quintile 4	37	21.8	20	16.1	25	17.0
Quintile 5	43	31.7	12	5.4	8	6.7

Abbreviations: NH, non-Hispanic; VA, Veterans Affairs; MSI, microsatellite instability; SES, socioeconomic status

[1] Proportions weighted by sampling fraction
[2]Percentages do not add to 100 because some younger patients were insured through Medicare; cell sizes too small (<5) to report
[3]MSI testing collected in 2010 only
[4]Reciept of chemotherapy includes both neoadjuvant and adjuvant (or both) chemotherapy
[5]Receipt of radiation therapy limited to rectal cancer patients (n=356); includes postoperative and preoperative radiation
[6]Socioeconomic status based on composite census-tract level indicators from Census 2000 and American Community Survey 2005-2009; limited to study years 2000, 2005, and 2010 and does not include Louisiana

Table 4.3: County-level socioeconomic indicators of 6,862 patients diagnosed with stages II and III colorectal cancer, 1990-2010, by age at diagnosis

	Age <50 Years (n=871)		Age 50-69 Years (n=3,018)		Age ≥70 Years (n=2,973)	
	n	weighted[1] %	n	weighted[1] %	n	weighted[1] %
Per Capita Income[2]						
<$32,000	133	15.4	665	23.2	718	24.8
$32,000-50,000	575	66.0	1909	62.8	1818	58.1
>$50,000	163	18.6	444	14.0	437	17.1
Median	39637		38380		38337	
Median Household Income[2]						
<$50,000	176	18.8	736	24.2	691	22.8
$50,000-75,000	558	67.6	1870	63.1	1838	61.9
>$75,000	137	13.6	412	12.7	444	15.3
Median	57755		56931		57745	
% Living in Poverty[3]						
<10	214	32.6	619	30.0	626	36.5
10-19	409	55.5	1165	54.4	1037	52.9
≥20	99	11.9	387	15.6	299	10.7
Median	12.8		12.7		11.9	
% Less than High School						
<10	378	37.3	1471	43.3	1614	48.7
10-19	377	45.9	1155	39.8	1042	36.4
≥20	116	16.8	392	16.9	317	14.8
Median	12.1		11.4		10.2	
Unemployment Rate (%)						
<5	273	30.6	976	32.6	1090	36.5
5-9	437	45.7	1550	43.1	1499	45.2
≥10	161	23.7	492	24.3	384	18.3
Median	5.8		5.7		5.3	
Total Physicians per 100,000						
<200	209	26.7	793	30.6	801	28.4
200-400	473	52.3	1560	48.4	1440	48.6
≥400	189	21.0	665	21.0	732	23.0
Median	268.2		268.3		268.4	
Total Gastroenterologists per 100,000[3]						
<3	201	30.3	625	32.0	581	32.0
3-5	310	41.7	957	42.6	845	38.1
>5	211	28.0	589	25.5	536	29.9
Median	3.9		3.7		3.7	

NOTE: Cutpoints based on approximate tertiles
[1] Proportions weighted by sampling fraction
[2]Adjusted to 2010 dollars
[3]Poverty and total number of gastroenterologists not collected in 1990/91

Table 4.4: Receipt of chemotherapy and radiation therapy among 6,862 patients diagnosed with stages II and III colorectal cancer, 1990-2010, by age and stage at diagnosis

	Stage II (n=3,329)			Stage III (n=3,533)		
	Age <50	Age 50-69	Age ≥70	Age <50	Age 50-69	Age ≥70
	n (wt. %)	n (wt. %)	n (wt. %)	n (wt. %)	n (wt. %)	n (wt. %)
Received Chemotherapy[1]						
Colon Cancer						
No	71 (29.2)	465 (57.4)	809 (86.0)	43 (18.0)	156 (17.9)	439 (50.9)
Yes	121 (70.8)	281 (42.6)	131 (14.0)	180 (82.0)	621 (82.1)	390 (49.1)
Rectal Cancer						
No	28 (17.6)	216 (34.3)	323 (52.3)	24 (6.0)	101 (9.1)	264 (42.3)
Yes	128 (82.4)	444 (65.7)	225 (47.7)	248 (94.0)	674 (90.9)	307 (57.7)
Received Radiation Therapy[2]						
Rectal Cancer						
No	45 (28.5)	241 (39.0)	342 (57.1)	67 (23.3)	253 (30.1)	357 (57.8)
Yes	112 (71.5)	431 (61.0)	214 (42.9)	209 (76.7)	531 (69.9)	227 (42.2)

NOTE: Proportions weighted by sampling fraction
Abbreviations: wt, weighted
[1]Reciept of chemotherapy includes both neoadjuvant and adjuvant (or both) chemotherapy
[2]Receipt of radiation therapy includes postoperative and preoperative radiation

CHAPTER 5
RESULTS: AN AGE-PERIOD-COHORT ANALYSIS OF THE CHANGING INCIDENCE PATTERNS OF COLORECTAL CANCER[2]

A. INTRODUCTION

The overall incidence and mortality of colorectal cancer (CRC) has declined in the U.S. in the past two decades.[8] In 1975, CRC incidence was 56.5 per 100,000 compared to 39.8 per 100,000 in 2012.[94] Incidence rates declined by 3.9% per year among adults age 50 years and older from 2001-2010, with the steepest declines among those over the age of 65. CRC mortality rates similarly declined 3.0% annually in both men and women during the same period.[8] Much of this decrease has been attributed to the use of screening.[9] Regular screening with the fecal occult blood test, sigmoidoscopy, and colonoscopy facilitates earlier detection of CRC and lowers incidence and mortality through removal of premalignant polyps.[11] Current guidelines recommend screening begin at age 50 for those at average-risk of CRC.[10]

Although CRC incidence and mortality rates in the total population have improved, recent evidence suggests the incidence of CRC in younger adults (age <50 years) is rapidly increasing. Incidence rates in younger populations have consistently increased since the 1970s, with annual increases of up to 4% for rectal and 2% for colon cancer.[2-6,17,68] Despite the overall population trends in aging, by 2030, approximately 11% of colon and 23% of rectal cancers may be diagnosed in patients younger than the current screening age.[68]

[2] At the time of this writing, a condensed version of this chapter has been submitted to *JAMA*. Authors include the committee members listed on the title page.

Reasons for the increase in young-onset CRC are poorly understood. Several risk factors for CRC, including unhealthy diet, obesity, physical inactivity, and smoking, have been proposed as the major drivers of the increase in younger adults.[48,49] Meta-analyses demonstrate significant associations between CRC risk and red and processed meat consumption,[50,51] adiposity (including measures of body mass index, waist circumference, and waist-to-hip ratio),[52,54,110] sedentary behavior and physical inactivity,[55] and cigarette smoking.[56] However, the extent to which changes in the prevalence of these risk factors explain increases in in the incidence of young-onset CRC has not been studied.

Given the changing incidence patterns of CRC, in which incidence has decreased in screening-aged populations but increased in younger populations, we used age-period-cohort (APC) analysis to estimate the differential contributions of age, period, and birth cohort to the incidence of CRC. APC analysis combines information on age, time period of observation, and birth cohort to track health outcomes over time. We also described period and cohort trends in the prevalence of CRC risk factors (obesity, physical inactivity, current smoking) in younger populations.

B. METHODS

Data Sources

CRC Incidence. Incidence of CRC was derived from the National Cancer Institute's (NCI) Surveillance, Epidemiology, and End Results (SEER) program. The SEER program routinely collects data on patient demographics, primary tumor site, tumor morphology, and stage for all cancers diagnosed in defined geographic regions. SEER 9 registries consists of Atlanta, Connecticut, Detroit, Hawaii, Iowa, New Mexico, San Francisco—Oakland, Seattle—Puget Sound, and Utah, approximately 9.5% of the U.S. population. CRC was defined anatomically as located in the colon (ascending, transverse, descending, hepatic flexure,

splenic flexure, sigmoid colon; International Classification of Disease for Oncology, Third Edition [ICD-O-3] codes 18.2-18.9) or rectum (rectum, rectosigmoid junction; ICD-O-3 codes 19.9, 20.9). Incidence rates per 100,000 persons within the population were analyzed from 1973-2012 for four race/sex subgroups (white males, white females, black males, black females).

Prevalence of CRC Risk Factors. To complement incidence data, we also described period and cohort trends in the prevalence of common risk factors for CRC among younger adults (age <50 years), including obesity, physical inactivity, and smoking. Obesity prevalence was determined using data from the National Health and Nutrition Examination Survey (NHANES) phases I (1971-1975), II (1976-1980), and III (1988-1994) and the continuous cycles (1999-2012). NHANES is the only study that provides estimates on anthropometric measures for the U.S. population. Data collection procedures include a standardized physical examination, where a trained health technologist collects a complete set of anthropometric measures (e.g., weight, height, upper leg, and arm length) from survey participants. Obesity was measured as a body mass index (BMI) ≥ 30 kg/ m^2.

Prevalence of physical inactivity was derived from the Behavioral Risk Factor Surveillance System (BRFSS) during years 1984-2010. Surveys are administrated continuously through the year by state health departments via random digit dialing. BRFSS provides a consistent measure of physical inactivity across all survey years, defined as no leisure-time activity or exercise in the last 30 days (i.e., “During the past month, did you participate in any physical activities or exercises such as running, calisthenics, golf gardening, or walking for exercise?”). Starting in 2011, BRFSS changed the weighting methodology and included cell phone only respondents in the sampling frame; data from 2011 forward are not directly comparable to previous survey years. Therefore, we used 2010 data as the estimate for the prevalence of physical activity during the corresponding 2010-2012 period.

We obtained the prevalence of current smoking from the National Health Interview Survey (NHIS) in years 1974, 1976-1980, 1983, 1985, 1987-1988, 1990-1995, and 1997-2012. Adult tobacco use was first measured on the NHIS as part of the 1965 smoking supplement and has been included annually in the adult health behaviors section of NHIS since 1997. The basic NHIS cigarette smoking questions consist of a screener question, “Have you ever smoked at least 100 cigarettes in your entire life?” and follow-up questions about current smoking practices. Although follow-up questions have varied slightly across survey years (e.g., from “Do you smoke now?” to “Do you smoke every day, some days, or not at all?”), the same smoking status variable recodes are included in the public use data files. Participants are categorized as current (including both every day smoker and some day smoker), former, or never smokers.

To estimate the prevalence of period-specific risk factors, we pooled prevalence estimates across all ages (age ≥ 21 years) for the four race/sex subgroups (white males, white females, black males, black females) for the corresponding NHANES, BRFSS, and NHIS years of data collection. Similarly, we estimated the prevalence of cohort-specific risk factors by averaging the prevalence in each birth cohort across the same periods of data collection. Because surveys were not conducted in certain years (e.g., NHANES data were not collected in 1981-1987), we interpolated prevalence by averaging the prevalences for the previous and subsequent periods or cohorts. To account for the possibility that CRC incidence increases linearly with BMI (vs. a threshold effect of BMI ≥ 30), we also estimated period- and cohort-specific mean and median BMI using a similar approach.

Statistical Analysis

Two-way graphical displays of age by time period, time period by age, and birth cohort by time period incidence rates were first created to describe trends in incidence across age, period, and cohort. Incidence trends were examined overall and separately in the younger (age <50 years) population.

Independent age (i.e., distribution of the outcome across the life course due to aging), period (i.e., secular trends in the prevalence of an outcome that occur in all ages), and birth cohort effects (i.e., variation in the outcome among those born in or around the same year) can be estimated with a variety of statistical methods. We used Poisson cross-classified random effects modeling (CCREM) specifications of hierarchical age-period-cohort (HAPC) models that account for variations in multi-level data by nesting age-specific incidence rates within periods and birth cohorts. The HAPC models avoid the identification problem[111] of linear APC regression models, in which age, period, and cohort are perfectly collinear, because the three effects are not assumed to be linear and additive at the same level of analysis.[81,87] CCREM allows for the possibility that individuals within the same periods and birth cohorts share some unobserved random variance. Although only a longitudinal panel study design provides data from true birth cohorts that follow identical individuals over time, this design allows for a synthetic cohort approach, used often in demography research.[81] The synthetic cohort approach has the advantage of simultaneously testing age and period effects because it is based on nationally representative data (e.g., SEER) collected regularly from one period to the next and covers several decades. We developed separate models for the overall, younger (age <50 years), and older (age ≥50 years) populations. We also plotted the random effects coefficients from these models to facilitate interpretation of the period and/or cohort effects. Model components and parameters are defined in detail in Chapter 3.

Birth cohort and period were divided into approximate 5-year categories for 21 birth cohorts and 9 periods. Continuous age was centered at the mean to reduce the association between the linear and quadratic age terms. Beta coefficients for the fixed-level effects (i.e., race, sex, and age) can be interpreted as the increase in incidence for every one-unit increase.

All statistical analyses were conducted using SAS version 9.3 (SAS Institute, Cary, North Carolina).

C. RESULTS

Two-way descriptive graphs of age, period, and cohort trends are shown in Figures 5.1-5.6. We created separate graphs of trends for the overall population and for younger ages (<50 years). As shown in Figure 5.1, age-period incidence rates in the overall population increased with increasing age. There were dramatic declines in incidence among older individuals by time period. The highest incidence rates for these older age groups were in 1973-74 and lowest in 2010-12. In younger ages (age <50 years), however, incidence rates were similar in all time periods (Figure 5.2). A similar trend was observed with period-age incidence rates (Figure 5.3). There were consistent declines in incidence for those ages 65 years and older after 1985, with the steepest decreases in the oldest age groups (ages 75-79, 80-85, 85+). In the younger population (Figure 5.4), incidence remained relatively flat over time, although there were small increases in the more recent time periods. Age-birth cohort incidence rates (Figure 5.5) increased in the early birth cohorts (1885-1925) for ages 65 years and older. Starting with approximately the 1920-24 birth cohort, incidence rates generally decreased across successive cohorts. Incidence rates in the younger population (age <50 years) slightly increased by birth cohorts, especially after the 1945-49 birth cohort (Figure 5.6).

Prevalence of CRC risk factors across period and birth cohort are shown for the four race/sex subgroups in Figures 5.7-5.16. For all race/sex subgroups, obesity increased across time periods, from 12.9% (1975-79) to 34.6% (2010-12) (Figure 5.7). Period-obesity prevalence was consistently higher among black females compared to other race/sex subgroups. Similar results were observed for mean (Figure 5.8) and median (Figure 5.9) BMI. Trends in cohort-obesity were more variable, although prevalence generally increased across successive birth cohorts, starting with individuals born around 1950 (Figure 5.10). The lowest prevalence was among white males in the 1935-39 birth cohort (10.8%), and the highest was black females in

the 1980-84 cohort (51.2%). There was a similar pattern for median (Figure 5.11) and mean (Figure 5.12) BMI.

Physical inactivity remained relatively stable over time in all race/sex subgroups combined, from 21.1% in 1985-89 to 20.1% in 2010-12 (Figure 5.13). There were small increases from 1985-89 to 1995-99 (24.6%) and subsequent decreases through the end of the study period. Prevalence was highest among black females in all years. The trend in cohort-physical inactivity was more consistent, with general declines across successive birth cohorts (Figure 5.14). For the majority of race/sex subgroups, the prevalence of cohort-physical inactivity was highest in the 1940-44 birth cohort (white males: 29.7%, white females: 28.1%, black males: 38.7%, black females: 38.8%).

For all race/sex subgroups, the prevalence of period-smoking decreased dramatically over time (Figure 5.15). The overall prevalence of current smoking was 43.5% in 1973-4 compared to 24.8% in 2010-12. Before 1990-94, smoking prevalence was highest among black males (59.6% in 1973-74 to 36.2% in 1990-94), but in more recent years, prevalence was highest among white males (31.3% 1995-99 to 27.2% in 2010-12). There were similar declines in smoking by birth cohort (Figure 5.16). Prevalence decreased across successive birth cohorts for all race/sex subgroups, although it increased slightly among white males in the youngest birth cohorts (1980-84).

Results of the CCREM analysis in the overall population are shown in Table 5.1 (Model A). Incidence was higher among blacks (vs. whites, $\beta= 0.20$, $p <0.001$) and lower among females (vs. males, $\beta= -0.32$, $p <0.001$). There was a significant age effect, where incidence rates increased 11% with each one-year increase in age ($\beta= 0.11$, $p<0.001$). There were also significant period ($\sigma= 0.02$, $p= 0.02$) and birth cohort ($\sigma=0.07$, $p<0.001$) effects. Incidence increased from the period 1973-74 to 1985-89 and then decreased through 2010-12 (Figure 5.17), independent of the birth cohort and age. Similarly, when adjusting for time period and

age, incidence increased slightly from the 1985-89 to the 1900-04 birth cohorts and declined through the 1945-49 cohort, with larger increases again in the most recent cohorts (Figure 5.18). In the stratified analysis by age group, incidence was also higher among blacks and lower among females in the both the younger (age <50 years; Table 5.1, Model B) and older (age ≥50 years; Table 5.1, Model C) populations. Age (β= 0.15, p<0.001) and birth cohort (σ= 0.01, p= 0.04) were statistically significant in the model restricted to age <50 years, but period was not (σ= 0.002, p= 0.09). In the older population, there were independent age (β= 0.10, p<0.001), period (σ= 0.02, p= 0.02), and birth cohort (σ= 0.04, p= 0.004) effects.

D. DISCUSSION

Our study shows large decreases in CRC incidence for older populations across time periods and birth cohorts that have important implications for the role of CRC screening. We observed a dramatic decline in incidence from the earliest (1973-74) to most recent (2010-12) time periods that was limited to the screening-age population. Unlike screening for other cancers (e.g., breast, prostate), CRC screening reduces incidence via polypectomy and removal of premalignant lesions. Population-based screening was first introduced in the late 1980s when the NCI developed working guidelines for the early detection of cancer[112] and was later included in the 2000 Objectives for the Health of the Nation. The use CRC screening has since increased (largely driven by the use of colonoscopy),[66,67,113,114] but the extent to which it explains declines in CRC incidence continues to be debated. Some argue screening accounts for much of the improvement in CRC outcomes, where the benefits of screening observed in clinical trials have led to significant reductions in both incidence and mortality rates.[9,115-117] However, others suggest it has had only a modest impact, and changes in risk factors may be more important than the adoption of screening.[115,118] Our results provide compelling evidence to suggest that reductions in incidence over time, which begin precisely at age 50, cannot be solely attributed to biology or risk factors. Rather, the striking finding that incidence curves diverge sharply in

screening-age populations demonstrate a screening effect in a way that has not been previously described.

Differences in the age distribution of CRC that have been frequently observed, in which the proportion of CRC patients diagnosed at younger ages is increasing, may simply be a consequence of declines in CRC in older adults (i.e., due to screening). The results of our study demonstrate only a small absolute increase in CRC incidence in younger populations. Young-onset CRC has gained the attention of many researchers and media outlets, in part because of the large relative increases (i.e., percentage change) in incidence that have been reported. Presenting increases in incidence rates in exclusively relative terms can have a dramatic effect on their interpretation and may lead to confusion about the magnitude of such increases. Our analyses show an increase of only one or two additional cases per 100,000 persons over a 30-year time period, while others have simultaneously suggested large relative increases of over 100%.[68] Some professional organizations have promoted changing screening guidelines to begin average-risk screening at a younger age.[1] Most recently, the American College of Gastroenterology[119] recommends screening begin at age 45 (vs. 50) for blacks based on relative differences in incidence,[120] and it may not be long before similar recommendations are made in the overall population. Our findings, however, highlight the importance of considering absolute (vs. relative) measures when making public health policy decisions. Relative measures of effect are expressed as a dimensionless ratio and do not capture information on background risk or base rates of health, and as a result, they may overestimate the perception of the extent to which young-onset CRC has increased.

Despite small increases in incidence, we observed a cohort effect in younger populations. Yet, in contrast to the incidence patterns in older adults, there was not a period effect. This finding gives additional support to the hypothesis that there may be different mechanisms involved in the development of CRC across the life course (see Chapter 4). Cohort

effects are evidence in many cancers and chronic diseases because birth cohorts often have different exposures to behavioral and environmental risk factors. When cumulative exposures or exposures during vulnerable ages to risk factors (known or unknown) vary in prevalence from one generation to the next, incidence rates in the population may also vary substantially by birth cohort. For example, the increase in lung cancer incidence and mortality through the 1990s and subsequent decline has been widely attributed to the parallel trends in smoking by birth cohort.[81] In gastrointestinal malignancies, it has been suggested that the decreasing cohort trends in *Helicobacter pylori* colonization,[121,122] which is inversely associated with risk of esophageal adenocarcinoma[123] and a known cause of gastric cancer,[124] may contribute to changes in incidence patterns.[125] Our study underscores the need for a better understanding of the critical factors involved with cohort-mechanisms, such as childhood obesity or changes in microbiota, which may be related to young-onset CRC.

The role of modifiable risk factors in the increase of young-onset CRC is not clear. Population surveys can be helpful in efforts to understand the context of incidence trends, but they do not provide the level of detail needed to identify risk factors for disease. Ecological data describes characteristics of a group, and conclusions cannot be made regarding individuals. Consistent with others,[126] our analysis of NHANES data shows increases in the prevalence of obesity in younger populations across both time periods and birth cohorts. It may be that increases in BMI, especially at younger ages, account for some of the increases in incidence, but the mechanisms involved in obesity and cancer risk have not been fully elucidated. For example, accumulating evidences shows there are differences in the relevant measure of adiposity by race/ethnicity, where waist-hip-ratio may be a better marker of CRC risk in black patients than BMI.[127,128] Second, we were surprised to find the prevalence of physical inactivity (based on BRFSS data) decrease by period and cohort. The relationship between CRC and sedentary behavior and/or physical activity is consistent across multiple studies and in different

settings and populations,[55] but the trend data do not correspond with changes in incidence. This may be a reflection of the challenges that arise when assessing physical activity with population surveys. Guideline recommendations have changed in the U.S. over time,[84] which has subsequently heightened awareness of what can be considered physical activity (e.g., gardening, yardwork). Finally, the results demonstrate overall declines in smoking trends, although there were some modest increases in smoking behavior among white males in the most recent birth cohorts. Large cohort studies suggest the elevated risk for rectal cancer among smokers may drive the association between smoking behavior and overall CRC risk.[129] Cancer registry data similarly show tumors in the rectum (vs. colon) are more common in young, white males.[2] Collectively, findings related to trends in obesity, physical inactivity, and smoking point to other factors that may increase risk of young-onset CRC.

In summary, our results suggest there are differential effects of time period and birth cohort in the changing incidence patterns of CRC. Declines in incidence among older populations can be attributed to screening, which was clearly demonstrated by the sharp break in incidence curves at age 50. The factors responsible for increases in young-onset CRC, albeit small on an absolute scale, remain unknown. It may be that other risk factors not measured through population surveys, such as intestinal microbiota, explain changes in incidence. When considered alongside evidence on the differences in the racial and tumor site distribution of young- and older-onset CRC, this study provides strong support for different etiologies of CRC by age.

Table 5.1: Poisson HAPC-CCREM estimates of incidence rates of colorectal cancer, SEER 9, 1973-2012

	Model A (all ages)			Model B (age <50 years)			Model C (age ≥ 50 years)		
Fixed Effects	Coefficient	SE	*p*	Coefficient	SE	*p*	Coefficient	SE	*p*
Model for Mean									
Intercept	-7.04	0.07	<0.001	-10.08	0.04	<0.001	-7.02	0.071	<0.001
Age (linear, centered)	0.11	<0.001	<0.001	0.15	0.002	<0.001	0.10	<0.001	<0.001
Age (quadratic)	-0.002	<0.001	<0.001	-0.001	<0.001	<0.001	-0.001	<0.001	<0.001
Female Sex	-0.32	0.003	<0.001	-0.12	0.01	<0.001	-0.33	0.003	<0.001
Black Race	0.20	0.01	<0.001	0.37	0.02	<0.001	0.18	0.01	<0.001
Variance Components	Variance	SE	*p*	Variance	SE	*p*	Variance	SE	*p*
Cohort									
Intercept	0.07	0.02	0.001	0.01	0.01	0.04	0.04	0.01	0.004
Period									
Intercept	0.02	0.01	0.02	0.002	0.002	0.09	0.02	0.01	0.02

Abbreviations: HAPC-CCREM, hierarchical age-period-cohort cross classified random effects model; SE, standard error

Figure 5.1. Age by time period incidence rates of colorectal cancer, all ages, SEER 9, 1973-2012

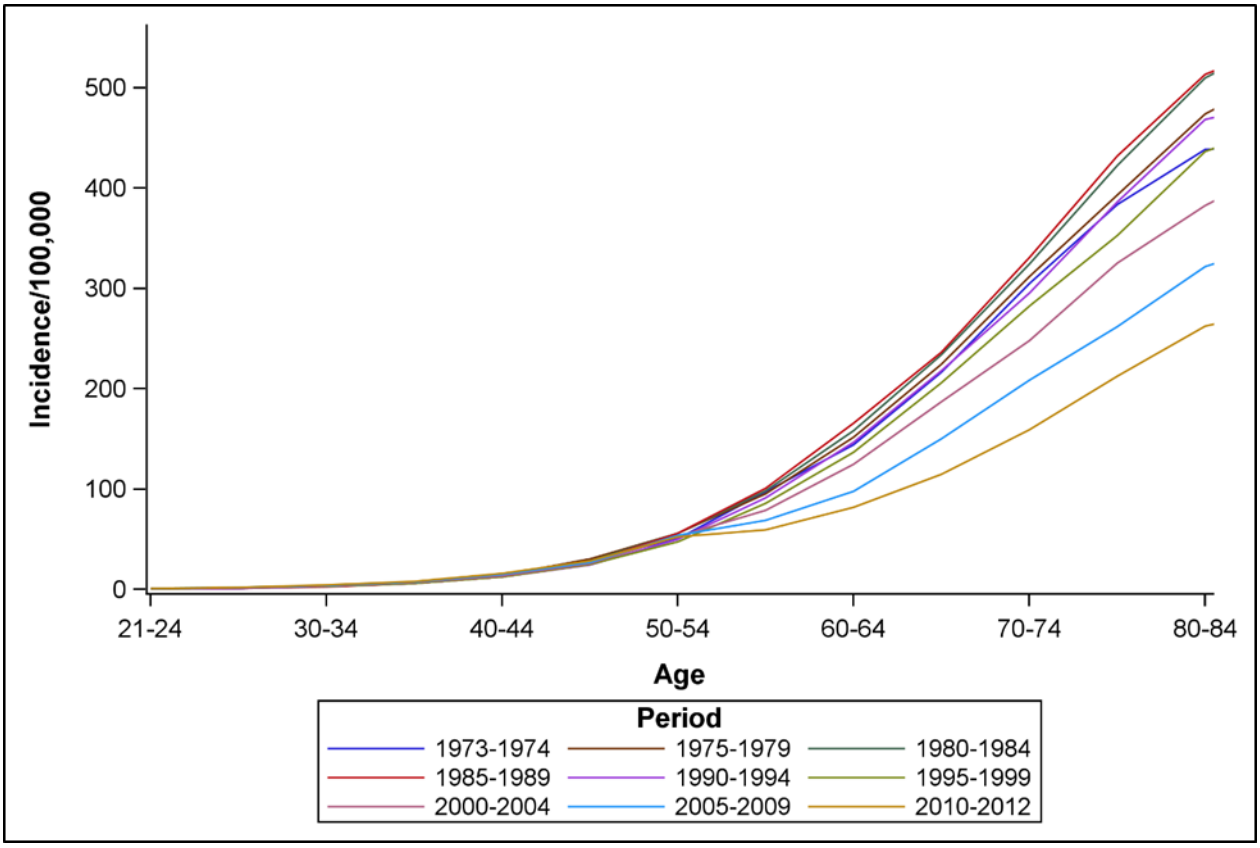


Figure 5.2. Age by time period incidence rates of colorectal cancer, age 21-49 years, SEER 9, 1973-2012

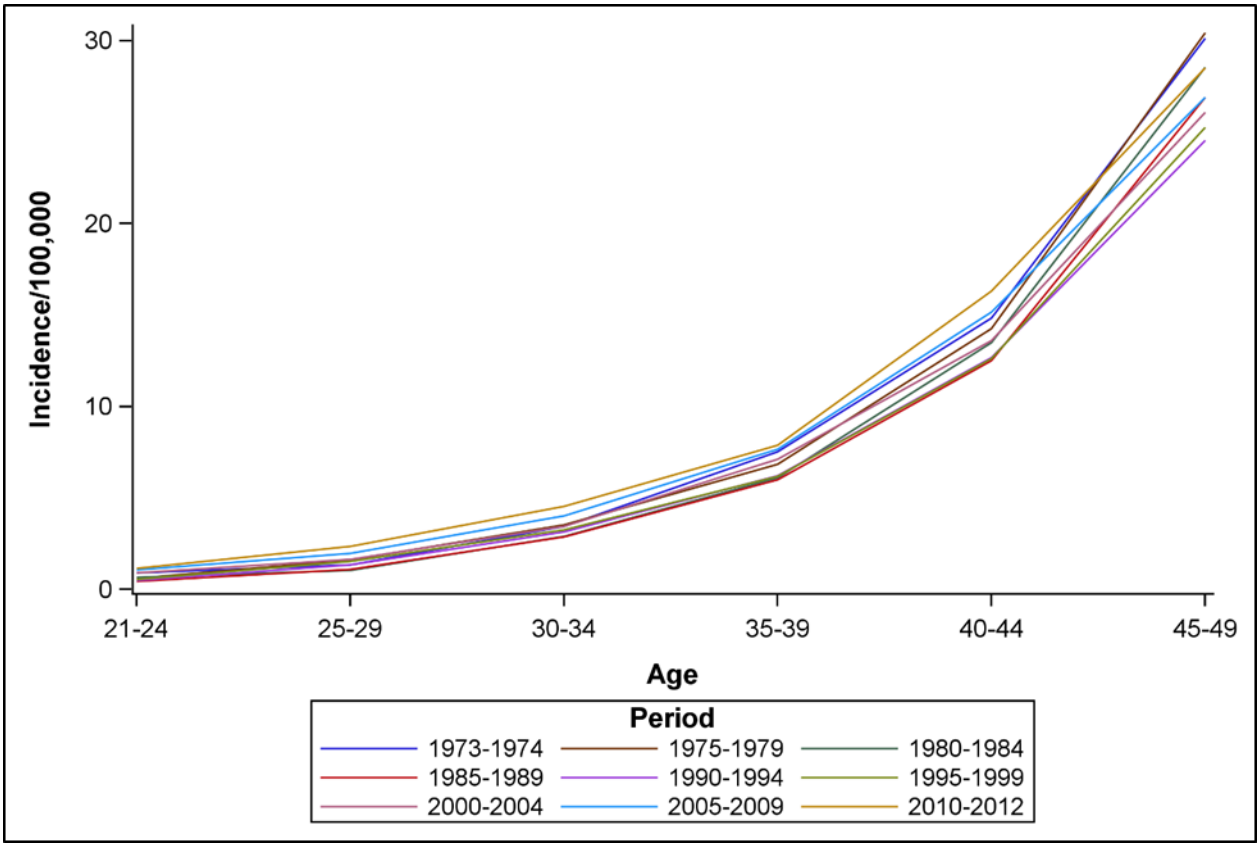


Figure 5.3. Time period by age incidence rates of colorectal cancer, all ages, SEER 9, 1973-2012

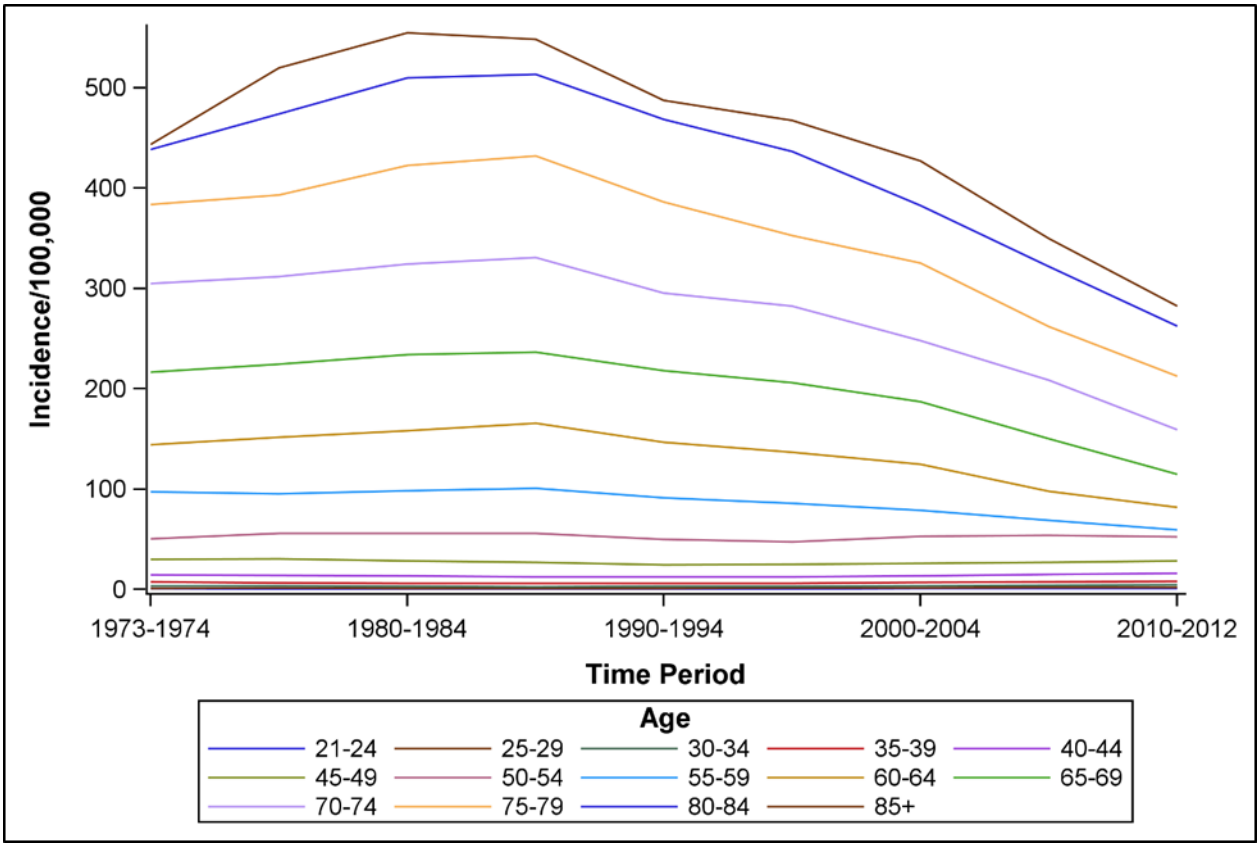


Figure 5.4. Time period by age incidence rates of colorectal cancer, ages 21-49 years, SEER 9, 1973-2012

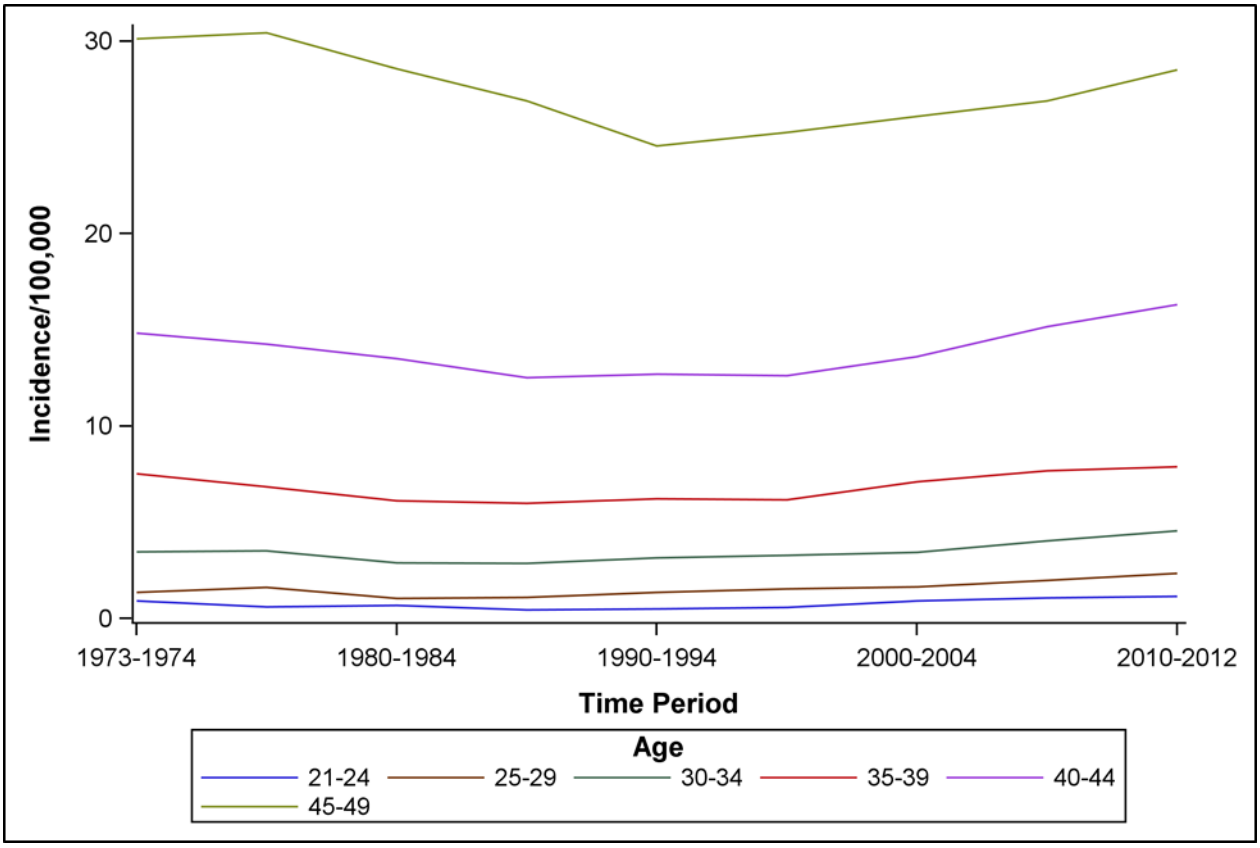


Figure 5.5. Birth cohort by age incidence rates of colorectal cancer, all ages, SEER 9, 1973-2012

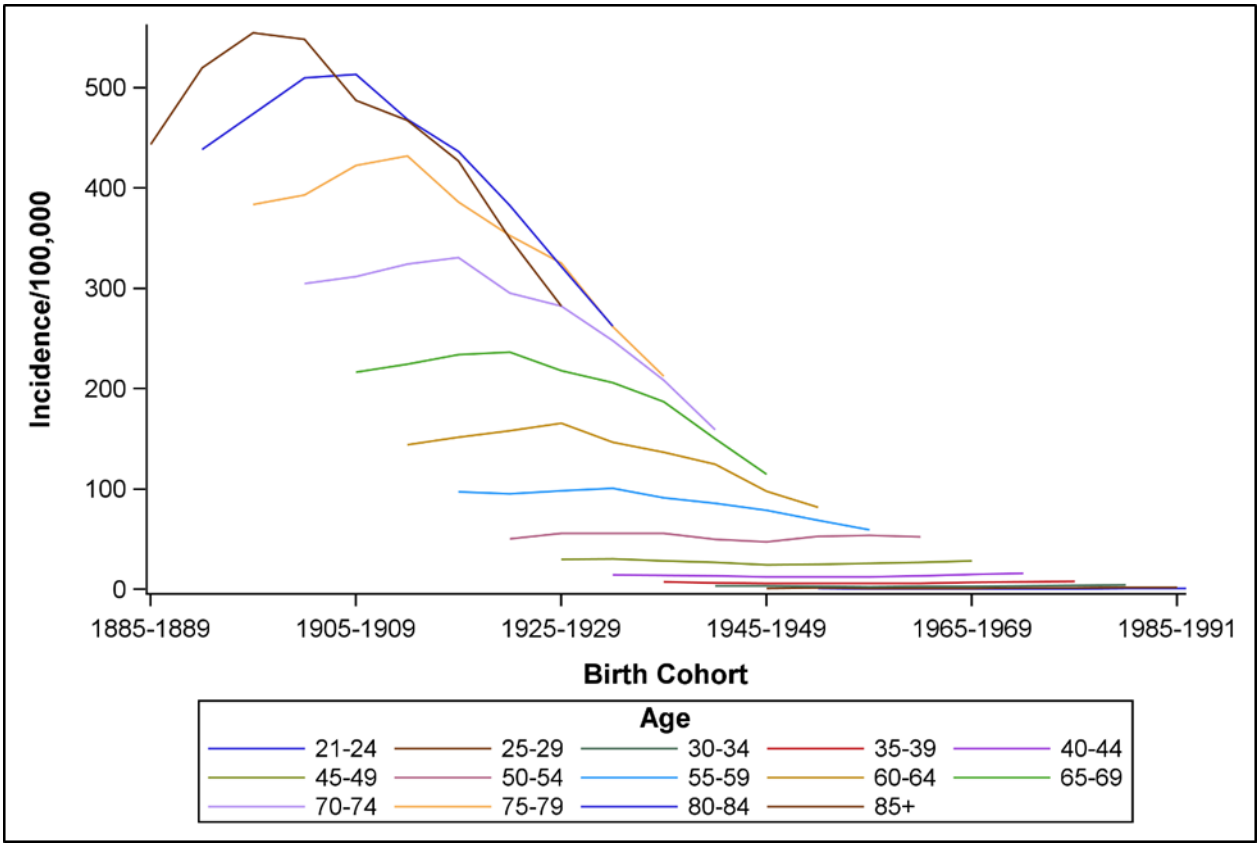


Figure 5.6. Birth cohort by age incidence rates of colorectal cancer, ages 21-49 years, SEER 9, 1973-2012

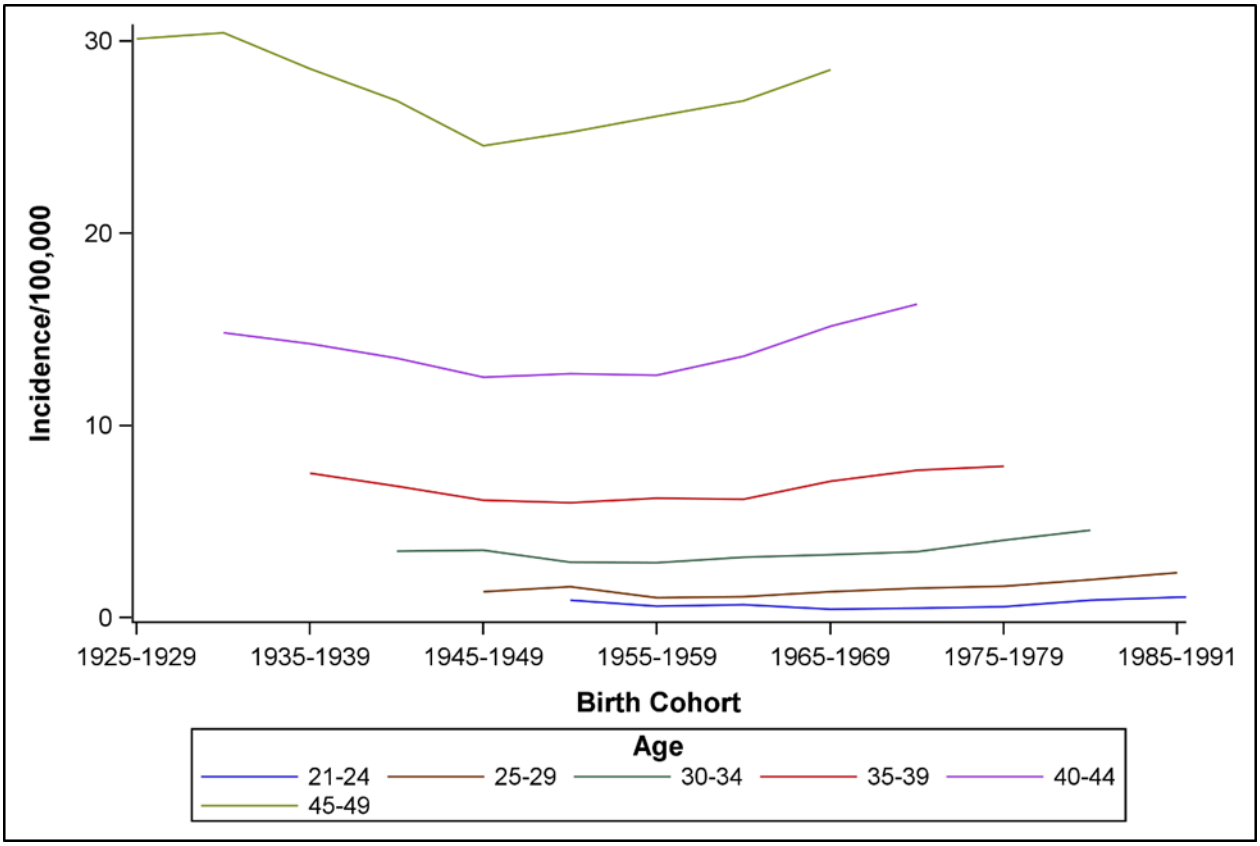


Figure 5.7. Prevalence of obesity by time period for four race/sex subgroups, ages 21-49 years, National Health and Nutrition Examination Survey, 1973-2012

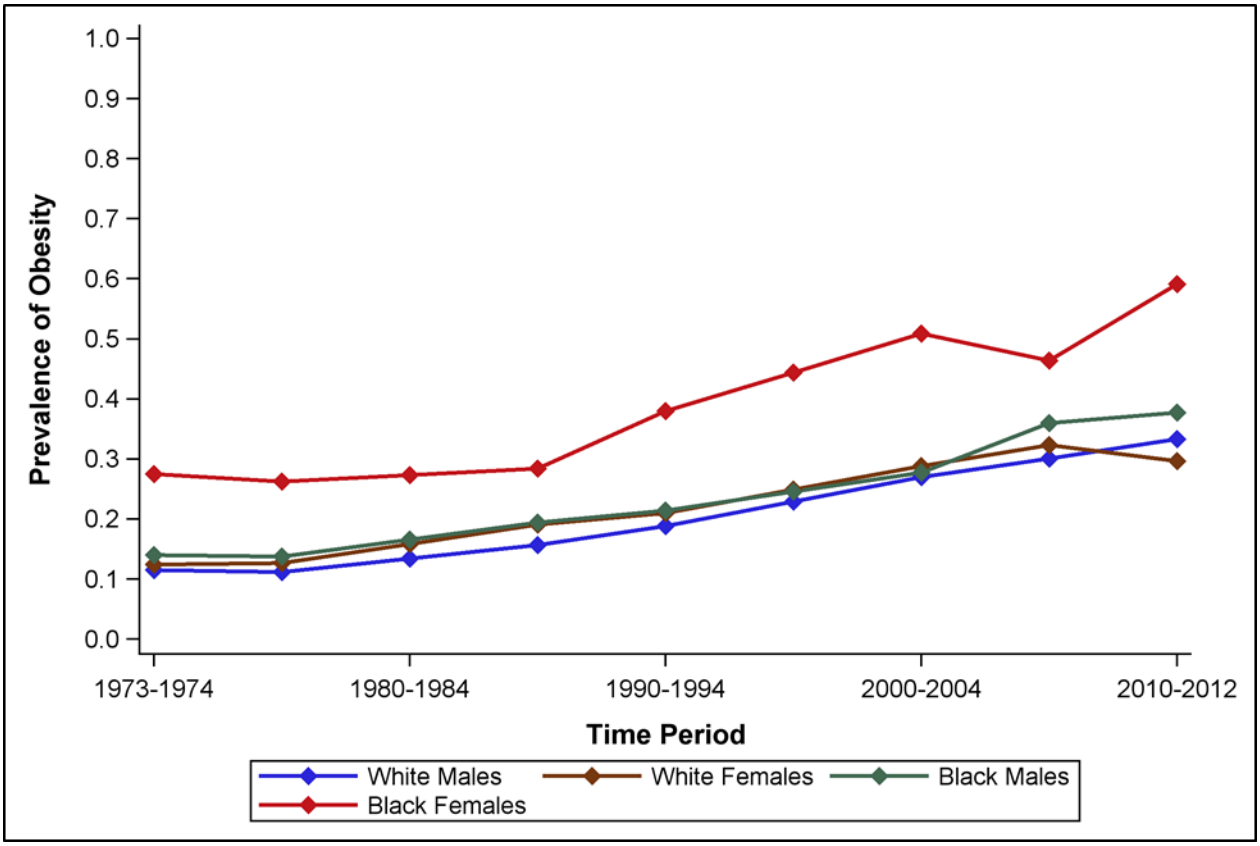


Figure 5.8. Mean body mass index by time period for four race/sex subgroups, ages 21-49 years, National Health and Nutrition Examination Survey, 1973-2012

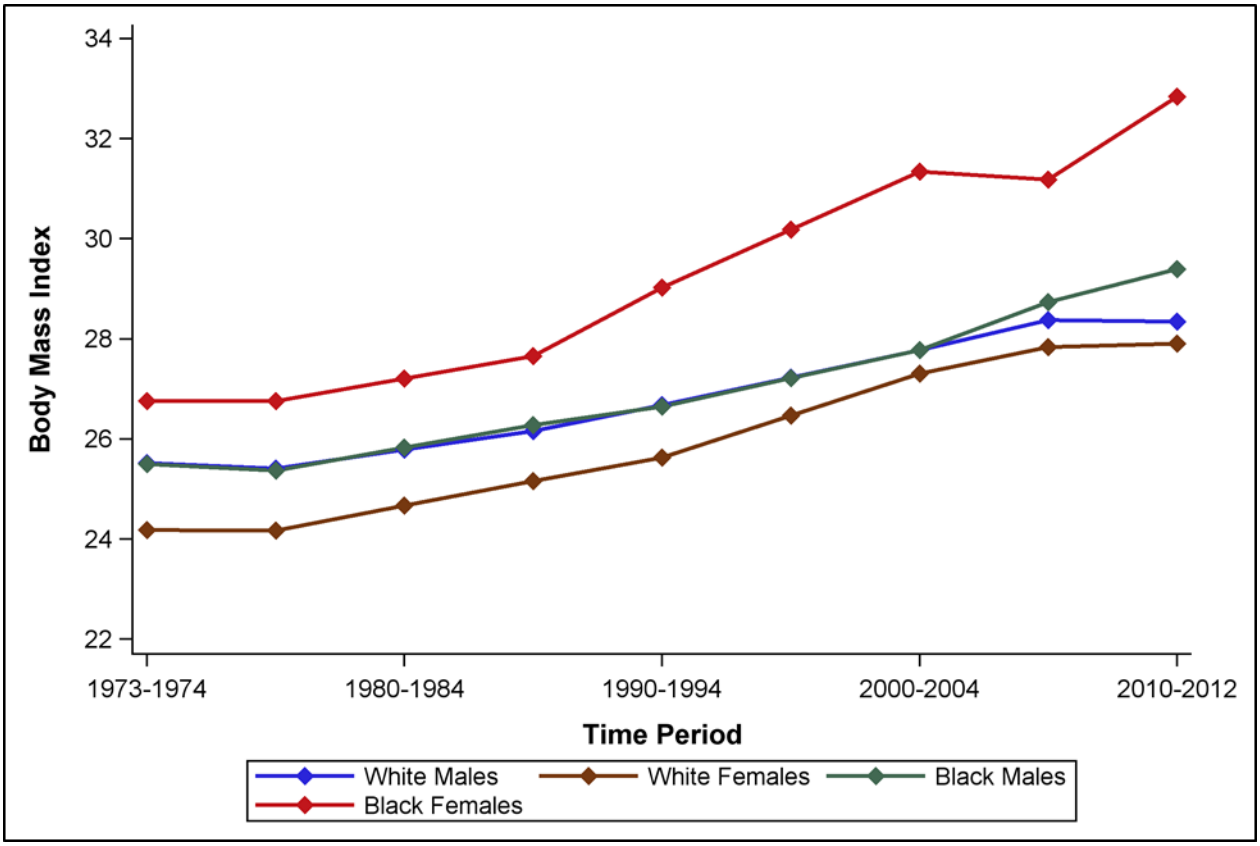


Figure 5.9. Median body mass index by time period for four race/sex subgroups, ages 21-49 years, National Health and Nutrition Examination Survey, 1973-2012

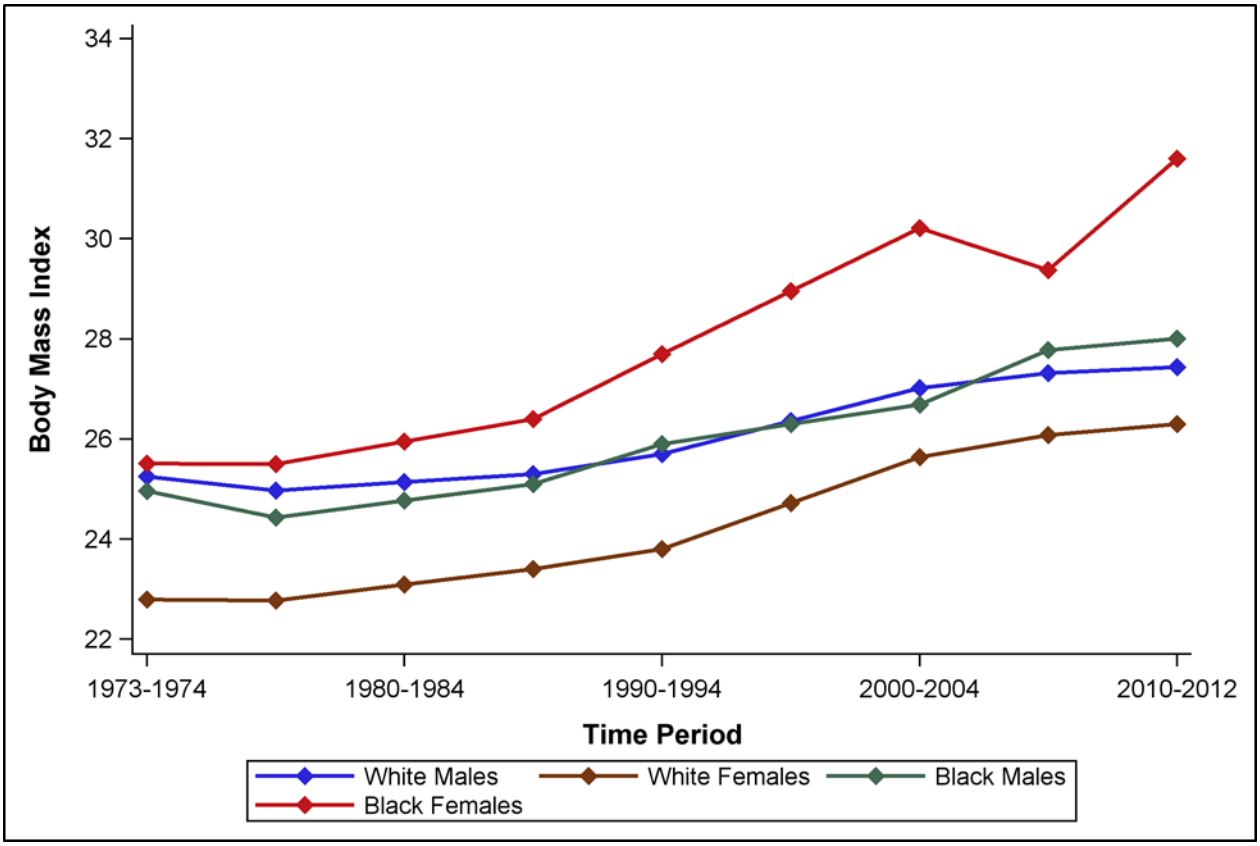


Figure 5.10. Prevalence of obesity by birth cohort for four race/sex subgroups, ages 21-49 years, National Health and Nutrition Examination Survey, 1973-2012

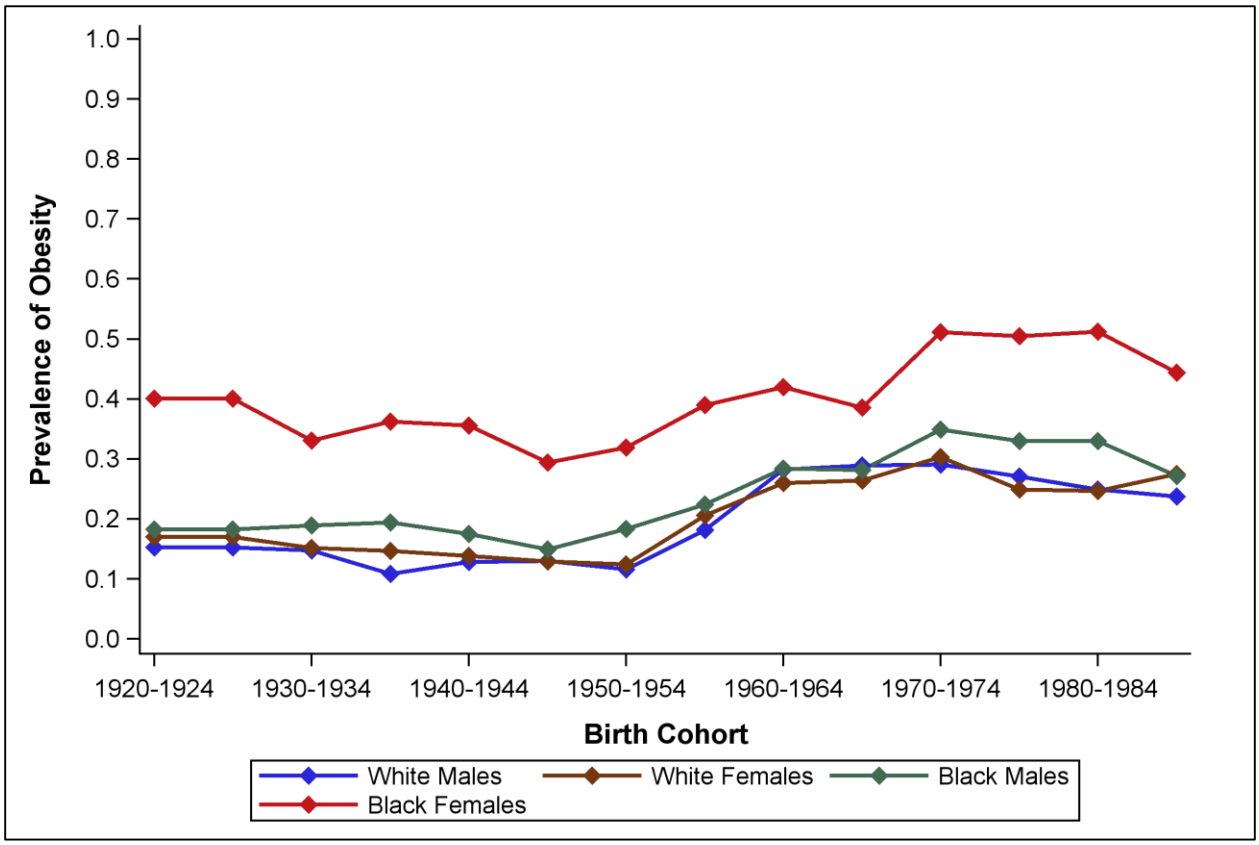


Figure 5.11. Mean body mass index by birth cohort for four race/sex subgroups, ages 21-49 years, National Health and Nutrition Examination Survey, 1973-2012

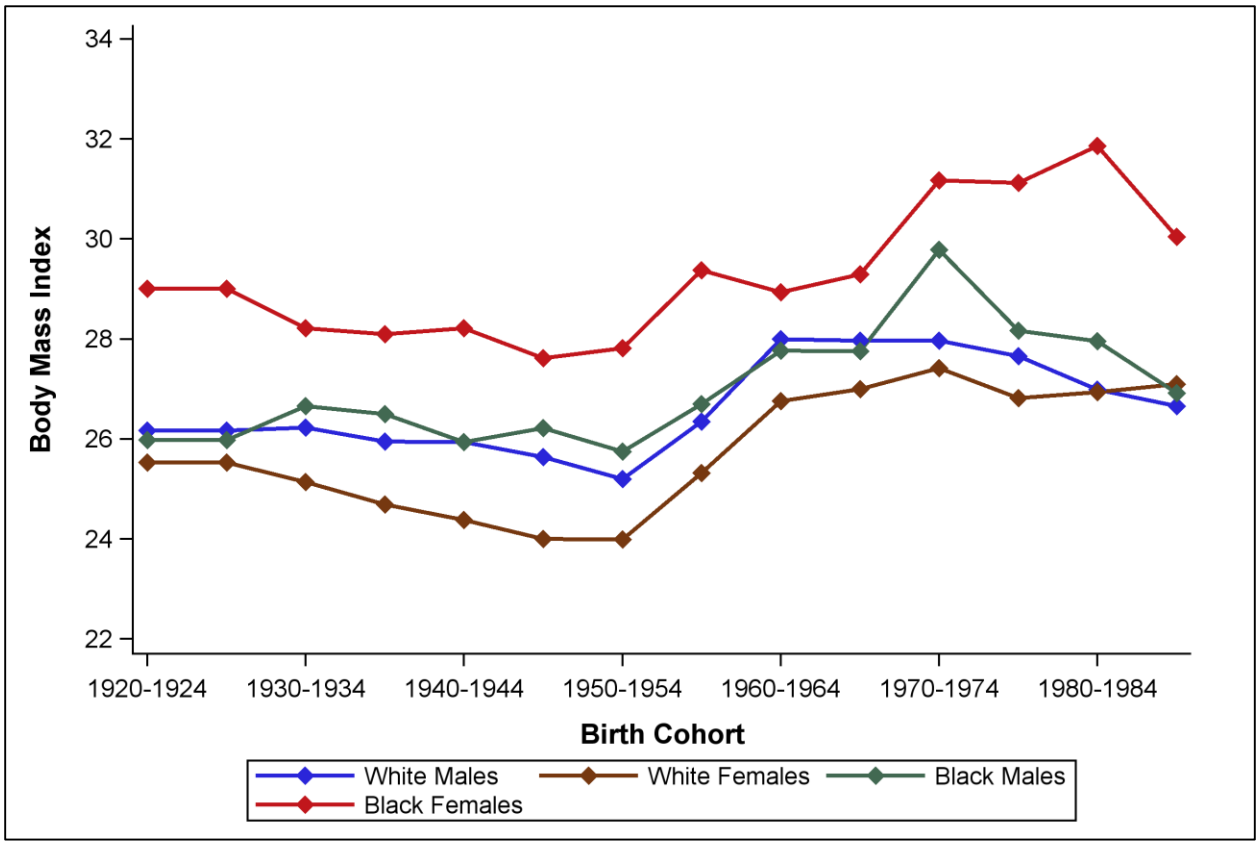


Figure 5.12. Median body mass index by birth cohort for four race/sex subgroups, ages 21-49 years, National Health and Nutrition Examination Survey, 1973-2012

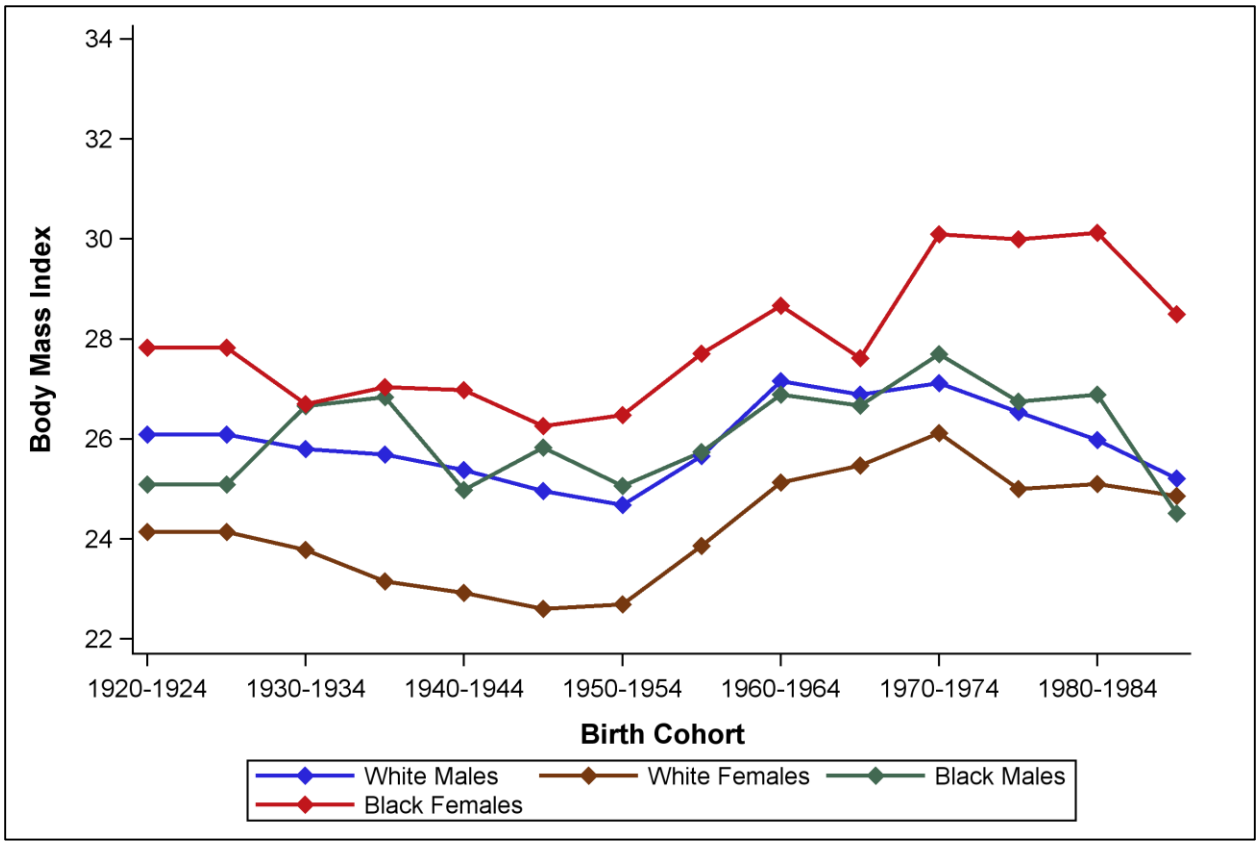


Figure 5.13. Prevalence of physical inactivity by time period for four race/sex subgroups, ages 21-49 years, Behavioral Risk Factor Surveillance System, 1984-2011

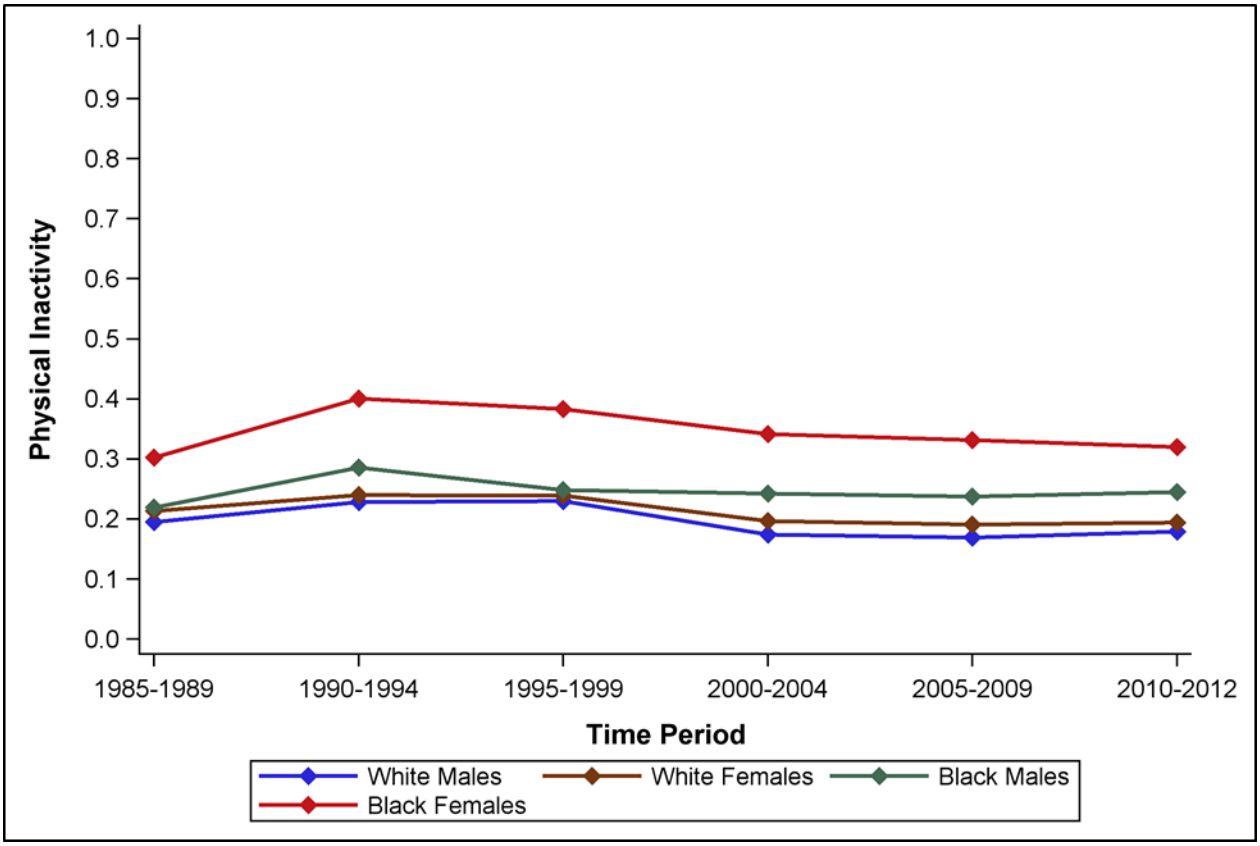


Figure 5.14. Prevalence of physical inactivity by birth cohort for four race/sex subgroups, ages 21-49 years, Behavioral Risk Factor Surveillance System, 1984-2011

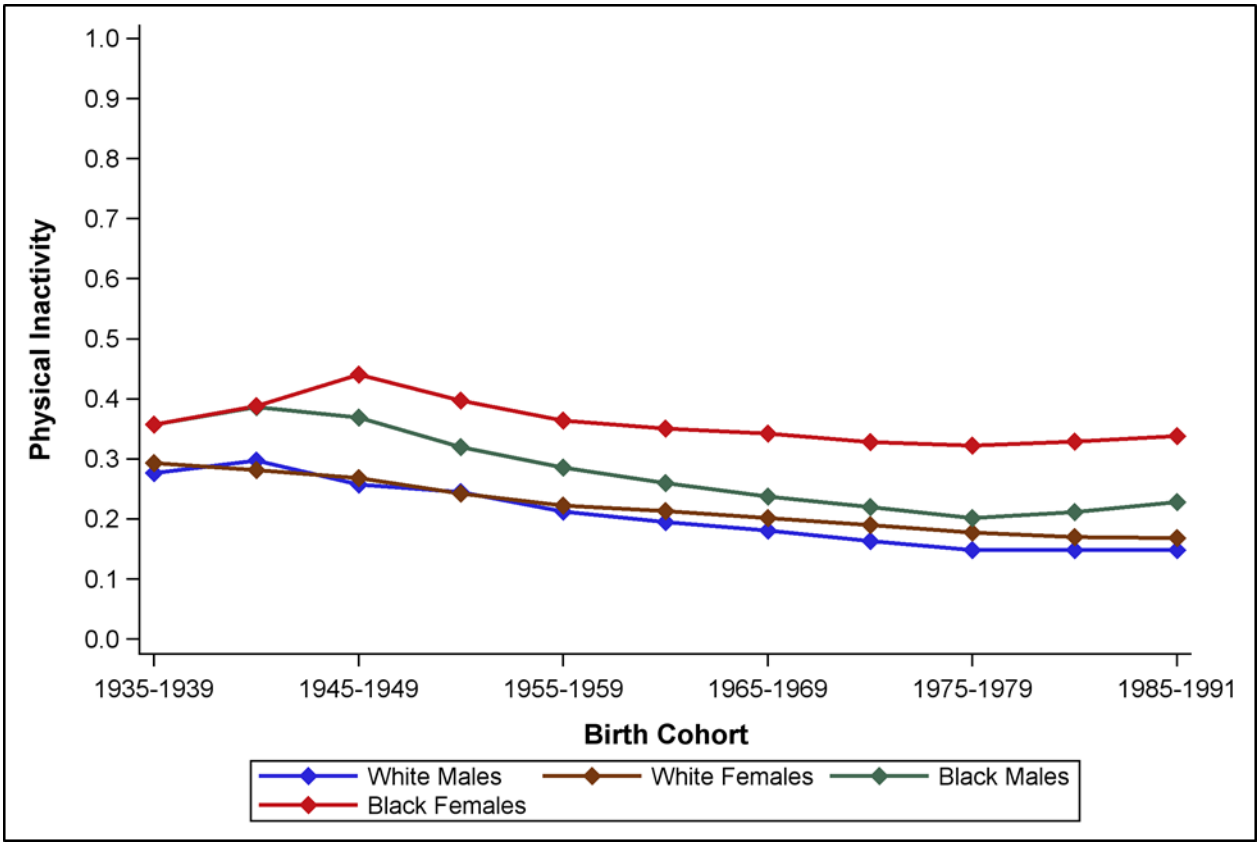


Figure 5.15. Prevalence of current smoking by time period for four race/sex subgroups, ages 21-49 years, National Health Interview Survey, 1974-2012

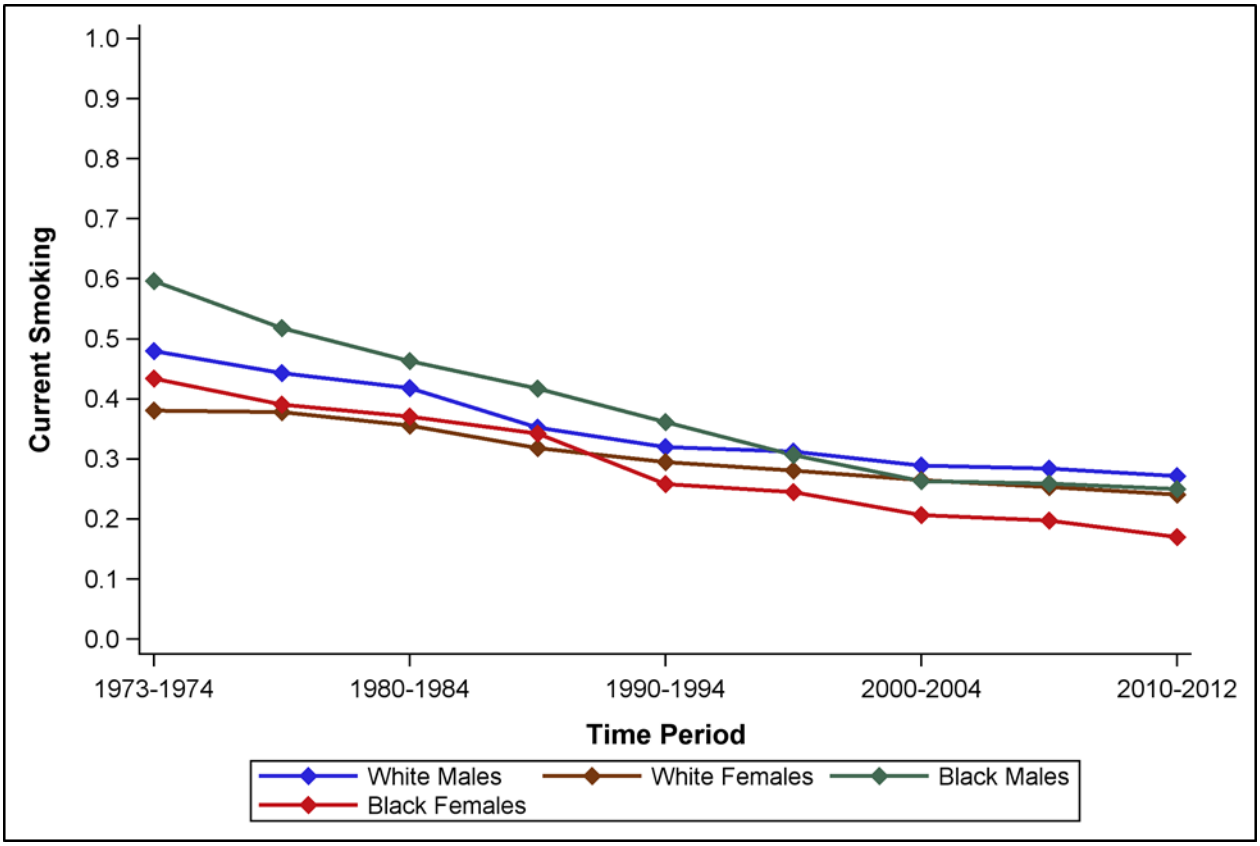


Figure 5.16. Prevalence of current smoking by birth cohort for four race/sex subgroups, ages 21-49 years, National Health Interview Survey, 1974-2012

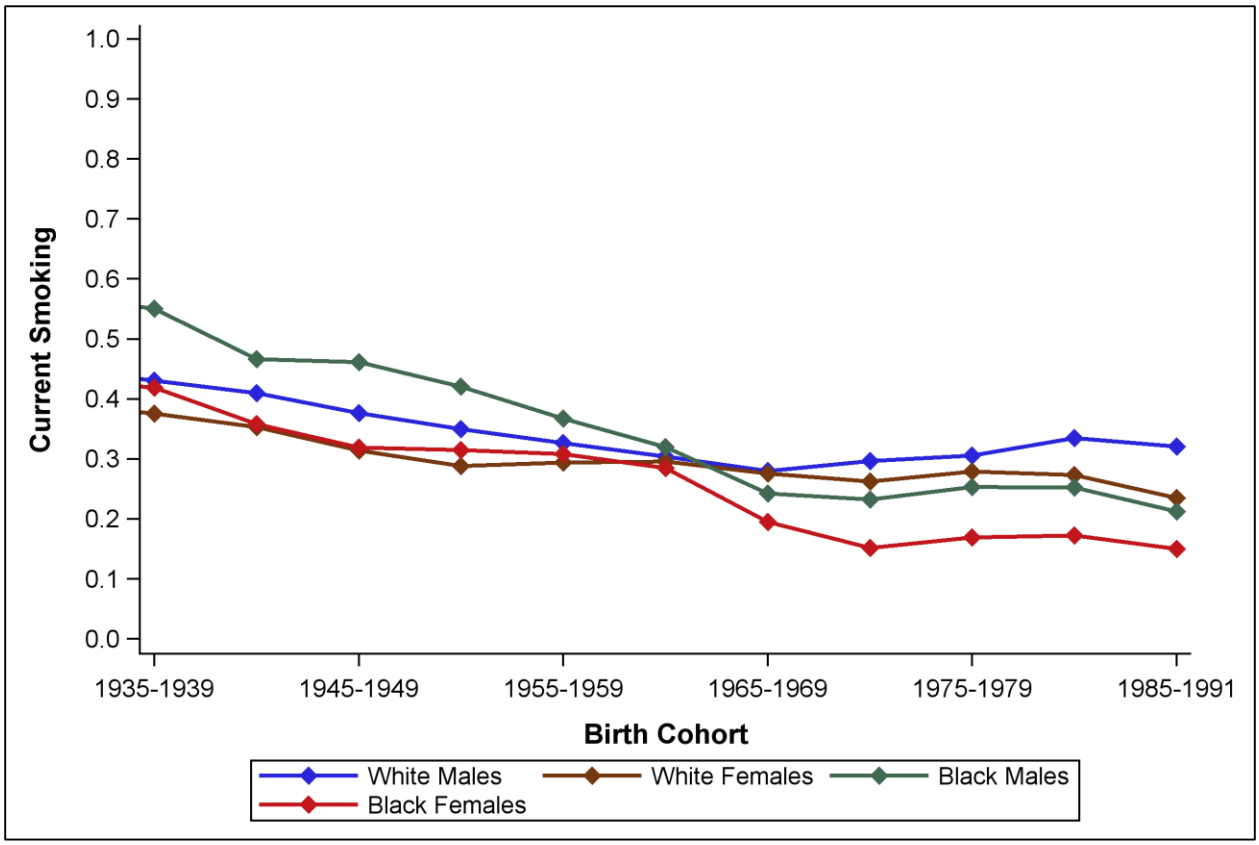


Figure 5.17. Predicted incidence of colorectal cancer by time period, all ages, SEER 9, 1973-2012

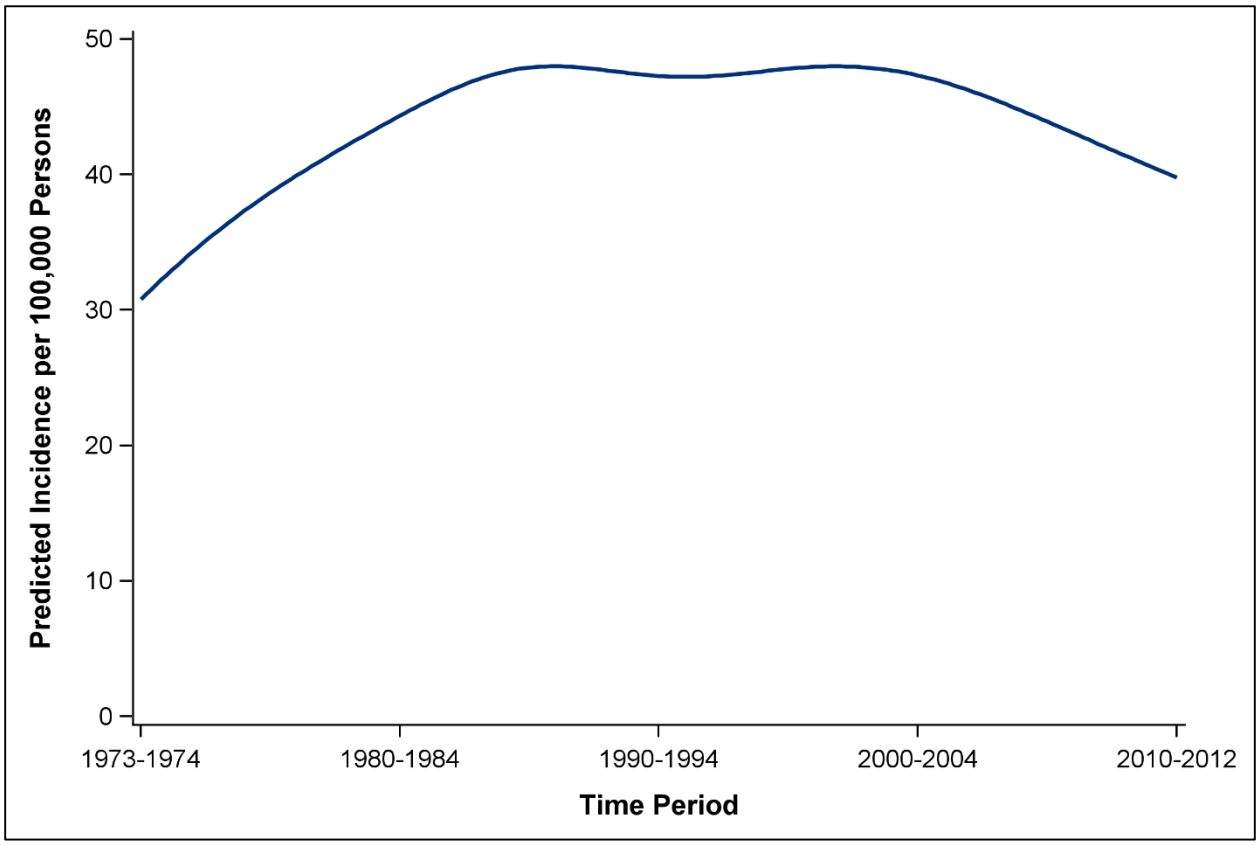
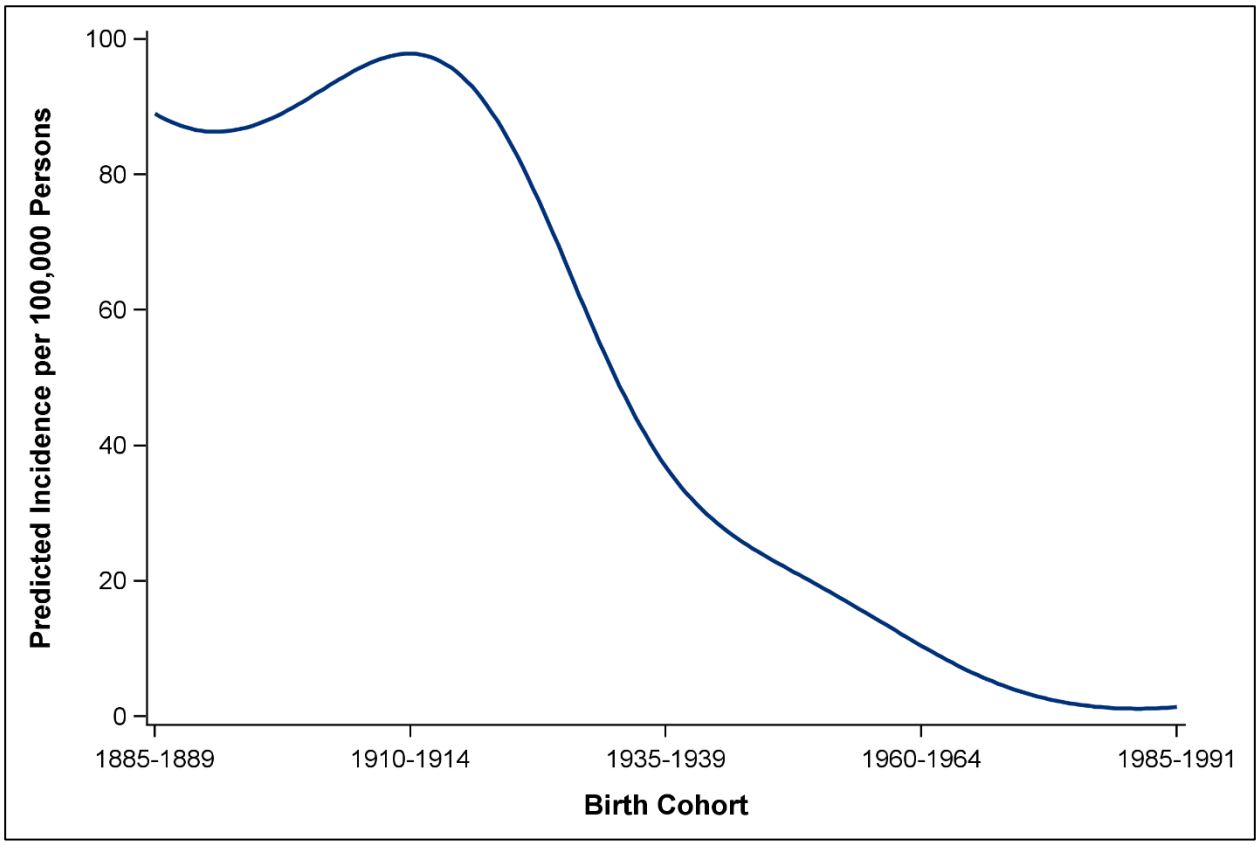


Figure 5.18. Predicted incidence of colorectal cancer by birth cohort, all ages, SEER 9, 1973-2012



CHAPTER 6
RESULTS: PATTERNS OF COLONOSCOPY USE IN YOUNGER ADULTS

A. INTRODUCTION

Recent research suggests the incidence of colorectal cancer (CRC) in younger populations (age <50 years) is rapidly increasing. Data from the Surveillance, Epidemiology, and End Results (SEER) program show the incidence of CRC has increased by up to 5% per year among younger adults since the 1970s.[1-6,68] Current projections suggest that, by 2030, approximately 11% of all colon and 23% of all rectal cancers may be diagnosed in patients younger than age 50.[68]

Trends in cancer incidence, such as those related to young-onset CRC, are a reflection of both changes in diagnoses (e.g., screening, case ascertainment, practice patterns) and exposures (e.g., risk factors). As a result, interpretation of cancer incidence statistics, particularly in cancers for which screening may lead to earlier detection, is subject to lead-time bias, length-biased sampling, and overdiagnosis.[130] Lead time bias and length-biased sampling occur when a greater detection of early-stage or slow-growing disease (via screening) artificially inflates incidence rates. Overdiagnosis, an extreme form of length-biased sampling, occurs when a cancer is detected at such an indolent stage by screening that it would have not been diagnosed otherwise. There is often debate over whether increases in cancer incidence are due to these biases or represent a true increase in clinically meaningful disease. For example, the rise in prostate cancer incidence that followed shortly after in the introduction of prostate-specific-antigen screening in the late 1980s is believed to be an artifact of earlier diagnoses

rather than an increase in disease burden.[131] It is important to consider the context and changes in diagnostic practices in which increases in incidence may arise.

Young-onset CRC is a unique opportunity to understand the relationship between changes in diagnoses and increases in incidence because of the availability of colonoscopy. Screening for CRC with colonoscopy is widely endorsed in older populations to reduce incidence and mortality. Although current guidelines recommend average risk individuals begin screening at age 50,[10] a number of colonoscopies may be performed in younger adults to evaluate symptoms related to gastrointestinal diseases (e.g., inflammatory bowel disease). More extensive use of colonoscopy in this setting could explain some of the recent increases in young-onset CRC; however, the frequency of colonoscopy use in younger populations has not been studied. As a first step toward understanding the impact of colonoscopy on trends in incidence, the purpose of this study was to describe patterns of colonoscopy use in a commercially insured population of adults under the age of 50. We hypothesized that colonoscopy use in younger adults has increased over time but would not fully account for trends in the incidence of young-onset CRC

B. METHODS

We characterized the patterns of colonoscopy use in younger adults (age <50 years) from 2001-2013 using MarketScan Claims and Encounters data (Truven Health Analytics, Ann Arbor, MI). MarketScan is a large employer-based claims database that includes 77 contributing employers and 12 contributing health plans, with 126 unique carriers and 8 Medicaid states. We summed the total number of months that individuals ages 18-49 years were enrolled in their insurance plan in each study year. For example, patients who were enrolled in their insurance plan for 8 months in the year 2001 (January 1, 2001 to December 31, 2001) contributed 8 months of “enrollee-time” during that year. Patient enrollee-time was counted in each study year without regard to enrollment in previous or subsequent years (i.e., patients could contribute

enrollee-time across all calendar years). Using standardized denominators of enrollee-time is a reasonable approach because the goal of the analysis was to estimate changes in colonoscopy use by year.

Any colonoscopy delivered in the outpatient setting was identified using International Classification of Disease, Ninth Revision, Clinical Modification (ICD-9-CM) and Current Procedure Terminology® (CPT) codes. Several validation studies of administrative data support the use of ICD-9-CM and CPT codes for identifying colonoscopy procedures.[90,91] We described time trends in colonoscopy use by calculating the rate of colonoscopy per 1,000 enrollee-years in each calendar year. We assumed constant rates within each calendar year. Colonoscopy rates were also examined by sex (male vs. female), age group (18-29 years, 30-39 years, 40-49 years), and geographic region (northeast, north central, west, south). We did not examine the indication for colonoscopy because our primary interest was in understanding CRC incidence trends (i.e., a colonoscopy would detect cancer regardless of why it was performed). A list of ICD-9-CM and CPT codes used to identify colonoscopy is provided in Table 3.4.

C. RESULTS

A total of 2,961,283 colonoscopies were performed in 173,774,053 enrollee-years during the study period. Mean age (range 35.2-36.5) and number of enrollee-months (range 9.8-10.4) contributed by individuals were similar across calendar years. The overall rate of colonoscopy increased from 2001 (14.4 per 1,000) to 2009 (18.8 per 1,000) and subsequently decreased through 2013 (15.6 per 1,000) (Figure 6.1). Females more frequently received colonoscopy in all years compared to males (Figure 6.2). At its peak in 2009, the colonoscopy rate among females was 20.6 per 1,000 and 16.6 per 1,000 among males. Colonoscopy rates were highest among the 40-49 year age group and lowest among the 18-29 year age group in all calendar years (Figure 6.3). Within each age group, rates generally increased from 2001 to 2009, with slight decreases through 2013 (Table 6.1). Colonoscopy rates were similar in the north central

and southern regions, but they remained consistently higher in the northeast and lower in the west (Figure 6.4) over time.

D. DISCUSSION

It is unclear from our findings how the use of colonoscopy in younger populations relates to trends in the incidence of young-onset CRC. We expected a more consistent trend in colonoscopy rates (i.e., no increase or monotonic increase), but the patterns were variable. The use of colonoscopy increased from 2001 through 2009 (relative increase of 30%) and accompanied increases in the incidence of CRC in younger populations during that period; however, rates of colonoscopy decreased in the latter half of the study period, while CRC incidence rates have continued to increase.

Screening can result in the earlier detection of small tumors that increase incidence and survival without changing mortality (i.e., lead-time bias). Early detection should lead to an increase in the incidence of early-stage disease that levels off and decreases a few years after mass screening is initiated, followed by a decrease in advanced disease. The amount of decrease in incidence and the time over which it occurs should reflect how much earlier most cancers are being detected. For example, the incidence of CRC in the overall population increased shortly after population-based screening was introduced in the 1980s. Incidence rates rose from 59.3 per 100,000 in 1975 to 66.3 per 100,000 in 1985.[94] Incidence rates have subsequently declined (largely due to removing premalignant polyps through polypectomy—see Chapter 5). The peak in incidence shows how the introduction of screening technologies can initially increase the earlier detection of cancers that may have otherwise been symptomatic in later years.

Increases in cancer incidence may also be a reflection of length-biased sampling because screening tests select for cancers that innately possess favorable prognoses. Cancer

screening tests work better to identify slower-growing lesions (i.e., those that exist for a longer period) than rapidly advancing ones. This is commonly observed in breast cancer, where screening (vs. symptomatic) mammography has been shown to be an important prognostic factor in survival.[132] Like lead-time bias, detection of these incidental or insignificant cancers increases the incidence of early-stage cancers and observed survival, but incidence rates should slowly level off over time and have no effect on mortality rates or advanced stage disease. Both local and regional stage CRC have increased in younger populations[2] as colonoscopy has become more widespread, which may give the impression that the increases in incidence are a reflection of screening-related biases. Mortality rates have remained relatively stable.[48] However, there is no evidence for the subsequent decline in incidence rates that would be expected in the context of these screening biases. Albeit small on the absolute (vs. relative) scale, the incidence of young-onset CRC has consistently increased since the 1990s (see Chapter 5). The steady increases in incidence, combined with the more variable trends in colonoscopy use, may suggest there are factors other than colonoscopy that account for the growing number of young patients diagnosed with CRC.

Despite the ambiguity of its relationship with incidence, we still observed an increase in the use of colonoscopy during the first half of the study period (2001-2009). Reasons for this increase are difficult to explain. The colonoscopy rates in younger patients parallel a striking rise in colonoscopy use among older patients that may have been fueled, in part, by Medicare coverage decisions regarding CRC screening. Medicare reimbursement was expanded to include screening colonoscopy for those at average-risk in July 2001, largely in an effort to reduce access and financial barriers to CRC screening[133] (the policy was revised in 2007 so that the Part B deductible no longer applied).[134] With more widespread use of colonoscopy to screen older adults, the thresholds for performing colonoscopy in young patients may have been lower. Further, there has been increased recognition of the high profit margins on endoscopy services.

Many ambulatory surgery centers were opened in the late 1990s and early 2000s to increase the capacity and efficiency of performing colonoscopy. Colonoscopy is the primary revenue source for many specialty centers and may account for as much as 60-70% of all revenue.[135] The increased endoscopy capacity may have provided easier access to colonoscopy for younger patients.

In 2010, however, there was a decline in colonoscopy use that persisted through the end of the study period. This decline was seen similarly across age, sex, and geographic region. One possibility is that the changes in national healthcare policy may have led to stricter reimbursement policies. Since it was first introduced in 2010, the Affordable Care Act has promoted an increased focus on preserving the quality of care by delivering the right service to right patient and right time. All health plans are required to cover preventive services rated as grade A or B by the U.S. Preventive Services Task Force (USPSTF), including screening colonoscopy, without charging copayment or coinsurance. The USPSTF recommendations make clear the appropriate age interval during which routine colonoscopy should be performed—between the ages of 50 and 75. Advances in healthcare reform may have significantly curtailed the use of endoscopy in populations (including those younger than the screening age) where routine use is not recommended.

Alternatively, the decline in colonoscopy use may be a reflection of the economic downturn that followed the 2008 financial crisis. Many Americans lost insurance coverage or had reduced benefits as a result of the recession. This was especially true among younger or employed populations, where the number of uninsured increased dramatically, and rates of employer-sponsored coverage simultaneously declined.[136] Downstream effects from the economic recession may have impacted willingness to see providers and pay for (expensive) colonoscopy copays performed for nonscreening indications. Colonoscopy use did not decrease among older persons or those covered under Medicare during this same period. Data from the

Behavioral Risk Factor Surveillance System show a comparable proportion (approximately 65%) of adults up-to-date with CRC screening from 2010-2012.[137,138] Similarly, the rates of colonoscopy remained relatively constant among adults ages 50-64 in the MarketScan data (results not shown). Considered alongside colonoscopy patterns in older adults, declines in colonoscopy use among younger populations could be due to the inability to pay for endoscopy procedures during a time of economic downturn.

We also observed differences in colonoscopy use by sex and geographic region, with higher rates among females and in the northeast, especially during the period 2001-2009. Patterns of colonoscopy overuse among older adults may help explain trends in overuse in these two younger population subgroups. Some studies of colonoscopy overuse have found a higher proportion of females (vs. males) receive colonoscopy in excess of guideline-recommended intervals,[139] while others show higher rates among males[140,141] or no differences by sex.[142] The exact reasons for the different associations reported in the literature are not fully understood. Meanwhile, research[139,143] in a wide variety of practice settings, including Medicare, the Veterans Health Administration, and integrated care delivery systems, consistently shows regional differences in colonoscopy use (for screening or surveillance), with up to 50% variation across states.[144-146] For example, a study of Medicare patients found substantial geographic variation in the frequency of early repeat colonoscopy after a negative screening colonoscopy; repeat colonoscopy rates ranged from 40-55%. A similar study among Medicare beneficiaries over the age of 70 reports potentially inappropriate colonoscopies that varied from 20-31% by health referral region. There was even more variation noted by health service area (i.e., cluster of counties that are relatively self-contained with respect to hospital care) within states. Geographic variations in care are often due to the availability of medical services rather than differences in population health or socioeconomic status.[147,148] Further, rates of colonoscopy overuse in these studies were generally consistent in groups of contiguous states, which suggests local physician practice patterns or supply may drive more frequent use of endoscopy.

Separately, our study also highlights the challenges of identifying indication for colonoscopy in administrative claims. It has long been recognized that health care data are a valuable resource for measuring use, safety, and effectiveness of colonoscopy, but they do no not distinguish well between colonoscopies performed for screening and those done for surveillance or diagnostic reasons. Differentiating colonoscopies performed to diagnose CRC (i.e., a CRC diagnosis is almost always confirmed with colonoscopy) in this population may have allowed for a better understanding of how background rates of colonoscopy use independently contribute to increases in incidence. A number of algorithms have been developed over the past decade or so to improve the classification of colonoscopy indications in administrative data.[90,91,149-155] Most of these algorithms use diagnosis and procedure codes that indicate the presence or absence of gastrointestinal procedures, signs, or symptoms (e.g., bleeding). Sensitivity and specificity remain suboptimal, and the algorithms can produce different results depending on the codes used or the length of time evaluated prior to the colonoscopy. As a result, we did not systematically evaluate the reasons for colonoscopy in this younger population. There may be even greater challenges to identifying indications for colonoscopy in younger adults because colonoscopies could be billed with symptom-related diagnosis codes for the purposes of reimbursement. Regression-based approaches that combine electronic health records and billing claims have more recently been used to determine colonoscopy indication and show promise for improving the accuracy of existing algorithms.[149]

A limitation of this study is the inability to reasonably exclude individuals with a family history or genetic predisposition to CRC who would qualify for earlier screening. We also did not exclude those with a previous diagnosis of CRC or history of colectomy, but this would be infrequent in a younger population. Greater awareness of a family or personal history of CRC—and the relevant recommendations for screening and surveillance—may have contributed to the increases in colonoscopy use we observed. In particular, the higher rates of colonoscopy use

among the 40-49 year age group could be due to more conscientious adherence to guidelines that advocate earlier screening in higher-risk populations. A recent study[156] that examined the prevalence of colonoscopy among privately insured (through health maintenance organizations) patients aged 40-49 years also found a moderate proportion of patients received screening colonoscopy. Consensus guidelines recommend CRC screening for those with a family history begin at age 40 or 10 years before the youngest relative was diagnosed with CRC.[119] The penetrance of genetic mutations (i.e., familial adenomatous polyposis, Lynch syndrome) is unlikely to have changed over time,[13,14,157] but there may be increased recognition of the benefits of earlier screening in higher-risk individuals. In contrast, the increasing number of colonoscopies performed in the 18-29 and 30-39 year age groups in this study are not likely to be accounted for by family history. The absolute risk of CRC among individuals younger than age 40 is very low.

In addition, we could not quantify the relationship with colonoscopy use and the incidence of young-onset CRC. We considered developing a model that estimated the association between endoscopy and cancer incidence (i.e., an ecologic analysis that compared rates across population subgroups), but we felt substantially limited in identifying a comparable population to derive incidence rates. For example, the SEER registries, while widely used in studies of cancer incidence, cover a more diverse population with differing levels of insurance and healthcare utilization. Persons in MarketScan data are all commercially insured through their employers. Our results likely overestimate colonoscopy rates in populations without insurance coverage or access to endoscopy. As a result, it would have been difficult to draw appropriate conclusions from these two populations together. Further, such a model would require multiple assumptions regarding length- and lead-times. Estimates of these parameters can vary greatly depending on population and context.[158,159] It is unknown whether the

assumptions related to lead and length time bias in the overall CRC population are suitable for young-onset CRC.

There are several strengths of this study. Colonoscopy is considered to be well-reported in claims because it is an expensive procedure. We used procedure codes that have been shown to accurately identify colonoscopy in outpatient settings. In addition, we used a large, employer-based claims database, which yielded a large enough study population to determine the use of a relatively uncommon procedure (i.e., the prevalence of colonoscopy in younger individuals is low). Although individuals represented in MarketScan are all insured through their employers, and findings cannot be generalized to uninsured populations or those without access to endoscopy, the data do have a number of advantages. MarketScan is fully integrated (inpatient, outpatient, enrollment) at the patient-level, represents all ages (i.e., compared to Medicare claims that generally only cover age ≥65), and is not limited to a single payor or geographic region.

Little is known about the underlying mechanisms and factors that have contributed to increases in CRC incidence in younger populations. The trends in colonoscopy use we observed among younger adults were similar to those in age-eligible (i.e., age ≥50 years) populations reported by others, where the use of colonoscopy for CRC screening and the use of any lower endoscopy (regardless of screening indication) have both increased since 2000. Although colonoscopy use in younger adults has increased, the declines in in more recent years could suggest there are other factors—beyond those related to detection and diagnosis—involved in the development of young-onset CRC.

Table 6.1:Number of colonoscopies performed (in thousands), enrollee-years (in thousands), and colonoscopy rate per 1,000 enrollee-years, overall and by sex, age, and geographic region

	Calendar Year												
	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Overall													
COL	35.4	71.5	110.7	157.9	180.6	210.3	210.8	320.9	315.1	330.5	364.9	365.7	287.1
Enrollee-Years	2,458.2	4,754.3	7,321.8	9,525.5	10,535.3	12,019.3	11,918.3	17,176.6	16,802.1	18,860.7	21,717.0	22,308.6	18,376.3
Rate	14.4	15.0	15.1	16.6	17.1	17.5	17.7	18.7	18.8	17.5	16.8	16.4	15.6
Male													
COL	14.9	29.3	46.1	65.8	75.5	87.7	88.0	136.2	132.4	137.7	153.3	153.1	119.1
Enrollee-Years	1,146.2	2,219.3	3,429.9	4,463.3	4,953.6	5,732.0	5,686.0	8,167.2	7,954.2	8,998.8	10,466.9	10,815.3	8,839.0
Rate	13.0	13.2	13.5	14.7	15.3	15.3	15.5	16.7	16.6	15.3	14.6	14.2	13.5
Female													
COL	20.5	42.2	64.6	92.1	105.0	122.6	122.8	184.7	182.7	192.7	211.5	212.6	168.1
Enrollee-Years	1,312.0	2,535.0	3,891.9	5,062.2	5,581.8	6,287.3	6,232.3	9,009.4	8,847.9	9,861.9	11,250.2	11,493.3	9,537.3
Rate	15.6	16.7	16.6	18.2	18.8	19.5	19.7	20.5	20.6	19.5	18.8	18.5	17.6
18-29 Years													
COL	3.0	6.9	10.7	16.3	19.0	23.1	24.2	38.8	39.9	42.3	52.8	55.8	43.7
Enrollee-Years	732.1	1,443.7	2,170.8	2,840.0	3,156.8	3,680.2	3,718.3	5,442.7	5,339.3	5,927.0	7,405.5	7,884.1	6,607.7
Rate	4.1	4.7	5.0	5.7	6.0	6.3	6.5	7.1	7.5	7.1	7.1	7.1	6.6
30-39 Years													
COL	7.6	16.5	25.3	36.2	41.2	48.6	48.9	75.6	75.3	79.6	86.5	87.9	68.2
Enrollee-Years	733.8	1,495.0	2,330.3	2,985.3	3,296.1	3,736.8	3,692.9	5,350.1	5,243.4	5,918.5	6,521.0	6,603.8	5,408.1
Rate	10.4	11.0	10.9	12.1	12.5	13.0	13.2	14.1	14.4	13.5	13.3	13.3	12.6
40-49 Years													
COL	24.8	48.2	74.7	105.5	120.4	138.6	137.7	206.4	199.9	208.5	225.6	222.0	175.2
Enrollee-Years	992.3	1,815.5	2,820.7	3,700.2	4,082.5	4,602.3	4,507.1	6,383.8	6,219.3	7,015.3	7,790.5	7,820.8	6,360.5
Rate	25.0	26.5	26.5	28.5	29.5	30.1	30.6	32.3	32.1	29.7	29.0	28.4	27.5
Northeast													
COL	5.6	8.1	13.6	17.4	20.6	23.9	21.6	58.8	48.5	60.4	79.3	79.2	65.8
Enrollee-Years	338.7	445.4	684.7	821.3	1,000.1	1,135.6	1,026.3	2,694.6	2,193.7	2,856.2	3,883.7	3,971.9	3,436.1
Rate	16.4	18.3	19.9	21.2	20.6	21.0	21.1	21.8	22.1	21.1	20.4	19.9	19.2
North Central													
COL	10.9	19.8	25.2	32.5	38.3	60.5	59.8	79.0	81.9	81.7	85.4	85.6	60.0
Enrollee-Years	686.6	1,219.3	1,493.1	1,875.9	2,112.5	3,335.7	3,291.7	4,222.0	4,350.7	4,570.8	5,045.6	5,143.8	3,833.8
Rate	15.9	16.3	16.9	17.3	18.1	18.1	18.2	18.7	18.8	17.9	16.9	16.6	15.7
West													
COL	15.2	35.1	48.8	78.8	89.7	99.8	102.7	147.6	147.9	137.9	139.5	137.1	104.5
Enrollee-Years	983.6	2,238.4	2,903.0	4,312.0	4,785.0	5,619.3	5,648.3	7,696.4	7,671.6	7,519.7	7,987.3	8,148.7	6,346.5
Rate	15.5	15.7	16.8	18.3	18.7	17.8	18.2	19.2	19.3	18.3	17.5	16.8	16.5

South													
COL	3.7	8.2	22.3	27.1	30.2	24.4	25.7	34.4	36.1	47.5	49.2	55.4	48.6
Enrollee-Years	445.9	838.1	2,182.2	2,396.6	2,534.6	1,839.7	1,901.1	2,490.3	2,531.6	3,838.3	4,125.2	4,597.7	4,290.6
Rate	8.2	9.8	10.2	11.3	11.9	13.2	13.5	13.8	14.3	12.4	11.9	12.0	11.3

NOTE: Colonoscopy rate reported as number of colonoscopies performed per 1,000 enrollee-years
Abbreviations: COL, colonoscopies

Figure 6.1. Colonoscopy rate per 1,000 enrollee-years by calendar year

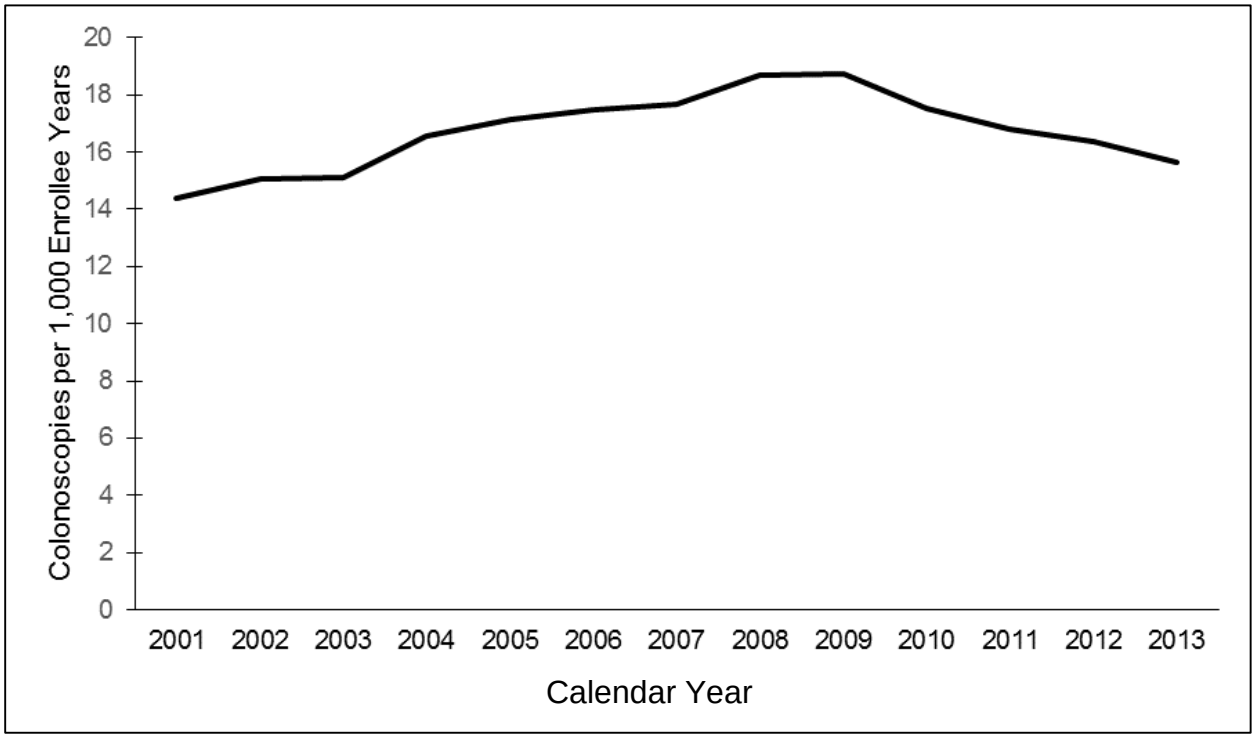


Figure 6.2. Colonoscopy rate per 1,000 enrollee-years by calendar year and sex

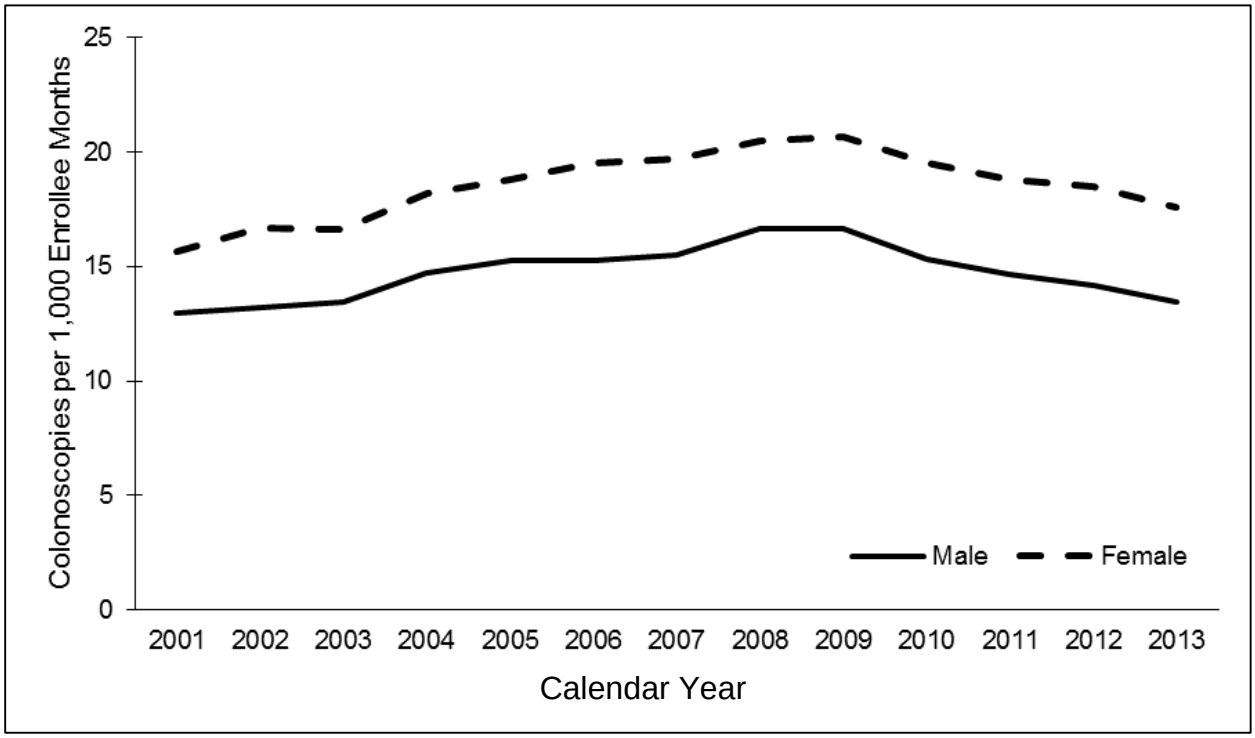


Figure 6.3. Colonoscopy rate per 1,000 enrollee years by calendar year and age

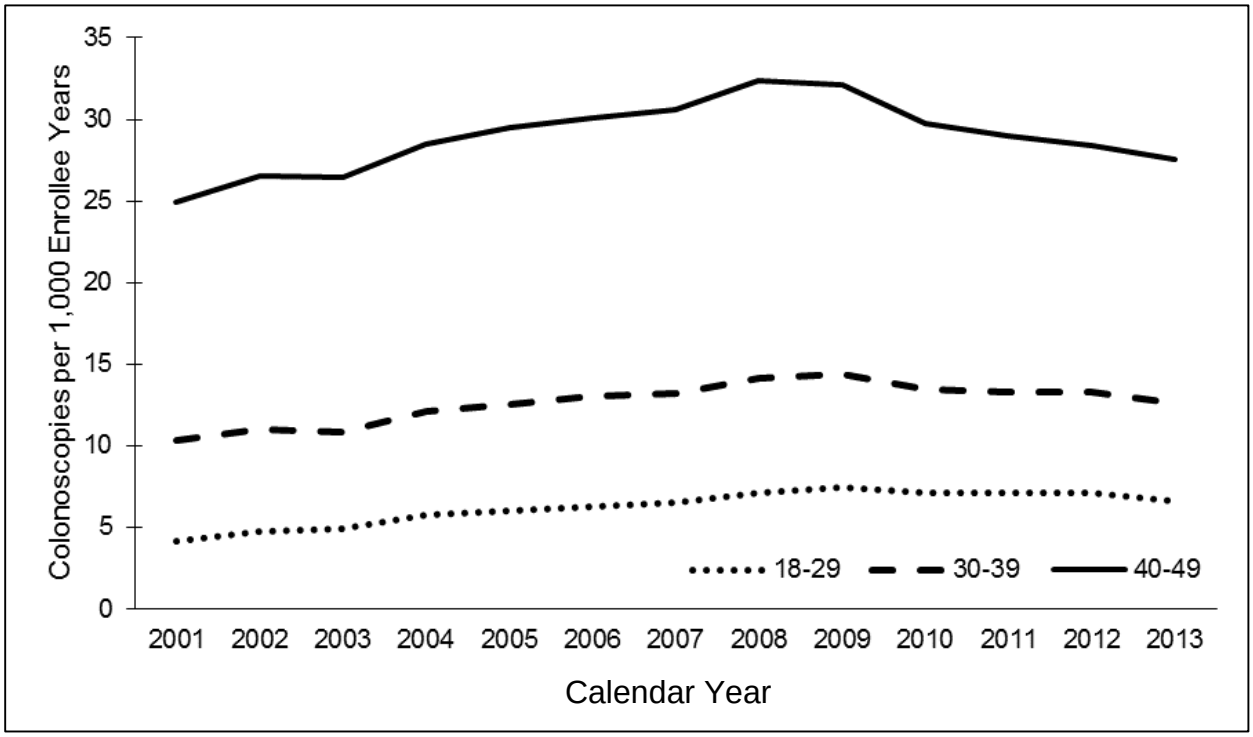
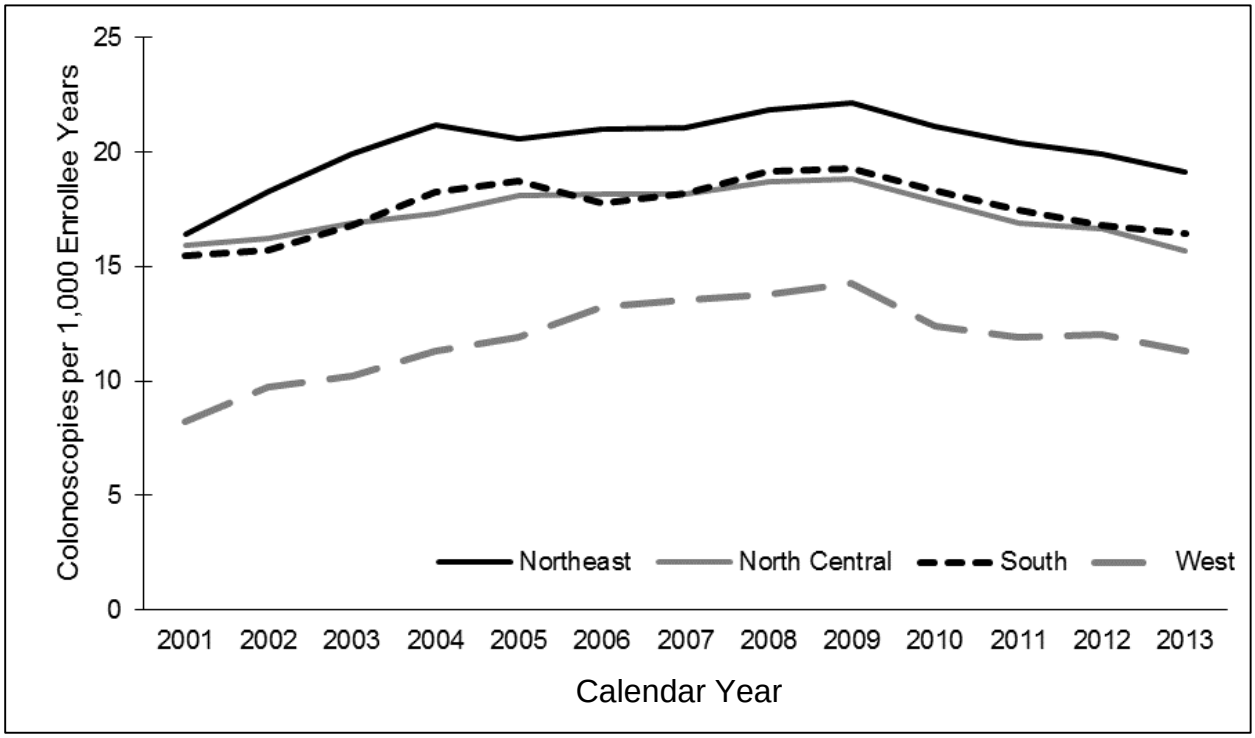


Figure 6.4. Colonoscopy rate per 1,000 enrollee years by calendar year and geographic region



CHAPTER 7
DISCUSSION

A. SUMMARY OF FINDINGS

This dissertation examined the increasing incidence of colorectal cancer (CRC) in younger (age <50 years) adults by leveraging large, population-based data sources. There were three primary goals: 1) to describe the demographic, clinicopathologic, and socioeconomic characteristics (*Specific Aim 1a*) and treatment patterns (*Specific Aim 1b)* of younger patients newly diagnosed with stages II and II CRC; 2) to estimate the contribution of age, time period, and birth cohort to increases in the incidence of young-onset CRC (*Specific Aim 2a*) and describe difference in the prevalence of lifestyle-related behavior risk factors by time period and birth cohort (*Specific Aim 2b*); and 3) to determine patterns of colonoscopy use in younger adults (*Specific Aim 3*).

To accomplish the first aim, Patterns of Care (POC) studies were used to compare the demographic, clinicopathologic, and socioeconomic characteristics of younger and older patients newly diagnosed with stages II and III CRC. POC data were linked with the Area Health Resource File (AHRF) to better describe socioeconomic indicators of the study population. Differences in treatment patterns (e.g., surgery, chemotherapy, and radiation therapy) were also examined across categories of age at diagnosis. Receipt of chemotherapy and radiation therapy was verified by the treating physician. Results showed important differences in the distribution of young-onset CRC by race/ethnicity. A higher proportion of black (13%) and Hispanic (15%) patients were diagnosed at age <50 years compared to whites. Young black patients more frequently had tumors located within the proximal colon (39%), whereas a higher proportion of

young white (41%) and Hispanic (33%) patients had rectal tumors. Combined with the growing literature on CRC tumor subtypes (i.e., showing differences in tumor subtypes by race), the differences we observed by race/ethnicity provide additional evidence for distinct mechanisms that may be driving CRC progression across population subgroups. In addition, younger CRC patients more frequently received “optimal” care, including better nodal counts at surgery, treatment at academic medical centers, and chemotherapy and radiation therapy. This may be a reflection of physician preferences to ambitiously treat disease simply based on younger age at presentation.

The second aim used Hierarchical Poisson models to estimate the independent effects of age, time period, and birth cohort on incidence of CRC. Incidence rates for all stages of CRC were derived from the Surveillance, Epidemiology, and End Results (SEER) program of cancer registries. When considered alongside *Specific Aim 1*, the findings from this aim give additional support to the hypothesis that there may be different mechanisms involved in the development of CRC across the life course. Specially, there was a differential contribution of time period and birth cohort on incidence patterns by age. Both a period (σ=0.02, p=0.02) and cohort (σ=0.04, p=0.004) effect were observed in the incidence of CRC in older (age ≥50 years) populations, but in contrast, there was only a cohort (σ=0.01, p=0.04) effect on young-onset CRC. In addition, the results show small absolute (vs. relative) increases in the incidence of young-onset CRC over time. Incidence rates increased by only a few additional cases per 100,000 persons over a 40-year time period, while others have simultaneously suggested *relative* increases of over 100%. Rather, the most dramatic changes in incidence occurred in the screening-age (≥50 years) population, with large declines in incidence from the earliest to most recent time period. Some professional organizations have promoted changing screening guidelines to being average-risk screening at a younger age, but these findings highlight the importance of evaluating absolute measures when making public health policy decisions.

As part of the second aim, trends in the prevalence of obesity, physical inactivity, and smoking were also examined across time period and birth cohort using data from the National Health and Nutrition Examination Survey (NHANES), Behavioral Risk Factor Surveillance System (BRFSS), and National Health Interview Survey (NHIS), respectively. The prevalence of obesity generally increased across time period (12.9% in 1975-79 to 34.6% in 2010-12) and birth cohorts (18.3% in 1924-29 to 27.6% in 1985-91), with the largest increases among black females. Physical inactivity remained relatively stable over time (21.1% in 1985-89 to 20.1% in 2010-12) in all race/sex subgroups combined. Trends in cohort-physical inactivity were more consistent, with general declines across successive birth cohorts (29.9% in 1940-44 to 17.6% in 1980-84). The prevalence of smoking decreased dramatically across time periods (43.5% in 1974-74 to 24.8% in 2010-12) and birth cohorts (42.0% in 1924-29 to 26.0% in 1985-91), although there were some modest increases in smoking behavior among white males in the most recent birth cohorts. Population surveys can be helpful in efforts to understand the context of incidence trends, but they do not provide the level of detail needed to identify risk factors for disease. The results underscore the need for a better understanding of the risk factors, particularly those that may be related to differences across birth cohort, involved in the development of young-onset CRC.

Finally, in the third aim, patterns of colonoscopy use in younger adults (age <50 years) were examined in MarketScan Claims and Encounters data. MarketScan is a large employer-based administrative claims database that includes 77 contributing employers and 12 contributing health plans. Any colonoscopy performed in the outpatient setting was identified using validated procedure codes. Time trends in colonoscopy were described by calculating a rate of colonoscopy per 1,000 enrollee-years in each calendar year. The overall rate of colonoscopy increased from 2001 (14.4 per 1,000) to 2009 (18.8 per 1,000) and subsequently decreased through 2013 (15.6 per 1,000). Colonoscopy rates were highest in females,

individuals aged 40-49 years, and the northeast. These results provide some evidence that the wider use of colonoscopy in younger populations may partially explain rises in the incidence of young-onset CRC. However, rates of colonoscopy decreased in the latter half of the study period, while CRC incidence rates in younger populations have continued to increase. The more recent decline in colonoscopy use may suggest increases in young-onset CRC cannot be solely attributed to changes in the detection and diagnosis of CRC (via colonoscopy) but could also be related to changes in risk factors, such as obesity, over time.

B. PUBLIC HEALTH IMPLICATIONS

The results of this dissertation research have several implications for public health and clinical practice. First, this research enhances our understanding of the factors involved in the development of young-onset CRC and provides a real world context for understanding age-related differences in CRC incidence. Given the decreases in CRC incidence in older populations (i.e., due to screening), the findings show the age distribution of CRC is changing. Others have similarly noted that nearly 1 in 7 patients diagnosed with CRC is under the age of 50.[160] Despite the overall population trends in aging, by 2030, approximately 11% of all colon and 23% of all rectal cancers may be diagnosed in younger patients.[68] Although the increases are small in absolute magnitude, changes in the age distribution of CRC point to the challenges of managing CRC in younger patients, including earlier recognition of symptoms at diagnosis and treatment approaches. Younger patients may also experience problems with fertility, employment and productivity loss, and long-lasting treatment sequelae.

This work also highlights differences in CRC by race/ethnicity. Racial disparities in CRC incidence and mortality have long been recognized; blacks continually experience higher incidence and mortality rates compared to whites. Racial differences in CRC incidence may be a reflection of differences in the uptake of CRC screening among whites and blacks. NHIS data have historically shown substantially lower screening rates in blacks than whites,[66] although

more recent evidence from BRFSS suggests screening rates have become comparable (~58%).[161] It may be that differences in follow-up of abnormal screenings or use of surveillance screening, which are not captured by population-based surveys, also account for disparities in incidence and stage. Another possibility is that blacks are more frequently screened with fecal occult blood testing (FOBT),[162] while a higher proportion of whites undergo colonoscopy, and the effectiveness of FOBT may be reduced when patients do not adhere to a regular schedule.[163-166] Our findings support efforts to target screening programs in the underserved or interventions

Others have noted racial differences in tumor subsite, which was also observed in our *Specific Aim 1*. These differences may be a reflection of differences in biology or CRC risk factors (e.g., higher prevalence of obesity, physical activity levels, and cultural differences in diet) rather than disparities in detection and treatment. However, little scientific progress has been made in improving our understanding of the biological differences that contribute to the distribution of CRC tumor subtypes by race.

Lastly, this work has implications for the national program of CRC screening. Population-based cancer screening is often cited as an important success of the public health system in the U.S. CRC screening was first introduced in the late 1980s when the National Cancer Institute developed working guidelines for the early detection of cancer and is currently recommended at age 50 for those at average-risk. Given the increases in young-onset CRC, some have argued in favor of changing screening guidelines to being average-risk screening at a younger age, such as age 40. The findings of this study, however, highlight the importance of considering absolute (vs. relative) measures when making public health policy decisions. Although incidence of young-onset CRC has increased, the absolute incidence rates have only increased by a few cases per 100,000 persons over a 40-year time period. Screening decisions should be based on the balance of risks and benefits and not solely on increases in incidence.

C. STRENGTHS

Research Guided by Conceptual Framework

This dissertation addressed several gaps in our understanding of young-onset CRC. Most studies have failed to consider the multiple influences that have likely contributed to the increase in CRC incidence among younger adults, instead only focusing on a single dimension of CRC (e.g., relative increases in incidence). We relied on a conceptual model to leverage complementary methods from the fields of epidemiology, demography, and health services use to examine the underlying mechanism that have led to increases in CRC incidence in younger populations.

Population-Based Data Sources (Specific Aims 1 and 2)

An important strength of this study is the use of several population-based data sources (POC, SEER, BRFSS, NHIS, NHANES). Previous research has been limited to single institution studies or relied on small sample sizes, in which results generally reflect the distribution of characteristics of CRC patients treated at that institution rather than true characteristics of young-onset CRC. The breadth of data used in this study, ranging from a population-based cancer registry to national surveys on health behaviors, provided a robust, real world context for examining the increasing incidence of CRC in younger populations. For example, data from POC studies offer a number of unique advantages for conducting population-based epidemiologic research because each participating SEER registry area has a defined population. The age and sex distributions of patients in POC reflect those of the U.S. population, and the SEER program includes registries with a higher percentage of African Americans (Detroit, Atlanta, Louisiana), Asians (Seattle, San Francisco, Los Angeles, San Jose-Monterey), and Hispanics (Los Angeles, Greater California, New Mexico). The large size and ethnic diversity of these data sources were also strengths that enabled us to examine differences

within population subgroups by race/ethnicity. The study findings are generalizable because the demographic profiles of patients include in this study are similar to that of the general U.S. population.

Physician Verification of Treatment Receipt (Specific Aim 1)

POC studies also provide a greater breadth and depth of information than that available solely from medical claims and/or SEER registries. This was particularly true for our assessment of receipt of chemotherapy and radiation therapy in *Specific Aim 1b*. Detailed tumor and treatment information was abstracted from patient medical records and verified by treatment physicians; doctor verification substantially improved the completeness of treatment ascertainment (or confirmed that no chemotherapy or radiation therapy were given).

Measurement of Colonoscopy (Specific Aim 3)

Lastly, a strength of the study was measurement of colonoscopy in *Specific Aim 3* in MarketScan data. Marketscan is fully integrated (inpatient, outpatient, enrollment) at the patient-level, represents all ages (i.e., compared to Medicare claims that generally only cover age ≥65), and is not limited to a single payor or geographic region. Colonoscopy is considered to be well-reported in administrative claims because it is an expensive procedure. We used procedure codes that have been shown to accurately identify colonoscopy performed in outpatient settings. In addition, we used a large employer-based claims database, which yielded a large enough study population to determine the use of a relatively uncommon procedure (i.e., prevalence of colonoscopy in younger adults is low).

D. LIMITATIONS

Unknown Family History of or Genetic Predisposition to CRC (Specific Aims 1 and 3)

A limitation of the study was the inability to exclude persons with a family history of or genetic predisposition to CRC. Although the prevalence of these syndromes in younger populations remains very low,[13,14] the data may have helped explain changes in the distribution of young-onset CRC we observed in *Specific Aim 1.* In addition, higher-risk individuals would quality for earlier screening. Greater awareness of a family history or personal history of CRC—and the relevant recommendations for screening and surveillance—may have contributed to the increases in colonoscopy use we observed in *Specific Aim 3.* The higher rates of colonoscopy use among the 40-49 year age group could be due to more conscientious adherence to guidelines that advocate earlier screening in higher-risk populations. The penetrance of genetic mutations is unlikely to have changed over time, but there may be increased recognition of the benefits of earlier screening in higher-risk persons.

Information on Molecular Markers Not Routinely Collected (Specific Aim 1)

We also lacked data on molecular markers of CRC that may have helped explain the results of *Specific Aim 1* on differences in young-onset CRC by race/ethnicity. There is growing recognition that CRC is a heterogeneous disease, often defined by combinations of DNA mismatch repair or microsatellite instability (MSI), *MLHI* methylation (i.e.,CpG island methylator phenoypte), and mutations in *BRAF* and *KRAS* oncogenes. However, few patients in POC studies received MSI testing, and even fewer received *KRAS* and *BRAF* testing (data not shown), which precluded our ability to report the prevalence of molecular markers in the study population. These data may have supported the differences in tumor subtypes across racial groups reported in other studies. Collection efforts have likely improved since the study period

(1990-2010), and SEER now includes site-specific factors on MSI and *KRAS* mutation. Our study highlights the continued need for robust sources of molecular data at the population-level.

Ecologic Analysis of CRC Risk Factors (Specific Aim 2)

Our analysis of CRC risk factors in *Specific Aim 2* was limited to aggregate measures of obesity, physical inactivity, and smoking. Population surveys can be helpful in efforts to understand the context of incidence trends, but they do not provide the level of detail needed to identify risk factors for the disease. Ecological data describes characteristics of a group, and conclusions cannot be made regarding individuals (i.e., ecologic fallacy). Further, the APC models used in our study are not models of etiology and cannot provide information on the specific mechanism by which these risk factors are related to carcinogenesis. The analysis does not distinguish between the many factors that have led to increases in obesity prevalence (e.g., diet) that could also be related to CRC risk. There may be other etiologic factors or exposures that could not be accounted for in this study (e.g., intestinal microbiota), either because their relationship with CRC risk is still unknown or the prevalence is not monitored in population-based sources with sufficiently long follow-up.

Synthetic Cohort Approach (Specific Aim 2)

We used a synthetic cohort approach in *Specific Aim 2* to estimate the effect of birth cohort on CRC incidence rates. Specifically, by pooling SEER incidence data, we created a rectangular age by period array where columns corresponded to age-specific incidence rates in each period, and rows were age-specific incidence rates across all periods (see Figure 3.3). Linking the diagonal cells of the array gave incidence rates that belonged to individuals born in the same calendar year and age together (i.e., birth cohort). Although this approach is often used in demographic research, only a longitudinal panel study design provides data from true birth cohorts that follow identical individuals over time. As a result, we lacked data at the tail

ends (i.e., in the earliest and most recent cohorts) because not all ages were equally represented across the study period. This was also true for birth cohort patterns of lifestyle-related risk factors that were derived from NHANES, BRFSS, and NHIS.

Generalizability of Results from MarketScan (Specific Aim 3)

Colonoscopy rates were estimated using data from MarketScan. Although MarketScan provides a number of advantages over other administrative claims, individuals are all insured through their employers, and findings cannot be generalized to uninsured populations or those without access to endoscopy. The results may overestimate colonoscopy rates in the general population, where a greater proportion of younger adults may lack insurance coverage for the procedure. In addition, because colonoscopy is predominantly performed for screening in older adults, much of the literature on patterns of colonoscopy use is focused on the Medicare population. This made it challenging to draw comparisons to our results and provide explanations for variations in colonoscopy use in younger adults.

E. FUTURE RESEARCH

There are two important areas of future research that may build upon these results and address the limitations described above. First, results from *Specific Aim 1* highlight the need for a better understanding of the differences in molecular markers of CRC (e.g., MSI and *BRAF* and *KRAS* mutations) by age and race/ethnicity. Research has only recently explored the molecular features of young-onset CRC and found no difference in the overall mutational rate among younger and older patients. However, the specific mutations involved in young-onset CRC appear to be distinct. An emerging body of literature also shows there differences in CRC tumor subtypes across racial groups. For example, *KRAS* mutations are more common in black patients, while the frequency of *BRAF* mutation tends to be higher in tumors from whites. There are likely distinct mechanism driving CRC progression in various population subgroups.

Population-based studies (vs. those done in single institution settings) of molecular differences in young- and older-onset CRC may identify pathways of CRC carcinogenesis that account for the growing number of young patients diagnosed with CRC.

In addition, these results leave room for a more critical analysis of the risk factors for young-onset CRC. For example, the analysis of NHANES data shows increases in the prevalence of obesity in younger populations across both time periods and birth cohorts. It may be that increases in central adiposity, especially at younger ages, account for some of the increases in incidence, but the mechanisms involved in obesity and cancer risk have not been fully elucidated. The findings related to trends in obesity, physical inactivity, and smoking point to additional factors that may increase risk of young-onset CRC. Increases in incidence in younger persons reflect changes in that population and may be a reasonable focus for future etiologic studies

F. CONCLUSIONS

This study provides strong support for differences in the etiology of CRC across the life course. There may be different mechanisms involved in the development of CRC in younger (vs. older) adults. Specifically, results from *Specific Aim 1* show differences in the distribution of CRC by age, race/ethnicity, and tumor subsite, and *Specific Aim 2* demonstrates age-related differences in the effect of time period and birth cohort on incidence patterns. The mechanisms responsible for increases in young-onset CRC, albeit small on an absolute scale, remain unknown. Collectively, the results of this dissertation work underscore the need for a better understanding of the critical factors involved in the onset of CRC by age and race/ethnicity.

APPENDIX A: FACTORS ASSOCIATED WITH YOUNG-ONSET COLORECTAL CANCER IN STUDIES WITH A COMPARISON GROUP

	Author (year)																	
	Chiang (2003)	Chou (2011)	Ganapathi (2011)	Griffen (1991)	Karsten (2008)	Li (2011)	Liang (2002)	Marble (1992)	Mitry (2000)	O'Connell (2004)	Paraf[1] (2000)	Parramore (1998)	Quah (2007)	Savas (2007)	Schellerer (2012)	Wang (2010)	Yantiss (2009)	You[2] (2011)
Sociodemographic Characteristics																		
Race/Ethnicity[3]				+	+			*		+		+				+		
Female Sex	*	+				*	*	*		+	*	-		-	*	-	*	*
Family History[4]					+		*	*			+		*				*	
Symptoms at Diagnosis																		
Rectal Bleeding		*			*			*										
Abdominal Pain		+			*			*										
Change in bowel Habits		*			*													
Weight Loss		+			*													
Constipation		*																
Diarrhea		*																
Nausea/ Vomiting		*																
Bowel Obstruction					*								*					
Tenesmus		*																
Bloating		*																
Poor Performance Status[5]															+			
History of Crohn's Disease/ Ulce. Colitis									*									
Clinicopathologic Characteristics																		
Higher Stage at Diagnosis																		
Tumor Site	*	*	*	*	+ (right)	*	*	+ (right)	+ (left)		*	+ (rectum)	+ (left)	+ (right)	+ (rectum)		+ (rectum)	
Larger Tumor Size						*	*										*	*
Poorly/ Undifferent. Grade	+	+	+	+	+	*	+	+		+	*	+	*		*		*	*

	Author (year)																	
	Chiang (2003)	Chou (2011)	Ganapathi (2011)	Griffen (1991)	Karsten (2008)	Li (2011)	Liang (2002)	Marble (1992)	Mitry (2000)	O'Connell (2004)	Paraf[1] (2000)	Parramore (1998)	Quah (2007)	Savas (2007)	Schellerer (2012)	Wang (2010)	Yantiss (2009)	You[2] (2011)
Lymphovas. Invasion		*	+		*	*	*				*		*		*		+	
Perineural Invasion		+			*						*		*		*			
Venus Invasion															*		+	
Infiltrative Tumor Growth Pattern											*							
Gross Type (Ulcerative vs. Polypoid)						*	*											
PNCA[6] Count											*							
Lymphocyte Infiltration							*											
Synchronous Tumor		*					+						*					
Peritoneal Metastasis															*			
CEA						*	*											
MMP-2						*												
P27 KIP1						*												
EGFR																	*	
P53							-										*	
BRAF Mutation																	*	
KRAS Mutation							*										*	
Microsatellite Instability							+											
Treatment																		
Surgery					*				+	*								
Emergent Surgery		*	+												-			
Total Colectomy		*											*					+ (APR)
Resection Status									*				*		*			
Greater Number of Lymph Nodes Sampled		+	+										+			+		
Surgical Mortality		*					+		+						+			
Surgical Morbidity															+			
Chemotherapy		+							+				+		+			+

	Author (year)																	
	Chiang (2003)	Chou (2011)	Ganapathi (2011)	Griffen (1991)	Karsten (2008)	Li (2011)	Liang (2002)	Marble (1992)	Mitry (2000)	O'Connell (2004)	Paraf[1] (2000)	Parramore (1998)	Quah (2007)	Savas (2007)	Schellerer (2012)	Wang (2010)	Yantiss (2009)	You[2] (2011)
Radiation									+	+								+
Survival																		
Overall Survival		*	*	*	*			-	*		*	*	*		+	+		
Disease-Free Survival		-	*			*							*					*
Cancer-Specific Survival		-					*						*		*	*		*
Recurrence[7]																		+

NOTE: *, no association; +, positive association; -, negative association
[1]Paraf (2000) was a case-control study matched on sex, stage at diagnosis, and tumor site
[2]You (2011) was limited to rectal cancer only
[3]In all studies except Karsten (2008), black race was associated with early-onset CRC. In Karsten (2008), Hispanic and Asian race/ethnicity were positively associated
[4]Family history does not include Familial Adenomatous Polyposis (FAP) or Lynch Syndrome
[5]Performance status measured by American Society of Anesthesiologists (ASA) physical status classification system
[6]PCNA, proliferative cell nuclear antigen
[7]No recurrence or only distant recurrence more common in younger patients

APPENDIX B: CHARACTERISTICS OF PATIENTS WITH YOUNG-ONSET COLORECTAL CANCER IN STUDIES WITH NO COMPARISON GROUP

Author (year)	Study Population	Race	Sex	Age	Symptoms at Diagnosis	Family History	Stage at Diagnosis	Tumor Site	Histology	Tumor Grade	Overall Survival
Cusack (1996)	186 patients age ≤ 40 years treated for primary colorectal adenocarcinoma 1961-1990; MD Anderson Cancer Center	75.8% white, 15.6% black, 7.5% Hispanic, 1.1% Asian	48.9% male, 51.1% female	34.3 (median), range 14-40	Abdominal pain (30.5%), rectal bleeding (43.7%), back pain (2.3%), nausea/vomiting (2.9%), weight loss (1.7%), bowel obstruction (5.2%)	2 (1.1%) FAP, 4 (2.2%) HNPCC	8.1% stage I, 26.3% stage II, 31.7% stage III, 33.9% stage IV	41.6% rectum, 21.1% sigmoid colon, 9.3% left colon, 10.3% transverse colon, 15.7% right colon, 2.2% multiple sites	78.6% intestinal type adenocarcinoma, 10.3% mucinous adenocarcinoma, 11.1% signet ring cell	12.3% well-differentiated, 46.7% moderately differentiated, 40.9% poorly differentiated	43% (5-year OS), 35 months (median survival time)
Dozois (2008)	1025 patients age ≤ 50 years diagnosed with primary CRC 1976-2002; no history of FAP, HNPCC, or IBD; Mayo Clinic		57.1% male, 42.9% female	42.4 (mean), st dev 6.4	Rectal bleeding (50.7%), abdominal pain (32.5%), change in bowel habits (18.0%), diarrhea (13.2%), rectal pain (7.0%), bloating (7.3%), constipation (8.4%), weight loss (13.2%), nausea/vomiting (6.6%)	181 (17.7%) first-degree relative with CRC, 79 (7.7%) second-degree relative with CRC	13.4% stage I, 20.4% stage II, 32.1% stage III, 34.0% stage IV	49.1% rectum, 29.1% left colon, 21.9% right colon	11% mucinous adenocarcinoma, 2% signet ring cell	2.0% grade I, 53.6% grade 2, 34.0% grade 3, 7.2% grade 4, 2.9% unknown grade	
Fancher (2011)[1]	45 patients age ≤ 50 years underwent CRC surgery 1998-2007; no history of FAP or UC; Waterbury, CT	66.7% white, 33.3% black	43% male, 57% female	43.6 (mean)		17 (37.8%) with family history of CRC (not defined)	31.9% stage I, 17.0% stage II, 23.4% stage III, 27.7% stage IV	53.3% left colon, 48.9% right colon			28.6 months (mean survival time)
Lee (1994)	62 patients age ≤ 40 years diagnosed with CRC 1968-1991; Oregon Health Science University/ Portland VA			34.5 (mean) range 18-40	Pain (78%), bleeding (70%), weight loss >10 lbs (53%), obstruction (26%), perforation (5%)	5 (8%) FAP	8% Dukes A, 20% Dukes B, 23% Dukes C, 48% Dukes D	28% right colon, 14% transverse colon, 21% left colon, 26% rectum, 9% multiple sites	22.6% mucinous adenocarcinoma		67.8% (5-year OS); by stage, 5-year OS 100% Dukes A, 85% Dukes B, 40% Dukes C, 7% Dukes D

Author (year)	Study Population	Race	Sex	Age	Symptoms at Diagnosis	Family History	Stage at Diagnosis	Tumor Site	Histology	Tumor Grade	Overall Survival
Leff (2007)	49 patients age ≤ 40 years diagnosed with CRC 1982-1992; United Kingdom		57% male, 43% female	32 (mean)	Rectal bleeding (69%), change in bowel habits (51%), abdominal pain (34%), weight loss (8%), anemia (8%)	2 (4%) FAP; 2 (4%) first-degree relative with CRC	14% Dukes A, 24% Dukes B, 44% Dukes C, 16% Dukes D	65% rectum, 12% right colon, 4% transverse colon, 2% left colon, 16% sigmoid colon		12% well-differentiated, 59.1% moderately differentiated, 22% poorly differentiated	58% (5-year OS), 46% (10-year OS)
Lin (2005)	45 patients age ≤ 39 years diagnosed with CRC 1992-2002; Taipei		49% male, 51% female	32 (mean), range 18-39	Abdominal pain (53%), bloody stools (24%), change in bowel habits (13%), protruding mass (6%), obstruction (2%)	2 (4.4%) first-degree and second-degree relatives with CRC	11.1% Dukes B, 24.4% Dukes C, 66.7% Dukes D	36% right colon, 23% left colon, 24% rectosigmoid, 17% transverse colon	24% mucinous adenocarcinoma, 76% typical adenocarcinoma	49% well-differentiated, 25% poorly differentiated	34% (2-year OS), 9% (5-year OS)
MacGillivray (1991)	50 patients age ≤ 39 years diagnosed with CRC 1962-1988; National Naval Medical Center		66% male, 34% female	20 (mean), 30.5 (median), range 19-39	Rectal bleeding (54%), abdominal pain (44%), change in bowel pattern (16%), weight loss (16%)	1 (2%) HNPCC; 8 (16%) family history (not defined)	4% Dukes A, 26% Dukes B, 42% Dukes C, 24% Duke D	36% rectum, 2% transverse colon, 31% right colon, 31% left colon			43% (5-year OS), 35% (10-year OS), 28 months (median survival time)
Makela (2010)	59 patients age ≤ 40 years diagnosed with CRC 1984-2003; Oulu, Finland		49.2% male, 50.8% female	36 (mean), range 23-40	Rectal bleeding, change in bowel habits, abdominal pain, obstruction (no prevalence given)	11 (19%) HNPCC, 3 (5%) FAP	23% Dukes A, 34% Dukes B, 16% Dukes C, 27% Dukes D	39% proximal colon, 61% distal colon	39% mucinous adenocarcinoma	29% grade I, 44% grade II, 27% grade III	59% (5-year OS), 75 months (mean survival time)
Minardi (1998)	37 patients age ≤ 40 years treated for CRC 1976-1997; Louisiana State University	35% white; 65% black	59% male, 41% female	30.5 (mean), range 11-40	Abdominal pain (65%), weight loss (57%), rectal bleeding (49%), nausea and vomiting (35%), constipation (24%), narrow caliber stools 19%)	1 (2.7%) family history (not defined)	2.7% Dukes A, 35% Dukes B, 38% Dukes C, 22% Dukes D (1 patient had carcinoma in situ)	21.6% cecum, 2.7% right colon, 13.5% hepatic flexure, 10.8% mid-transverse colon, 10.8% splenic flexure, 10.8% left colon, 16.2% sigmoid colon, 2.7% rectosigmoid junction, 10.8% rectum	42% mucinous adenocarcinoma (3 had additional signet ring cell features); 58% typical adenocarcinoma	15% well-differentiated, 38% moderately differentiated, 46% poorly differentiated	26% (5-year OS)

Author (year)	Study Population	Race	Sex	Age	Symptoms at Diagnosis	Family History	Stage at Diagnosis	Tumor Site	Histology	Tumor Grade	Overall Survival
Palmer (1991)	105 patients age ≤ 39 years treated for CRC 1973-1985; Roswell Park Cancer Institute		49% male, 51% female	32 (mean), 34 (median), range 14-39		4 (3.8%) FAP	0.9% Dukes A, 18.1% Dukes B, 50% Dukes C, 33% Dukes D				42 months (mean survival time)
Yilmazlar (1995)	46 patients age ≤ 40 years underwent CRC surgery 1986-1993; no history FAP or HPNCC; Turkey		50% male, 50% female	33 (mean), range 16-40	Rectal bleeding (78%), abdominal pain (28%), weight loss (26%), change in bowel habits (22%), vomiting (7%)	6 (13%) family history (not defined)	4.3% stage I, 19.6% stage II, 43.5% stage III, 32.6% stage IV	69% rectum, 11% sigmoid colon, 9% left colon, 11% right colon	28% mucinous adenocarcinoma	15% well-differentiated, 37% moderately differentiated, 20% poorly differentiated	43.5% (5-year OS); 21.4 months (median survival time)

Abbreviations: CRC, colorectal cancer; FAP, familial adenomatous polyposis; HNPCC, hereditary non-polyposis colorectal cancer; UC, ulcerative colitis; IBD, irritable bowel syndrome; CEA, carcinoembryonic antigen

[1]Results reported stratified by sex in Fancher (2011) but summarized across sex in table

APPENDIX C: EXAMPLE OF MEDICAL RECORD ABSTRACTION FORM USED FOR PATTERNS OF CARE STUDIES

SEER PATTERNS OF CARE STUDY - 2010
Leave no boxes blank

IDENTIFICATION/PATIENT DATA

A1. SEER PARTICIPANT

A2. CASE No.

A3. QUALITY CONTROL

A4. RECORD NUMBER

A5. SEQUENCE NUMBER

A6. PRIMARY SITE

A7. MORPHOLOGY HIST BEH GR

A8. DIAGNOSTIC CONFIRMATION

A9. HOSPITAL CODE

A10. INSURANCE STATUS
- No insurance
- Medicare
- Medicaid/Medicaid pending
- Private/HMO/IPA Plan/Managed Care
- Tricare/VA/Military
- IHS
- Other, specify _____

A11. Treatment protocol registration

A12. Protocol sponsor and number

A13. Therapy verified with physician

A14. Height/Weight Ht. Units _____

Wt. Units _____

COLORECTAL CANCER

B1. Perforation

B2. Obstruction

B3. KRAS mutation

B4. BRAF status

B5. Microsatellite instability

B6. Laparoscopic surgery

B7. Date of cancer-directed surgery to primary site
month day year

B8. Number of positive lymph nodes

B9. Number of lymph nodes examined

B10. Extension of primary tumor

B11. Pathological margins

B12. Date radiation began
month day year

B13. Radiation sequence with surgery

B14. Radiation sequence with systemic therapy

B15. Systemic therapy sequence with surgery

SYSTEMIC THERAPY AGENTS / REGIMES

	month	day	year
B16. FOLFOX			
B17. FOLFIRI			
B18. 5-fluorouracil (5-FU)			
B19. Bevacizumab (Avastin)			
B20. Capecitabine (Xeloda)			
B21. Cetuximab (Erbitux)			
B22. Folinic acid (Leucovorin)			
B23. Irinotecan (CPT-11, Camptosar)			
B24. Levamisole (Ergamisol)			
B25. Oxaliplatin (Eloxatin)			
B26. Panitumumab (Vetibux)			
B27. Other, specify _____			

CODING INFORMATION

Coder ID

Date Abstracted
month day year

CO-MORBID CONDITIONS

List all diagnoses on the second page.

PLEASE TURN OVER

APPENDIX D: EXAMPLE OF PHYSICIAN VERIFICATION FORM FOR PATTERNS OF CARE STUDIES

Patterns of Care Study
Physician Verification Form

Patient Name _____

Patient Identification No._____ Physician _____

Date of Initial Diagnosis_____ Type of Cancer COLORECTAL

PLEASE DESCRIBE ALL CANCER-DIRECTED THERAPY WHICH HAS BEEN GIVEN TO THE PATIENT, REGARDLESS OF WHERE, WHEN OR BY WHOM THE THERAPY WAS ADMINISTERED. Please respond even if the patient did not receive cancer-targeted treatment. If you did not see this patient please go to question # 6.

1. Was this patient actively enrolled in an active/open clinical trial? ____Yes
   ____ No (SKIP TO Q3)

2. Please provide the name of the clinical trial sponsor and the clinical trial number.

   Sponsor (example: SWOG) _____

   Number (example: 8711) _____

3. Did the patient receive RADIATION therapy?

   ____Yes Date of radiation ____ / ____ / _____
   ____ No
   ____ Refused
   ____ Unknown

4. Did this patient receive any SYSTEMIC THERAPY AGENTS?

   ____Yes (Please mark all that apply in Q.5)
   ____Patient/Guardian refused ALL systemic therapy (Skip to Q.6)
   ____No (Skip to Q.6)
   ____Unknown (Skip to Q.6)

5. Please indicate whether or not each of the agents listed below was given and the date first given or refused.

FOLFOX	Yes (Date ____/____/____)	No	Patient/guardian refused
FOLFIRI	Yes (Date ____/____/____)	No	Patient/guardian refused
5-Fluorouracil	Yes (Date ____/____/____)	No	Patient/guardian refused
Bevacizumab	Yes (Date ____/____/____)	No	Patient/guardian refused
Capecitabine	Yes (Date ____/____/____)	No	Patient/guardian refused
Cetuximab	Yes (Date ____/____/____)	No	Patient/guardian refused

(TURN OVER)

REFERENCES

1. Davis DM, Marcet JE, Frattini JC, et al: Is it time to lower the recommended screening age for colorectal cancer? J Am Coll Surg 213:352-61, 2011.

2. Siegel RL, Jemal A, Ward EM: Increase in incidence of colorectal cancer among young men and women in the United States. Cancer Epidemiol Biomarkers Prev 18:1695-8, 2009.

3. O'Connell JB, Maggard MA, Liu JH, et al: Rates of colon and rectal cancers are increasing in young adults. Am Surg 69:866-72, 2003.

4. Meyer JE, Narang T, Schnoll-Sussman FH, et al: Increasing incidence of rectal cancer in patients aged younger than 40 years: an analysis of the surveillance, epidemiology, and end results database. Cancer 116:4354-9, 2010.

5. Fairley TL, Cardinez CJ, Martin J, et al: Colorectal cancer in U.S. adults younger than 50 years of age, 1998-2001. Cancer 107:1153-61, 2006.

6. Cress RD, Morris C, Ellison GL, et al: Secular changes in colorectal cancer incidence by subsite, stage at diagnosis, and race/ethnicity, 1992-2001. Cancer 107:1142-52, 2006.

7. Bailey CE, Hu CH, You N, et al: Disparities in age-related incidence of colon and rectal cancer in the United States, 1975-2010. J Clin Oncol 32:suppl; abstr 6579, 2014.

8. Siegel R, Desantis C, Jemal A: Colorectal cancer statistics, 2014. CA Cancer J Clin 64:104-17, 2014.

9. Edwards BK, Ward E, Kohler BA, et al: Annual report to the nation on the status of cancer, 1975-2006, featuring colorectal cancer trends and impact of interventions (risk factors, screening, and treatment) to reduce future rates. Cancer 116:544-73, 2010.

10. Levin B, Lieberman DA, McFarland B, et al: Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology*†. CA Cancer J Clin 58:130-160, 2008.

11. Zauber AG, Lansdorp-Vogelaar I, Knudsen AB, et al: Evaluating test strategies for colorectal cancer screening: a decision analysis for the U.S. Preventive Services Task Force. Ann Intern Med 149:659-69, 2008.

12. Schellerer VS, Hohenberger W, Croner RS: Is it time to lower the recommended screening age for colorectal cancer? J Am Coll Surg 214:377-8; author reply 378-9, 2012.

13. Limburg PJ, Harmsen WS, Chen HH, et al: Prevalence of alterations in DNA mismatch repair genes in patients with young-onset colorectal cancer. Clin Gastroenterol Hepatol 9:497-502, 2011.

14. Samowitz WS, Curtin K, Lin HH, et al: The colon cancer burden of genetically defined hereditary nonpolyposis colon cancer. Gastroenterology 121:830-8, 2001.

15. Marble K, Banerjee S, Greenwald L: Colorectal carcinoma in young patients. J Surg Oncol 51:179-82, 1992.

16. O'Connell JB, Maggard MA, Livingston EH, et al: Colorectal cancer in the young. Am J Surg 187:343-8, 2004.

17. You YN, Xing Y, Feig BW, et al: Young-onset colorectal cancer: is it time to pay attention? Archives of internal medicine 172:287-289, 2012.

18. Parramore JB, Wei JP, Yeh KA: Colorectal cancer in patients under forty: presentation and outcome. Am Surg 64:563-7; discussion 567-8, 1998.

19. Schellerer VS, Merkel S, Schumann SC, et al: Despite aggressive histopathology survival is not impaired in young patients with colorectal cancer : CRC in patients under 50 years of age. Int J Colorectal Dis 27:71-9, 2012.

20. Yantiss RK, Goodarzi M, Zhou XK, et al: Clinical, pathologic, and molecular features of early-onset colorectal carcinoma. Am J Surg Pathol 33:572-82, 2009.

21. Karsten B, Kim J, King J, et al: Characteristics of colorectal cancer in young patients at an urban county hospital. Am Surg 74:973-6, 2008.

22. Savas N, Dagli U, Akbulut S, et al: Colorectal cancer localization in young patients: should we expand the screening program? Dig Dis Sci 52:798-802, 2007.

23. Chou CL, Chang SC, Lin TC, et al: Differences in clinicopathological characteristics of colorectal cancer between younger and elderly patients: an analysis of 322 patients from a single institution. Am J Surg 202:574-82, 2011.

24. Ganapathi S, Kumar D, Katsoulas N, et al: Colorectal cancer in the young: trends, characteristics and outcome. Int J Colorectal Dis 26:927-34, 2011.

25. Li M, Li JY, Zhao AL, et al: Do young patients with colorectal cancer have a poorer prognosis than old patients? J Surg Res 167:231-6, 2011.

26. Griffin PM, Liff JM, Greenberg RS, et al: Adenocarcinomas of the colon and rectum in persons under 40 years old. A population-based study. Gastroenterology 100:1033-40, 1991.

27. Chiang JM, Chen MC, Changchien CR, et al: Favorable influence of age on tumor characteristics of sporadic colorectal adenocarcinoma: patients 30 years of age or younger may be a distinct patient group. Dis Colon Rectum 46:904-10, 2003.

28. Liang JT, Huang KC, Cheng AL, et al: Clinicopathological and molecular biological features of colorectal cancer in patients less than 40 years of age. Br J Surg 90:205-14, 2003.

29. You YN, Dozois EJ, Boardman LA, et al: Young-onset rectal cancer: presentation, pattern of care and long-term oncologic outcomes compared to a matched older-onset cohort. Ann Surg Oncol 18:2469-76, 2011.

30. O'Connell JB, Maggard MA, Liu JH, et al: Are survival rates different for young and older patients with rectal cancer? Dis Colon Rectum 47:2064-9, 2004.

31. Wang L, Hollenbeak CS, Stewart DB: Node yield and node involvement in young colon cancer patients: is there a difference in cancer survival based on age? J Gastrointest Surg 14:1355-61, 2010.

32. O'Connell JB, Maggard MA, Liu JH, et al: Do young colon cancer patients have worse outcomes? World J Surg 28:558-62, 2004.

33. Quah HM, Joseph R, Schrag D, et al: Young age influences treatment but not outcome of colon cancer. Ann Surg Oncol 14:2759-65, 2007.

34. Paraf F, Jothy S: Colorectal cancer before the age of 40: a case-control study. Dis Colon Rectum 43:1222-6, 2000.

35. Phipps AI, Limburg PJ, Baron JA, et al: Association between molecular subtypes of colorectal cancer and patient survival. Gastroenterology 148:77-87.e2, 2015.

36. Arends MJ: Pathways of colorectal carcinogenesis. Appl Immunohistochem Mol Morphol 21:97-102, 2013.

37. Weisenberger DJ, Levine AJ, Long TI, et al: Association of the colorectal CpG island methylator phenotype with molecular features, risk factors, and family history. Cancer Epidemiol Biomarkers Prev 24:512-9, 2015.

38. Weisenberger DJ, Siegmund KD, Campan M, et al: CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. Nat Genet 38:787-793, 2006.

39. You N, Chang K, Chen K, et al: Somatic mutations in young-onset colorectal cancer unrelated to hereditary syndromes: a comparitive study using high-depth targeted exome sequencing. J Clin Oncol 33:abstr 623, 2015.

40. Abbott AM, Kothari N, Teer JK, et al: Genetic analysis of colorectal cancers in young patients. J Clin Oncol 33:abstr 632, 2015.

41. Murphy CC, Harlan LC, Lund J, et al: Patterns of colorectal cancer care in the United States: 1990-2010. J Natl Cancer Inst 107(10):pii:djv198, 2015.

42. Benson AB, 3rd, Venook AP, Bekaii-Saab T, et al: Colon cancer, version 3.2014. J Natl Compr Canc Netw 12:1028-59, 2014.

43.Schrag D, Cramer LD, Bach PB, et al: Age and adjuvant chemotherapy use after surgery for stage III colon cancer. J Natl Cancer Inst 93:850-7, 2001.

44. Kirkpatrick HM, Aitelli CL, Qin H, et al: Referral patterns and adjuvant chemotherapy use in patients with stage II colon cancer. Clin Colorectal Cancer 9:150-6, 2010.

45. Wirtzfeld DA, Mikula L, Gryfe R, et al: Concordance with clinical practice guidelines for adjuvant chemotherapy in patients with stage I-III colon cancer: experience in 2 Canadian provinces. Can J Surg 52:92-7, 2009.

46. Mitry E, Benhamiche AM, Jouve JL, et al: Colorectal adenocarcinoma in patients under 45 years of age: comparison with older patients in a well-defined French population. Dis Colon Rectum 44:380-7, 2001.

47. Kneuertz PJ, Chang GJ, Hu CY, et al: Overtreatment of young adults with colon cancer: more intense treatments with unmatched survival gains. JAMA Surg 150:402-9, 2015.

48. Ahnen DJ, Wade SW, Jones WF, et al: The increasing incidence of young-onset colorectal cancer: a call to action. Mayo Clin Proc 89:216-24, 2014.

49. Inra JA, Syngal S: Colorectal cancer in young adults. Dig Dis Sci 60:722-33, 2015.

50. Larsson SC, Wolk A: Meat consumption and risk of colorectal cancer: a meta-analysis of prospective studies. Int J Cancer 119:2657-2664, 2006.

51. Chan DS, Lau R, Aune D, et al: Red and processed meat and colorectal cancer incidence: meta-analysis of prospective studies. PloS one 6:e20456, 2011.

52. Larsson SC, Wolk A: Obesity and colon and rectal cancer risk: a meta-analysis of prospective studies. Am J Clin Nutr 86:556-65, 2007.

53. Moghaddam AA, Woodward M, Huxley R: Obesity and risk of colorectal cancer: a meta-analysis of 31 studies with 70,000 events. Cancer Epidemiol Biomarkers Prev 16:2533-47, 2007.

54. Hong S, Cai Q, Chen D, et al: Abdominal obesity and the risk of colorectal adenoma: a meta-analysis of observational studies. Eur J Cancer Prev 21:523-31, 2012.

55. Wolin K, Yan Y, Colditz G, et al: Physical activity and colon cancer prevention: a meta-analysis. Br J Cancer 100:611-616, 2009.

56. Liang PS, Chen TY, Giovannucci E: Cigarette smoking and colorectal cancer incidence and mortality: systematic review and meta-analysis. Int J Cancer 124:2406-15, 2009.

57. Slattery ML: Physical activity and colorectal cancer. Sports Med 34:239-252, 2004.

58. Guthrie JF, Lin B-H, Frazao E: Role of food prepared away from home in the American diet, 1977-78 versus 1994-96: changes and consequences. J Nutr Educ Behav 34:140-150, 2002.

59. Flegal KM, Carroll MD, Kuczmarski RJ, et al: Overweight and obesity in the United States: prevalence and trends, 1960-1994. Int J Obes Rel Metabol Disord 22:39-47, 1998.

60. Ogden CL, Carroll MD, Kit BK, et al: Prevalence of childhood and adult obesity in the United States, 2011-2012. JAMA 311:806-14, 2014.

61. Li C, Ford ES, McGuire LC, et al: Increasing trends in waist circumference and abdominal obesity among US adults. Obesity 15:216-24, 2007.

62. Eheman C, Henley SJ, Ballard-Barbash R, et al: Annual Report to the Nation on the status of cancer, 1975-2008, featuring cancers associated with excess weight and lack of sufficient physical activity. Cancer 118:2338-2366, 2012

63. Health UDo, Services H: Results from the 2010 National Survey on Drug Use and Health: Summary of national findings. Substance Abuse and Mental Health Services Administration: Rockville, MD, USA, 2011.

64. Seeff LC, Richards TB, Shapiro JA, et al: How many endoscopies are performed for colorectal cancer screening? Results from CDC's survey of endoscopic capacity. Gastroenterology 127:1670-7, 2004.

65. Peery AF, Dellon ES, Lund J, et al: Burden of gastrointestinal disease in the United States: 2012 update. Gastroenterology 143:1179-87.e1-3, 2012.

66. Klabunde CN, Cronin KA, Breen N, et al: Trends in colorectal cancer test use among vulnerable populations in the United States. Cancer Epidemiol Biomarkers Prev 20:1611-21, 2011.

67. Shapiro JA, Klabunde CN, Thompson TD, et al: Patterns of colorectal cancer test use, including CT colonography, in the 2010 National Health Interview Survey. Cancer Epidemiol Biomarkers Prev 21:895-904, 2012.

68. Bailey CE, Hu CY, You YN, et al: Increasing disparities in the age-related incidences of colon and rectal cancers in the United States, 1975-2010. JAMA Surg 150:17-22, 2015.

69. Potosky AL, Harlan LC, Kaplan RS, et al: Age, sex, and racial differences in the use of standard adjuvant therapy for colorectal cancer. J Clin Oncol 20:1192-202, 2002.

70. Cronin DP, Harlan LC, Potosky AL, et al: Patterns of care for adjuvant therapy in a random population-based sample of patients diagnosed with colorectal cancer. Am J Gastroenterol 101:2308-18, 2006.

71. Compton CC: Updated protocol for the examination of specimens from patients with carcinomas of the colon and rectum, excluding carcinoid tumors, lymphomas, sarcomas, and tumors of the vermiform appendix: a basis for checklists. Cancer Committee. Arch Pathol Lab Med 124:1016-25, 2000.

72. Compton CC, Fielding LP, Burgart LJ, et al: Prognostic factors in colorectal cancer. College of American Pathologists Consensus Statement 1999. Arch Pathol Lab Med 124:979-94, 2000.

73. Byrd JC, Bresalier RS: Mucins and mucin binding proteins in colorectal cancer. Cancer Metastasis Rev 23:77-99, 2004.

74. Ogino S, Brahmandam M, Cantor M, et al: Distinct molecular features of colorectal carcinoma with signet ring cell component and colorectal carcinoma with mucinous component. Mod Pathol 19:59-68, 2006.

75. Plesec TP, Hunt JL: KRAS mutation testing in colorectal cancer. Adv Anat Pathol 16:196-203, 2009.

76. Feigelson HS, Goddard KA, Johnson MA, et al: Reliability of KRAS mutation testing in metastatic colorectal cancer patients across five laboratories. BMC Res Notes 5:196, 2012.

77. Jarry A, Masson D, Cassagnau E, et al: Real-time allele-specific amplification for sensitive detection of the BRAF mutation V600E. Mol Cell Probes 18:349-52, 2004.

78. Umar A, Boland CR, Terdiman JP, et al: Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. J Natl Cancer Inst 96:261-8, 2004.

79. Yost K, Perkins C, Cohen R, et al: Socioeconomic status and breast cancer incidence in California for different race/ethnic groups. Cancer Causes Control 12:703-11, 2001.

80. Yu M, Tatalovich Z, Gibson JT, et al: Using a composite index of socioeconomic status to investigate health disparities while protecting the confidentiality of cancer registry data. Cancer Causes Control 25:81-92, 2014.

81. Yang Y, Land KC: Age-period-cohort analysis: new models, methods, and empirical applications, CRC Press, 2013.

82. Murphy CC, Yang YC, Shaheen NJ, et al: An age-period-cohort analysis of obesity and incident esophageal adenocarcinoma. Dis Esoph, under review.

83. American College of Sports Medicine: ACSM's guidelines for exercise testing and prescription, Lippincott Williams & Wilkins, 2013.

84. Pate RR, Pratt M, Blair SN, et al: Physical activity and public health: a recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. JAMA 273:402-407, 1995.

85. Mason WM, Fienberg SE: Cohort analysis in social research: beyond the identification problem. 1985.

86. Yang Y, Land KC: A mixed-models approach to age-period-cohort analysis of repeated cross-section surveys: Trends in verbal test scores. Sociol Methodol 36:75-97, 2006.

87. Yang Y, Land KC: Age–period–cohort analysis of repeated cross-section surveys: Fixed or random effects? Sociol Methods Res 36:297-326, 2008.

88. Raudenbush SW, Bryk AS: Hierarchical linear models: Applications and data analysis methods, Sage, 2002.

89. Chubak J, Pocobelli G, Weiss NS: Tradeoffs between accuracy measures for electronic health care data algorithms. J Clin Epidemiol 65:343-349.e2, 2012.

90. El-Serag HB, Petersen L, Hampel H, et al: The use of screening colonoscopy for patients cared for by the Department of Veterans Affairs. Arch Intern Med 166:2202-8, 2006.

91. Fisher DA, Grubber JM, Castor JM, et al: Ascertainment of colonoscopy indication using administrative data. Dig Dis Sci 55:1721-5, 2010.

92. Lieu CH, Renfro LA, de Gramont A, et al: Association of age with survival in patients with metastatic colorectal cancer: analysis from the ARCAD Clinical Trials Program. J Clin Oncol 32:2975-84, 2014.

93. Murphy CC, Harlan LC, Lund JL, et al: Patterns of Colorectal Cancer Care in the United States: 1990-2010. J Natl Cancer Inst 107(10):pii:djv198, 2015.

94. Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov) SEER*Stat Database: Incidence - SEER 9 Regs Research Data, Nov 2014 Sub (1973-2012) <Katrina/Rita Population Adjustment> - Linked To County Attributes - Total U.S., 1969-2013 Counties, National Cancer Institute, DCCPS, Surveillance Research Program, Surveillance Systems Branch, released April 2015, based on the November 2014 submission.

95. Howe HL, Wu X, Ries LA, et al: Annual report to the nation on the status of cancer, 1975–2003, featuring cancer among US Hispanic/Latino populations. Cancer 107:1711-1742, 2006.

96. Ennis SR, Ríos-Vargas M, Albert NG, et al: The hispanic population: 2010, US Department of Commerce, Economics and Statistics Administration, US Census Bureau, 2011.

97. American Cancer Society: Cancer Facts & Figures for Hispanics, American Cancer Society., 2009.

98. Pinheiro PS, Sherman RL, Trapido EJ, et al: Cancer incidence in first generation U.S. Hispanics: Cubans, Mexicans, Puerto Ricans, and new Latinos. Cancer Epidemiol Biomarkers Prev 18:2162-9, 2009.

99. Singh KE, Taylor TH, Pan CG, et al: Colorectal cancer incidence among young adults in California. J Adolesc Young Adult Oncol 3:176-184, 2014.

100. Sinicrope FA, Shi Q, Smyrk TC, et al: Molecular markers identify subtypes of stage III colon cancer associated with patient outcomes. Gastroenterology 148:88-99, 2015.

101. Comprehensive molecular characterization of human colon and rectal cancer. Nature 487:330-7, 2012.

102. Guinney J, Dienstmann R, Wang X, et al: The consensus molecular subtypes of colorectal cancer. 21:1350-6, 2015.

103. Yoon HH, Shi Q, Alberts SR, et al: Racial differences in BRAF/KRAS mutation rates and survival in stage III colon cancer patients. J Natl Cancer Inst 107, 2015.

104. Kumar K, Brim H, Giardiello F, et al: Distinct BRAF (V600E) and KRAS mutations in high microsatellite instability sporadic colorectal cancer in African Americans. Clin Cancer Res 15:1155-61, 2009.

105. Sylvester BE, Huo D, Khramtsov A, et al: Molecular analysis of colorectal tumors within a diverse patient cohort at a single institution. Clin Cancer Res 18:350-9, 2012.

106. Siegel RL, Jemal A, Thun MJ, et al: Trends in the incidence of colorectal cancer in relation to county-level poverty among blacks and whites. J Natl Med Assoc 100:1441-4, 2008.

107. Ward E, Jemal A, Cokkinides V, et al: Cancer disparities by race/ethnicity and socioeconomic status. CA Cancer J Clin 54:78-93, 2004.

108. Doubeni CA, Field TS, Buist DS, et al: Racial differences in tumor stage and survival for colorectal cancer in an insured population. Cancer 109:612-20, 2007.

109. Du XL, Fang S, Vernon SW, et al: Racial disparities and socioeconomic status in association with survival in a large population-based cohort of elderly patients with colon cancer. Cancer 110:660-669, 2007.

110. Ben Q, An W, Jiang Y, et al: Body mass index increases risk for colorectal adenomas based on meta-analysis. Gastroenterology 142:762-72, 2012.

111. Fienberg SE, Mason WM: Specification and implementation of age, period and cohort models, Springer, 1985.

112. Vernon SW: Participation in colorectal cancer screening: a review. J Natl Cancer Inst 89:1406-22, 1997.

113. Shapiro JA, Seeff LC, Thompson TD, et al: Colorectal cancer test use from the 2005 National Health Interview Survey. Cancer Epidemiol Biomarkers Prev 17:1623-30, 2008.

114. Seeff LC, Nadel MR, Klabunde CN, et al: Patterns and predictors of colorectal cancer test use in the adult U.S. population. Cancer 100:2093-103, 2004.

115. Vogelaar I, van Ballegooijen M, Schrag D, et al: How much can current interventions reduce colorectal cancer mortality in the U.S.? Mortality projections for scenarios of risk-factor modification, screening, and treatment. Cancer 107:1624-33, 2006.

116. Zauber AG, Winawer SJ, O'Brien MJ, et al: Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. N Engl J Med 366:687-96, 2012.

117. Meester RG, Doubeni CA, Lansdorp-Vogelaar I, et al: Colorectal cancer deaths attributable to nonuse of screening in the United States. Ann Epidemiol 25:208-213.e1, 2015.

118. Cronin KA, Krebs-Smith SM, Feuer EJ, et al: Evaluating the impact of population changes in diet, physical activity, and weight status on population risk for colon cancer (United States). Cancer Causes Control 12:305-16, 2001.

119. Rex DK, Johnson DA, Anderson JC, et al: American College of Gastroenterology guidelines for colorectal cancer screening 2009 [corrected]. Am J Gastroenterol 104:739-50, 2009.

120. Agrawal S, Bhupinderjit A, Bhutani MS, et al: Colorectal cancer in African Americans. Am J Gastroenterol 100:515-23; discussion 514, 2005.

121. Roosendaal R, Kuipers EJ, Buitenwerf J, et al: Helicobacter pylori and the birth cohort effect: evidence of a continuous decrease of infection rates in childhood. Am J Gastroenterol 92:1480-1482, 1997.

122. Grad YH, Lipsitch M, Aiello AE: Secular trends in Helicobacter pylori seroprevalence in adults in the United States: evidence for sustained race/ethnic disparities. Am J Epidemiol 175:54-59, 2012.

123. Islami F, Kamangar F: Helicobacter pylori and esophageal cancer risk: a meta-analysis. Cancer Prev Res 1:329-338, 2008.

124. World Health Organization: IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Schistosomes, Liver Flukes and Helicobacter Pylori, International Agency for Research on Cancer, 1994.

125. Blaser MJ: Disappearing microbiota: Helicobacter pylori protection against esophageal adenocarcinoma. Cancer Prev Res 1:308-11, 2008.

126. Flegal KM, Carroll MD, Ogden CL, et al: Prevalence and trends in obesity among US adults, 1999-2008. JAMA 303:235-41, 2010.

127. Murphy CC, Martin CF, Sandler RS: Racial differences in obesity measures and risk of colorectal adenomas in a large screening population. Nutr Cancer 67:98-104, 2015.

128. Thompson CL, Berger NA, Chak A, et al: Racial differences in measures of obesity and risk of colon adenoma. Obesity 20:673-7, 2012.

129. Paskett ED, Reeves KW, Rohan TE, et al: Association between cigarette smoking and colorectal cancer in the Women's Health Initiative. J Natl Cancer Inst 99:1729-35, 2007.

130. Croswell JM, Ransohoff DF, Kramer BS: Principles of cancer screening: lessons from history and study design issues. Semin Oncol 37:202-15, 2010.

131.Dennis LK, Resnick MI: Analysis of recent trends in prostate cancer incidence and mortality. Prostate 42:247-252, 2000.

132. Burrell HC, Sibbering DM, Wilson A, et al: Screening interval breast cancers: mammographic features and prognosis factors. Radiology 199:811-817, 1996.

133. Shih YC, Zhao L, Elting LS: Does Medicare coverage of colonoscopy reduce racial/ethnic disparities in cancer screening among the elderly? Health Aff 25:1153-62, 2006.

134. Luchtefeld MA, Kim DG: Colonoscopy in the office setting is safe, and financially sound ... for now. Dis Colon Rectum 49:377-81; discussion 381-2, 2006.

135. Vicari JJ: The future value of ambulatory endoscopy centers in the United States: challenges and opportunities. Gastrointest Endosc 76:400-5, 2012.

136. Holahan J: The 2007–09 recession and health insurance coverage. Health Affairs:1003, 2010.

137. Joseph DA, King JB, Miller JW, et al: Prevalence of colorectal cancer screening among adults—behavioral risk factor surveillance system, United States, 2010. MMWR Morb Mortal Wkly Rep 61:51-56, 2012.

138. Vital signs: colorectal cancer screening test use--United States, 2012. MMWR Morb Mortal Wkly Rep 62:881-8, 2013.

139. Murphy CC, Sandler RS, Grubber JM, et al: Underuse and overuse of colonoscopy for repeat screening and surveillance in the Veterans Health Administration. Clin Gastroenterol Hepatol, 14:436-222, 2016.

140. Cooper GS, Kou TD, Barnholtz Sloan JS, et al: Use of colonoscopy for polyp surveillance in Medicare beneficiaries. Cancer 119:1800-7, 2013.

141. Hol L, Sutradhar R, Gu S, et al: Repeat colonoscopy after a colonoscopy with a negative result in Ontario: a population-based cohort study. CMAJ Open 3:E244-50, 2015.

142. Singh A, Kuo YF, Goodwin JS: Many patients who undergo surgery for colorectal cancer receive surveillance colonoscopies earlier than recommended by guidelines. Clin Gastroenterol Hepatol 11:65-72.e1, 2013.

143. Salz T, Weinberger M, Ayanian JZ, et al: Variation in use of surveillance colonoscopy among colorectal cancer survivors in the United States. BMC Health Serv Res 10:256, 2010.

144. Goodwin JS, Singh A, Reddy N, et al: Overuse of screening colonoscopy in the Medicare population. Arch Intern Med 171:1335-43, 2011.

145. Sheffield KM, Han Y, Kuo YF, et al: Potentially inappropriate screening colonoscopy in Medicare patients: variation by physician and geographic region. JAMA Intern Med 173:542-50, 2013.

146. Cooper GS, Koroukian SM: Geographic variation among Medicare beneficiaries in the use of colorectal carcinoma screening procedures. Am J Gastroenterol 99:1544-50, 2004.

147. Fisher ES, Wennberg DE, Stukel TA, et al: The implications of regional variations in Medicare spending. Part 1: the content, quality, and accessibility of care. Ann Intern Med 138:273-87, 2003.

148. Fisher ES, Wennberg DE, Stukel TA, et al: The implications of regional variations in Medicare spending. Part 2: health outcomes and satisfaction with care. Ann Intern Med 138:288-98, 2003.

149. Adams KF, Johnson EA, Chubak J, et al: Development of an Algorithm to Classify Colonoscopy Indication from Coded Health Care Data. EGEMS (Wash DC) 3:1171, 2015.

150. Fassil H, Adams KF, Weinmann S, et al: Approaches for classifying the indications for colonoscopy using detailed clinical data. BMC Cancer 14:95, 2014.

151. Gupta S, Tong L, Anderson P, et al: Measurement of colorectal cancer test use with medical claims data in a safety-net health system. Am J Med Sci 345:99-103, 2013.

152. Haque R, Chiu V, Mehta KR, et al: An automated data algorithm to distinguish screening and diagnostic colorectal cancer endoscopy exams. J Natl Cancer Inst Monogr:116-8, 2005.

153. Ko CW, Dominitz JA, Neradilek M, et al: Determination of colonoscopy indication from administrative claims data. Med Care 52:e21-9, 2014.

154. Schenck AP, Klabunde CN, Warren JL, et al: Data sources for measuring colorectal endoscopy use among Medicare enrollees. Cancer Epidemiol Biomarkers Prev 16:2118-27, 2007.

155. Sewitch MJ, Jiang M, Joseph L, et al: Developing model-based algorithms to identify screening colonoscopies using administrative health databases. BMC Med Inform Decis Mak 13:45, 2013.

156. Rutter CM, Greenlee RT, Johnson E, et al: Prevalence of colonoscopy before age 50. Prev Med, 2015.

157. Schofield L, Watson N, Grieu F, et al: Population-based detection of Lynch syndrome in young colorectal cancer patients using microsatellite instability as the initial test. Int J Cancer 124:1097-102, 2009.

158. Draisma G, Etzioni R, Tsodikov A, et al: Lead time and overdiagnosis in prostate-specific antigen screening: importance of methods and context. J Natl Cancer Inst 101:374-83, 2009.

159. Etzioni R, Gulati R, Mallinger L, et al: Influence of study features and methods on overdiagnosis estimates in breast and prostate cancer screening. Ann Intern Med 158:831-8, 2013.

160. Abdelsattar ZM, Wong SL, Regenbogen SE, et al: Colorectal cancer outcomes and treatment patterns in patients too young for average-risk screening. Cancer, 2016.

161. Sabatino SA, White MC, Thompson TD, et al: Cancer screening test use - United States, 2013. MMWR Morb Mortal Wkly Rep 64:464-8, 2015.

162. Centers for Disease Control and Prevention: Vital signs: colorectal cancer screening test use--United States, 2012. MMWR. Morb Mortal Wkly Rep 62:881, 2013.

163. Mandel JS, Church TR, Ederer F, et al: Colorectal cancer mortality: effectiveness of biennial screening for fecal occult blood. J Natl Cancer Inst 91:434-7, 1999.

164. Hardcastle JD, Armitage NC, Chamberlain J, et al: Fecal occult blood screening for colorectal cancer in the general population: results of a controlled trial. Cancer 58:397-403, 1986.

165. Hardcastle JD, Thomas WM, Chamberlain J, et al: Randomised, controlled trial of faecal occult blood screening for colorectal cancer. Results for first 107,349 subjects. Lancet 1:1160-4, 1989.

166. Hardcastle JD, Chamberlain JO, Robinson MH, et al: Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. Lancet 348:1472-7, 1996.