2020

Epidemiology of Colorectal Cancer Comorbidities and Stage at Diagnosis, Survival, and Second Primary Malignancies in Kentucky, 2003-2016

Nikita Leigh Vundi
University of Kentucky, nikita.vundi@uky.edu

STUDENT AGREEMENT:

I represent that my capstone and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's capstone including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Nikita Leigh Vundi, Student

Steven T Fleming, Committee Chair

Richard Ingram, Director of Graduate Studies

ABSTRACT OF CAPSTONE

Nikita Leigh Vundi

The College of Public Health

University of Kentucky

2020

Epidemiology of Colorectal Cancer Comorbidities and Stage at Diagnosis, Survival, and Second Primary Malignancies in Kentucky, 2003-2016

A Capstone project submitted in partial fulfillment of the requirements for the degree of Doctor of Public Health in the College of Public Health
at the University of Kentucky

By:

_____

Nikita Leigh Vundi

_____

Lexington, Kentucky

Chair: (Dr. Steven Fleming)
Lexington, Kentucky

Co-Chair: (Dr. W. Jay Christian)
Lexington, Kentucky

Co-Chair: (Dr. Jaclyn McDowell)
Lexington, Kentucky

ABSTRACT OF CAPSTONE

Epidemiology of Colorectal Cancer Comorbidities and Stage at Diagnosis, Survival, and Second Primary Malignancies in Kentucky, 2003-2016

Background: Colorectal cancer (CRC) is the third most common type of cancer and the third most common cause of cancer death among men and women in the United States.[1-3] The American Cancer Society estimates that there will be 147,950 new cases of CRC and 53,200 CRC related deaths in the U.S. for the year 2020.[3] Kentucky CRC incidence for 2012-2016 was the highest in the nation, and the mortality rate for years 2013-2017 was ranked 5th in the nation.[4-6] Risk factors for CRC include lifestyle factors, genetics, and disease status (comorbidities and treatment).[2, 7] Diabetes has been found to be the most prevalent comorbidity among CRC patients, and the risk of developing CRC in patients with diabetes is 25% higher than those without diabetes.[8, 9]

Aim: The purpose of this study is to explore if comorbidities impacts CRC progression, CRC outcomes, and the development of second primary malignancy among CRC patients age 18 and older in Kentucky diagnosed between January 1, 2003 and December 31, 2016.

Methods: Two studies were performed using CRC data from Kentucky Cancer Registry, one was a retrospective cohort study and the other was a case control study. There were 20,571 cases included in the cohort study with the primary outcomes was all-cause mortality, CRC mortality, and second primary cancer. There were 18,170 total, 9,085 cases and controls in the second study. This study examined the geographical distribution of late-stage CRC and comorbidities.

Results Chapter 3: Logistic regression models show that comorbidities increased the odds of death or late-stage CRC. The Cox proportional hazard models of all-cause and CRC mortalities and second primary show that comorbidities, patient factors, and treatments can be protective or increase the

hazards of dying or having a second primary cancer. The Kaplan Meier curve demonstrates the survival of early-stage at diagnosis CRC versus late-stage at diagnosis CRC.

Results Chapter 4: The geographical distribution maps of the four positively associated morbidities (electrolyte disorders, liver disease, weight loss, and deficiency anemia) do not demonstrate any patterns resembling the cluster, the comorbidity distribution appears to be random. The map of comorbidities among CRC patients show that a large percentage experience a burden of two or more comorbidities.

Conclusion: The results indicate that comorbidities do play a role in the stage of CRC diagnosis, with the data showing greater odds of being diagnosed with early-stage cancer for many of the individual comorbidities. The space-time analysis found a significant high rate cluster of late-stage CRC, however, mapping the distribution of positively associated comorbidities did not demonstrate a pattern matching the cluster. Further research is needed to examine the impact of comorbidities and CRC stage at diagnosis.

KEYWORDS: (colorectal cancer, comorbidities, second primary cancers, mortality)

(Student's Signature) Nikita Vundi

(Date) 04/24/2020

Epidemiology of Colorectal Cancer Comorbidities and Stage at Diagnosis, Survival, and Second Primary Malignancies in Kentucky, 2003-2016

By

Nikita Leigh Vundi

2020

(Signature of Capstone Chair)

April 24, 2020

(Date)

(Signature of Director of Doctoral Studies)

4/24/20

(Date)

RULES FOR THE USE OF CAPSTONE PROJECTS

Unpublished capstone projects submitted for the Doctor's degree and deposited in the University of Kentucky library are as a rule open for inspection but are to be used only with due regard to the rights of the authors. Bibliographical references may be noted, but quotations or summaries of parts may be published only with the permission of the author, and with the usual scholarly acknowledgements.

Extensive copying or publication of the capstone project in whole or in part also requires the consent of the Dean of the College of Public Health of the University of Kentucky.

A library that borrows this capstone project for use by its patrons is expected to secure the signature of each user.

Name **Date**

_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____

Epidemiology of Colorectal Cancer Comorbidities and Stage at Diagnosis, Survival, and Second Primary Malignancies in Kentucky, 2003-2016

Nikita Leigh Vundi

College of Public Health

University of Kentucky

Table of Contents

List of Figures

Page

List of Tables

Page

Acknowledgements

This work would not have been possible first and foremost without the support of my family. I'd like to thank my loving and supportive husband, Abednego, our sons, Noah and Viktor, who fill my life with laughter, love, and encourage me to be a better person for them. My husband helped motivate me to apply for the epidemiology program, and he was right, this program was exactly where I needed to be. I'd also like to thank my parents for teaching me the importance of education, loving and serving our neighbors, and for their support and encouragement when I wasn't sure I could do this.

I am grateful for all of the wonderful faculty, staff, and students that have helped me along the way. The College of Public Health has some of the most phenomenal people working for them. I'd like that thank Laverne Carter for her reassurance, she may not have realized just how a small conversation with her convinced me that public health was my calling and that I could do this. I am so thankful for the pleasure of having wonderful professors that inspire students and provide an environment that not only furthers learning and invites questions, but is instrumental in illustrating their passion for public health, data, and research. I'd like to thank Dr. Erin Abner, Dr. Wayne Sanderson, Dr. Bin Huang, Dr. Sabrina Brown, Dr. Phillip Westgate, Dr. Steve Browning, and last but certainly not least my committee members, Dr. Jaclyn McDowell, Dr. W. Jay Christian, and Dr. Steve Fleming for being a support for student learning and especially shaping my academic career while at UK College of Public Health. Without the encouragement and support of Dr. Fleming, I am not sure I would have been able to finish. I am forever grateful for your guidance.

I'd also like to thank UK's Center for Health Services Research for their support while working and obtaining my degree. My bosses, Dr. Mark Williams, Dr. Jing Li, and Jess Clouser have been instrumental in protecting my academic time to pursue my goals. Thank you for your leadership and recognizing the importance of the academic careers of your employees.

Chapter 1

Introduction

Background and Statement of the problem

Colorectal cancer (CRC), includes any cancer that arises in the colon or rectum, the part of the gastrointestinal system making up the large intestines. CRC is sometimes termed colon cancer, bowel cancer, or rectal cancer.[7] Several decades ago CRC had a low incidence rate, however, it is now the third most common type of cancer and the third most common cause of cancer death among men and women in the United States.[1-3] The American Cancer Society estimates that in 2020 there will be 147,950 new cases of CRC and 53,200 CRC related death in the U.S..[3]

Kentucky ranks number one in per capita cancer incidence and mortality rates.[5] The risk of developing any cancer increases with age, the same is true for CRC, older adults (50 years and older) have the most CRC burden than any age group.[3, 10] CRC incidence reported for 2012-2016 in Kentucky was the highest in the nation at 49.2 (per 100,000) and mortality for years 2013-2017, Kentucky ranked 5th in the nation with a rate of 16.4 (per 100,000).[4-6] Eastern Kentucky, part of the Appalachia region, makes up about 15% of the population of Kentucky, yet experiences a higher rate of mortality and morbidity than the rest of the state.[5] The area also has high prevalence rates of lung disease, heart disease, and diabetes.[5]

While healthcare professional do not know the cause of CRC, there are many known risk factors associated with CRC.[2] Risk factors for CRC include family history, being over the age of 50, African American race, history of polyps, radiation therapy, inherited and inflammatory diseases, and lifestyle factors like low physical activity, smoking, alcohol, obesity, and high-fat diets.[2, 7] Diabetes has

been found to be the most prevalent comorbidity among CRC patients.[8, 9] The risk of developing CRC in patients with diabetes is estimated to be more than 25% higher than those without diabetes.[9]

There are studies addressing 1) the prevalence of CRC screening and comorbidities within Appalachian Kentucky and 2) non-adherence to the standard of care as a contributing factor for Kentucky's high mortality rate.[11, 12] To date, there is no known study that has compared CRC outcomes and comorbidities across Kentucky.

Purpose of the Study

To address these gaps in the literature, we conducted secondary data analysis on CRC patients in Kentucky using data from the Kentucky Cancer Registry (KCR). KCR is a population-based central cancer registry that collects data on cancer, treatment, death, and individual demographic data.

The overarching purpose of the current study is to explore if comorbidities impact CRC progression, CRC outcomes, and the development of second primary malignancy among CRC patients age 18 and older in Kentucky diagnosed between January 1, 2003 and December 31, 2016.

There are three specific aims for this study:

1. Aim 1 was to characterize the patient factors of socio-demographic and comorbidity by stage of diagnosis.
2. Aim 2 was to examine if comorbidity status is associated with mortality and having second primary cancers.
3. Aim 3 was to perform a space-time cluster analysis of late-stage at diagnosis to investigate its relationship with comorbidities at the population level.

Research hypotheses

H_1: CRC patients in Kentucky without comorbidity are more likely to be diagnosed with late-stage cancer compared to patients with comorbidity.

H_0: There is no difference in the diagnosis of late-stage CRC patients in Kentucky with or without comorbidity.

H_2: CRC patients in Kentucky with comorbidity are more likely to be diagnosed with a second primary malignancy compared to patients without comorbidity.

H_0: There is no difference in the diagnosis of second primary malignancy in CRC patients with or without comorbidity.

H_3: CRC patients in Kentucky with comorbidity have higher all-cause mortality compared to patients without comorbidity.

H_0: There is no difference in all-cause mortality in CRC patients with or without comorbidity.

H_4: CRC patients in Kentucky with comorbidity have higher CRC mortality compared to patients without comorbidity.

H_0: There is no difference in CRC mortality in CRC patients with or without comorbidity.

H_5: High rate late-stage CRC diagnoses in Kentucky will be spatially and temporally correlated with low rates of CRC morbidity.

H_0: There is no spatial or temporal correlation between high rates of late-stage diagnosis and CRC comorbidities among CRC cases in Kentucky.

Significance of the study

This study will add to the extant literature by examining the relationship between comorbidity and cancer outcomes including survival and development of second primary malignancy across Kentucky to inform comprehensive prevention programs targeting populations identified at risk. The results of this study will also be useful for informing future CRC and comorbidity research.

Limitations and delimitations of the study

There are several potential methodological limitations that should be considered when interpreting data on second primaries when using a population-based cancer registry. One major limitation is the lack of a standard definition of multiple primaries that can be applied across all diagnosed cases, a person diagnosed in 2003 would have different classification rules from someone diagnosed in 2005, 2007, or 2018.[13-17] This makes it extremely difficult to be accurate and compare multiple primaries across years.[17] The data set includes second primary according to the rules during the time period of the diagnosis. The biggest limitation of these data is with regards to changes in the rules for diagnosis of second primaries during different time periods. Second primaries would not mean the same thing between those time periods and this would affect the interpretation of the results. For the study time period 2003-2016, KCR's multiple primary rules remained relatively similar. Other potential limitations include not all cases being captured, incomplete data due to clerical errors, and missing data due to unavailability.[18] The missing comorbidities data could be due to the reality that many reporting hospitals are not part of the Commission of Cancer (CoC), a group that requires comprehensive standardized data collection.[19, 20] Another limitation would be patients receiving surgery for something other than a malignancy, such as receiving resection of the colon that may or may not contain cancer, and may or may not be tested because it was removed for other reasons.

Overview of Project Processes

Descriptive statistics were used to compare the socio-demographic and comorbidity factors in CRC patients included and excluded from the study. Odds ratios and 95% confidence intervals were estimated from Logistic regression. Multivariable Cox proportional hazard regressions were used to examine if comorbidity status was associated with CRC survival and second primary malignancy. The statistical software SAS version 9.4 was used for all of the above analyses.[21] The spatial software, SaTScan and ArcGIS, were also used to perform spatial analysis and visualize the prevalence of comorbidity among CRC patients in Kentucky.[22, 23]

Definition of Terms in context of this study

- **Colorectal Cancer** (CRC)- an initial primary cancer that arises in the colon or rectum
- **Comorbidity** – a chronic health condition in the presence of primary CRC
- **Multi-morbidity** – the presence of two or more comorbidities in the presence of primary CRC
- **Second primary malignancy** – an additional primary cancer that arises after the first primary cancer has been diagnosed and treated
- **Kentucky Cancer Registry** (KCR) – a population-based registry in Kentucky that collects cancer related personal health, and treatment information

Chapter 2

Literature Review

The purpose of this literature review is to introduce and orient the topics of interest for this CRC study. The goal is to identify any gaps in knowledge surrounding CRC, comorbidities, and the use of spatial analysis to view disease distribution over a geographical area.

Cancer Surveillance

Cancer surveillance is the routine continuous systematic collection and analysis of data on new cancer incidences, morbidity, treatment, survival, and mortality.[24, 25] Cancer surveillance quantifies the incidence of cancer and its related factors (e.g., genetic or behavioral factors) in a defined population to provide a means by which the observations can be used in research to facilitate interventions and reduce the burden of cancer.[24, 25] Cancer surveillance, like that of other Public Health Surveillance programs have strict inclusion criteria that could include diagnosis, timing, and/or be laboratory confirmed to be considered a case, while a clinical diagnosis may not be as involved for a patient to receive the diagnosis and treatment.[26]

In the United States there is no nationally recognized single surveillance program.[25] While there are several smaller registries that are created from doctors' offices, hospital and healthcare system registries, to state population-based registries, there are two important national cancer surveillance programs, National Program of Cancer Registries (NCPR) and The Surveillance, Epidemiology and End Results Program (SEER).[25]

In Kentucky, the state cancer registry, Kentucky Cancer Registry (KCR), began as a voluntary reporting system until legislation mandated reporting starting in 1991.[27] KCR is funded by NPCR and SEER. Data collected by KCR is sent to the umbrella organization, North American Association of

Central Cancer Registries (NAACR) to be independently evaluated for completeness, accuracy, and timeliness.[20, 27] Kentucky Cancer registry is very thorough, and among the most accurate and complete population-based registries.[27]

Colorectal Cancer

CRC, includes any cancer that initiates in the colon or rectum, the portion of the gastrointestinal system that makes up the majority of the large intestines, other names for CRC are colon cancer, bowel cancer, or rectal cancer.[7] The anus is the final part of the large intestine but because of the cell types are different that make up the anus, any cancers originating in the anus is classified as anal cancer.[3] The colon is made up of four sections, the ascending colon, transverse colon, descending colon, and the sigmoid colon, it is about 5 feet long and is connected at the bottom to the rectum.[3] There are three functions of the large intestine including absorption of electrolytes and water from food being digested, production and absorption of vitamins, and the formation and elimination of fecal waste from the body.[3, 28]

CRC is ranked third for most commonly diagnosed cancer and it is also the third leading cause of cancer death in both men and women.[3] While there are no certain causes of CRC, there are many known risk factors.[2] Risk factors for CRC include advanced age, African American race, history of polyps, family history of colon cancer, sedentary lifestyle, high-fat diet, diabetes, obesity, smoking, radiation therapy, alcohol, and inherited and inflammatory diseases.[2, 7]

Patient Factors

Age

The median age at diagnosis for rectal cancer is 63 years old and median age of diagnosis for colon cancer is 68 in men and 72 in women.[29] The majority of CRCs are diagnosed in people over the age of 50, with only around 12% being diagnosed in people under the age of 50.[3] The incidence of

CRC in patients under the age of 50 has been increasing, however healthcare providers are not sure the reason behind the increase.[2, 10] One author found that patients diagnosed at younger than 50 years of age presented with advance stage and higher recurrence of CRC than older patients but the two groups had comparable survival.[10] Right-sided colon cancer seems to be more prevalent in older adults and women and this type of CRC usually presents at more advanced stages with lower survival rate.[30]

Sex

The lifetime risk of developing cancer is similar in both men and women, about 1 in 23 men and 1 in 25 women, a difference of 0.3% incidence.[7] Socioeconomic factors seem to disproportionately affect CRC incidence rates in men. One study from England found that the most deprived areas had a 13% higher incidence rate compared to the least deprived areas; there was no difference found in women.[31] The author also noted that men are less aware of cancer signs and symptoms compared to women.[31] Another study noted that genetic and environmental factors are believed to play a role in sex-associated differences in CRC, with high-fat diet being associated with the risk and development on CRC.[32] The biological responses to diet are different among men and women.[32] Studies have associated women with a higher proportion of right-sided colon cancer, which usually presents at a more advanced stage, which might account for women having a lower 5-year survival rate.[30, 32]

Second Primary

A second primary is a new primary cancer that arises in a person that has had a diagnosis and treatment for a non-related cancer in the past.[33] Second primaries comprise almost 19% of incident cancer cases.[34] Patients can have multiple primaries, the requested data included the number of primary cancers, including the first primary cancer of CRC. Patients that have had CRC often have

several health problems, including a higher risk of secondary cancers.[34, 35] People that have had colon or rectal cancers can develop any second primary cancer but are at an increased risk of developing colon, rectal, stomach, small intestine, anal, or lung cancer.[35, 36] While the risk of secondary primary cancers is complex, genetics, previous cancer treatments, and environmental exposures have been recognized as risks to developing a second primary.[34, 37]

Comorbidities

Elixhauser groups

A comorbidity is defined as a disease or condition that exists simultaneously with another index condition of interest.[38, 39] The presence of comorbidity in addition to an index condition such as cancer has become increasingly more common with much evidence supporting the majority of the comorbidity burden is concentrated in patients that are older, those in minority groups, and those patients living in poverty-stricken areas.[39] The presence of comorbidities varies by cancer site and is difficult to determine an accurate prevalence.[38, 39] Comorbidity is usually assessed as a contributor to a health outcome, like cancer survival, using methods such as individual disease indexes or counts and weighted grouped variables to help describe overall disease burden and health status of a patient.[38, 40]

In this study, comorbidities will be looked at both on an individual level and an index, using the Elixhauser comorbidity index. The Elixhauser comorbidity index includes 29 individual comorbid conditions that were initially selected and refined by examining the literature.[41-43] The final use of the Elixhauser index was modified to only include 26 groupings. Diabetes with and without chronic complications was combined into one group. Three of the groupings, metastatic cancer, lymphoma, and solid tumor without metastasis were not evaluated. Although the Kentucky Cancer Registry

captures these data, these cancers should not be considered comorbid conditions as part of the index since the disease of interest is also cancer.

Prior to the newest Elixhauser measure, there were 31 groupings, the updated version has since collapsed hypertension (complicated and uncomplicated) and removed cardiac arrhythmias, as research has shown cardiac arrhythmia was not a good indicator of readmission, and questions remain around its reliability as a comorbidity.[41, 44] The older Elixhauser measure that used ICD-9-CM diagnosis codes was converted to the new Elixhauser version, which utilized ICD-10-CM codes, and combined according to Agency for Healthcare Research and Quality's (AHRQ) guidelines.[41, 45] Traditionally the Elixhauser index has been used in research as a count index, more eligible diagnoses would lead to a higher count and risk to the patient.[46-48] This index, like the Charlson Comorbidity index, has been used to predict in-hospital mortality, high-risk patients, and scenarios that may need a higher intervention of care both while in the hospital and when transitioning out of the hospital to prevent readmissions.[44, 45, 49] One study found that using the individual Elixhauser comorbidities in a regression gave slightly better results compared to the Elixhauser index score and it is possible "that a comorbidity measure with more variables can lose more information than one with a smaller number of variables in finite sample sizes".[48] While the index score is important to use, the objective of this study is not focused on hospital utilization and readmission; the individual and grouped comorbid conditions are more important to use in this study.

Colorectal cancer and comorbidities

A number of studies have shown that cancer patients with comorbidity have lower survival compared to cancer patients without comorbidity.[50] Although the pattern of comorbidities and their risk factors among CRC patients is not well documented globally, there is however, consistent evidence illustrating the effect of comorbidities on CRC outcomes.[8] Morbidities are often associated

with the elderly, but have recently been occurring in younger patients living in socioeconomically deprived areas.[51] A study of adults with hypertension found that in a year of visits with their primary care provider, only one third were related to hypertension while the next most common reason for their care visit was for diabetes.[51] Studies have found that morbidities, both physical and mental health conditions, do not exist in isolation and are influenced by an individuals' society and family.[8, 51] Studies have found that diabetes is the most prevalent comorbidity among CRC patients.[8, 9] The risk of CRC in patients with diabetes is estimated to be more than 25% higher than those without diabetes.[9]

Another study found that CRC patients can be grouped into four classes based on defined clusters of comorbid conditions.[40] Class one represented the largest part of the sample and included patients with no Charlson-defined comorbidities or only one morbidity.[40] Classes two and three were similar in size and age at diagnosis, however class two patients comorbid conditions were primarily characterized as cardiovascular or cardiorespiratory diseases, while class three comorbid conditions were primarily diabetes with complications such as kidney disease.[40] Class four consisted of less than 8% of the study population and were comprised of the patients with the presence of four or more comorbidities.[40] Class one patients had the highest survival probability followed by class three, class two, and then class four, with the lowest survival probability -- 43% lower than class one.[40] The majority of class four patients were older with a higher burden of comorbidity.[40] Despite the increasing importance of comorbidity among cancer patients, many challenges and questions remain.[50] Cancer patients with comorbidities have compromised treatment plans, effectiveness, and compliance, and we do not know the duration and severity of the influence of comorbidity on cancer prognosis or how comorbidity is most accurately measured in cancer patients.[50, 52]

Geospatial topics

Appalachia

The Appalachian region includes all of West Virginia and parts of 12 other states, including 54 counties in the southeastern and eastern area of Kentucky.[53, 54] Within the Appalachian region, and more specifically rural Appalachia, health disparities have been well documented.[53] Appalachia as a whole experiences higher rates of mortality and chronic diseases such as diabetes, chronic obstructive pulmonary disease, heart disease, stroke, and cancers such as lung, breast, and CRCs than non-Appalachian areas.[53-55] Health disparities and disease in Appalachia are exacerbated by socioeconomic, behavioral, and geographical factors such as environmental exposures, poverty, low literacy rates, lack of health insurance coverage, long distances from home to clinics and healthcare providers, high rates of obesity and smoking, low physical activity, and many other multifactorial issues.[53, 55] Appalachia has high rates of CRC incidence and mortality and CRC is one of the leading causes of cancer deaths in Appalachia in both men and women.[56, 57]

Interaction of patient factors and geospatial data

Geographic information systems (GIS) are used in epidemiological research to identify the "where" of disease.[58] Spatial data in public health studies allows for researchers to visualize disease and patient attributes across geographic areas which can help to identify and characterize health trends over time.[59, 60] Spatial analysis of patient data can help to determine clustering or patterns in geographic areas that will help to understand patient populations at higher risk, determine any socioeconomic factors, and highlight areas that would need intervention in addressing health disparities.[59, 60] Cluster analysis is useful in producing estimates where limited data is available and providing statistical evidence of diseases.[58] Creating maps based on disease information more easily

reveals geographic-related information about disease distribution than typical research tables showing data.[58]

Utilizing SaTScan in this study to perform a cluster analysis was helpful in detecting areas with high or low rates of statistical significance during time period, 2003-2016.[22] SaTScan was required to perform the analysis because it is not available in standard GIS software packages.[61] For mapping purposes, the resulting cluster analysis from SaTScan was exported and layered with a map of Kentucky in ArcGIS 10.7.1.[62] ArcGIS was also used to map the proportion of comorbidities and late-stage cancers within each county.[23] These maps are necessary for us to visualize the geographic distribution of these factors and evaluate them relative to one another.

Chapter 3

Paper 1: The Effect of Comorbidities on Colorectal Cancer Stage at Diagnosis, Mortality, and Second Primary Cancers among Colorectal Cancer Cases in Kentucky (2003-2016)

Background

Colorectal cancer (CRC) is cancer that begins in the colon or rectum, the portion making up the large intestines of the gastrointestinal system, other names are colon cancer, bowel cancer, or rectal cancer.[7] CRC is the third most common type of cancer and the third most common cause of cancer death among men and women in the United States.[1-3] By 2020, The American Cancer Society estimates that there will be 147,950 new cases of CRC and 53,200 CRC related death in the U.S..[3] Kentucky ranks number one in overall cancer per capita incidence and mortality rates.[5] Kentucky had the highest CRC incidence in nation with 49.2 (per 100,000) for years 2012-2016 and ranked 5th in the nation in CRC mortality at a rate of 16.4 (per 100,000) for years 2013-2017.[3-6]

Eastern Kentucky, consisting of much of Kentucky's Appalachia region, is estimated to include slightly under 15% of the population of the state, yet the poverty-stricken area experiences a higher rate of mortality and morbidity than the rest of the state and the nation.[5] The area is also known for its high prevalence rates of chronic illnesses such as lung disease, heart disease, and diabetes.[5]

A comorbidity is defined as a chronic illness that exists concurrently with an index condition of interest.[38, 39] In this study, the index condition is primary CRC. The presence of comorbidity in addition to an index condition, like CRC has become increasingly more common with considerable evidence supporting a larger part of the burden is concentrated in older patients, minority groups, and patients living in poverty-stricken or deprivation areas.[39]

The presence of comorbidity can vary by cancer site making it difficult to determine an accurate prevalence.[38, 39] In this study, comorbidities will be examined on an individual level as well as aggregated into groups, using the Elixhauser comorbidity index. The Elixhauser comorbidity index includes 29 individual comorbid conditions, that were grouped according to similar body systems (i.e. grouping acute heart failure with chronic heart failure in the congestive heart failure group) reference table 1-1.[41, 42] Comorbidity is usually assessed as a contributor to health outcomes, like cancer survival, using methods such as individual disease indexes or scores and weighted grouped variables to help describe a patient's overall disease burden and health status.[38, 40] As such, the Elixhauser index was chosen for use in the current study.

The exact cause of CRC is not known, however there are many known risk factors associated with CRC.[2] The risk associated with developing any cancer increases with age, the same is true for CRC, older adults (50 years and older) have the most CRC burden than any age group.[3, 10] Other known risk factors for CRC include family history of CRC, being of African American race, history of polyps, history of radiation therapy, other inherited and inflammatory diseases, and lifestyle factors like low physical activity, smoking, alcohol, obesity, and high-fat diets.[2, 7] Diabetes has been found to be the most prevalent comorbidity among CRC patients.[8, 9] The risk of developing CRC in patients with diabetes is estimated to be more than 25% higher than those without diabetes.[9]

The purpose of this study is to explore if comorbidities impact CRC progression, CRC outcomes, and the diagnosis with a second primary malignancy among CRC patients aged 18 and older in Kentucky diagnosed between January 1, 2003 and December 31, 2016. There are two specific aims for this study. Aim 1 was to characterize patient factors of socio-demographic and comorbidity by stage of diagnosis. Aim 2 was to examine whether comorbidity status is associated with stage at

diagnosis, mortality, and the development of second primary cancers. To our knowledge, this will be the first study that has compared CRC outcomes and comorbidities across Kentucky.

Methods

Study Design and Data Source

This is a retrospective cohort study of CRC cases in Kentucky. We started out with 28,229 incident cases of first primary CRC diagnosed between January 1, 2003 and December 31, 2016 and excluded a combined total of 7,658 cases due to missing information. The excluded cases included 6,054 because of missing morbidity information and 2,730 (1,126 of these were also missing morbidity information and included in the above number) because of missing stage. The final study population included 20,571 CRC cases. Table 1-2 compares included and excluded cases. All cases were identified from the Kentucky Cancer Registry (KCR). KCR is funded in part by Surveillance, Epidemiology, and End Results Program (SEER) and National Program of Cancer Registries (NCPR), and North American Association of Central Cancer Registries (NAACR).[20, 27] KCR data is sent to the umbrella organization, North American Association of Central Cancer Registries (NAACR) to be independently evaluated for completeness, accuracy, and timeliness. Approval for this study was granted by the University of Kentucky Internal Review Board.

Variables

Sex, age at diagnosis, race, ethnicity, marital status at diagnosis, number of primaries, survival for primary and a subsequent second primary, Appalachian status, vital status, primary payer, best stage group, treatment composite, comorbidity (up to 10 independent variables of ICD-9-CM diagnosis codes), and secondary diagnosis (up to 10 independent variables of ICD-10-CM diagnosis codes) were provided by KCR. Age at diagnosis was categorized into five age groups, 18 - 34, 35 - 44,

45 – 54, 55 – 64, and 65+ years. Number of primaries was coded as 0 for the initial CRC primary of interest and 1 for any patient that had been diagnosed for any subsequent primaries not related to their initial primary of CRC. Race was categorized as white, black, and other. Ethnicity was categorized as non-Hispanic or Hispanic. Marital status at diagnosis was categorized at married, single, or other. Primary payer was categorized as Medicaid, Medicare, military/other, private pay, and not insured. The variable treatment had 15 combinations of treatment, including no treatment, and variations of surgery, radiation therapy, chemotherapy, and other therapies. The other therapies were not expressly specified, but typically include immunotherapy. Treatment coding in this study was based on the available dataset and a CRC study by Rane *et al.*[63] The final coding included six classes: no treatment, surgery at primary site only, chemotherapy only, radiation only, chemotherapy and radiation, and surgery at primary site and chemotherapy/or radiation.

Comorbidity was measured using the diagnosis codes from the variables comorbidity and secondary diagnosis and entering into the Elixhauser Comorbidity Software, Version 3.7 for ICD-9-CM and the Elixhauser Comorbidity Software for ICD-10-CM from Healthcare Cost and Utilization Project by the Agency for Healthcare Research and Quality.[41, 45] The software classifies certain diagnoses codes as part of the Elixhauser Comorbidity index, outputting individual variables for the 31 (Version 3.7)/29 Elixhauser groups. The final variables were combined to match the most up-to-date Elixhauser index, removing arrhythmias and combining hypertension.[41] Elixhauser groups related to cancer, solid tumors without metastases, metastatic cancer, and lymphoma were removed. Cancer is the outcome of interest, and therefore it cannot be a comorbidity. Diabetes with and without chronic complications were combined after observing some patients had both diagnosis codes.

The Adult Comorbidity Evaluation-27 (ACE-27) index morbidity groupings were considered, however, available data only includes diagnosis codes, indicating the presence of disease but not

disease severity.[64] ACE-27 grades the extent of organ decompensation on three grades of severity, mild, moderate, and severe, data from KCR did not allow for measurement of such decompensation.[65] Nonetheless, studies have also shown that the Elixhauser measure performs better than other comorbidity indexes.[46, 48, 49, 66] Table 1-1 shows the morbidity mapping from ACE-27 and Elixhauser Comorbidity index to the final inclusion of comorbidities (individual and grouped).

The KCR variables with diagnoses codes include coding for patients with no known morbidities (comorbidity diagnosis code of 0000 or a secondary diagnosis entry of 0). Patients with these entries in any diagnoses code variables were treated as having no morbidity. Patients with diagnosis code(s) in the comorbidity/secondary diagnosis variables that are not part of the Elixhauser groups were also treated as having no morbidity. Cases with unknown morbidity status and unknown stage at diagnosis were excluded from the study. Figure 1-1 shows flow chart for case inclusion within the study.

Statistical Analysis

All analyses were carried out with SAS 9.4 statistical software.[21] Included and excluded patients were compared on demographic and disease characteristics using column percentages to observe comparisons between groups (see Table 1-2). Included patient demographic characteristics stratified by cancer stage, early and late using row percentages as comparison can be found in Table 1-2b.

To explore the relationship between morbidities, stage, and survival, a series of bivariate models were fitted. Two logistic regression models were fitted (tables 1-3 and 1-4), one with stage at diagnosis (late versus early) and the other with vital status (died versus alive). A cancer specific

survival model was estimated using Cox proportional hazard regression. Kaplan-Meier survival curves for primary CRC survival were also analyzed using SAS 9.4 software and can be found in figure 1-2.[21]

Next, a series of multivariable fully adjusted statistical models for estimating the risk of mortality or second primary were specified. These models were Cox proportional hazard models that were fitted to explore the relationship between patient factors and morbidities on survival and second primary malignancy. The first set of three Cox proportional hazard models looked as all-cause mortality and can be found in table 1-6. The second set of three Cox proportional hazard models looked as CRC-cause mortality and can be found in table 1-7. The third set of three Cox proportional hazard models looked at second primary cancer and can be found in table 1-8. The three sets of models included sociodemographic data, cancer stage, and followed different categories of Elixhauser comorbidities; Model 1 used the individual Elixhauser morbidities, Model 2 used Elixhauser grouped comorbidities based on Table 1-1, and then Model 3 used a total count of the number of morbidities that a patient would have (no morbidities/0, one morbidity/1, or comorbidities/2+).

Table 1-1. Morbidity Mapping ACE-27 Index, Elixhauser, and Final Inclusion Study Comorbidity and Groupings

Ace-27 Index	Elixhauser ICD 10*	Final Inclusion
Cardiovascular System	Congestive Heart Failure	**Cardiovascular System**
Myocardial Infarct	Valvular disease	Congestive Heart Failure
Angina / Coronary Artery Disease	Pulmonary circulation disorders	Hypertension
Congestive Heart Failure (CHF)	Peripheral vascular disease	Peripheral Vascular Disorder
Arrhythmias	Hypertension (Complicated &	Valvular Heart Disease
Hypertension	Uncomplicated)	**Respiratory System**
Venous Disease	Paralysis	Pulmonary Circulation Disorders
Peripheral Arterial Disease	Other neurological disorders	Chronic Pulmonary
Respiratory System	Chronic pulmonary disease	**Gastrointestinal System**
Restrictive Lung Disease or COPD	Diabetes Uncomplicated	Liver Diseases
Other markers not diagnoses	Diabetes Complicated	Peptic Ulcer Disease
Gastrointestinal System	Hypothyroidism	**Renal System**
Hepatic	Renal failure	Renal Failure
Stomach / Intestine	Liver disease	**Endocrine System**
Pancreas	Chronic peptic ulcer disease	Diabetes (Complicated & Uncomplicated)
Renal System	HIV and AIDS	Hypothyroidism
End-stage renal disease	Lymphoma	**Neurological System**
Endocrine System	Metastatic cancer	Paralysis
Diabetes Mellitus	Solid tumor without metastasis	Other Neurological Disorders
Neurological System	Rheumatoid arthritis/ collagen vascular	**Psychiatric**
Stroke	diseases	Psychoses
Dementia	Coagulation deficiency	Depression
Paralysis	Obesity	**Rheumatologic**
Neuromuscular	Weight loss	Rheumatoid Arthritis/Collagen
Psychiatric	Fluid and electrolyte disorders	**Immunological System**
Recent suicidal attempt	Blood loss anemia	AIDS/HIV
Schizophrenia	Deficiency anemias	**Substance Abuse**
Depression or bipolar disorder	Alcohol abuse	Alcohol Abuse
Rheumatologic	Drug abuse	Drug Abuse
Rheumatoid Arthritis	Psychoses	**Body weight**
Systemic Lupus	Depression	Obesity
Mixed Connective Tissue Disorder		**Blood System**
Polymyositis		Coagulopathy
Rheumatic Polymyositis		Blood Loss Anemia
Immunological System		Deficiency Anemia
AIDS		**Remain but not grouped**
Malignancy		Weightloss
Solid Tumor including melanoma		Fluid & Electrolyte Disorders
Leukemia and Myeloma		
Lymphoma		
Substance Abuse		
Alcohol		
Illicit Drugs		
Body Weight		
Obesity		

*Elixhauser ICD-9 (31 Variables) conversion to ICD-10 (29 Variables) version changes: Removal of Cardiac Arrhythmia and Combining of Hypertension with and without complications. Some data were collected prior to the 2015 ICD-10 activation, combination and conversion of ICD-9 and ICD-10 Elixhauser adjustment occurred.

Figure 1-1. Flow chart for participant selection and inclusion/exclusion criteria for study.

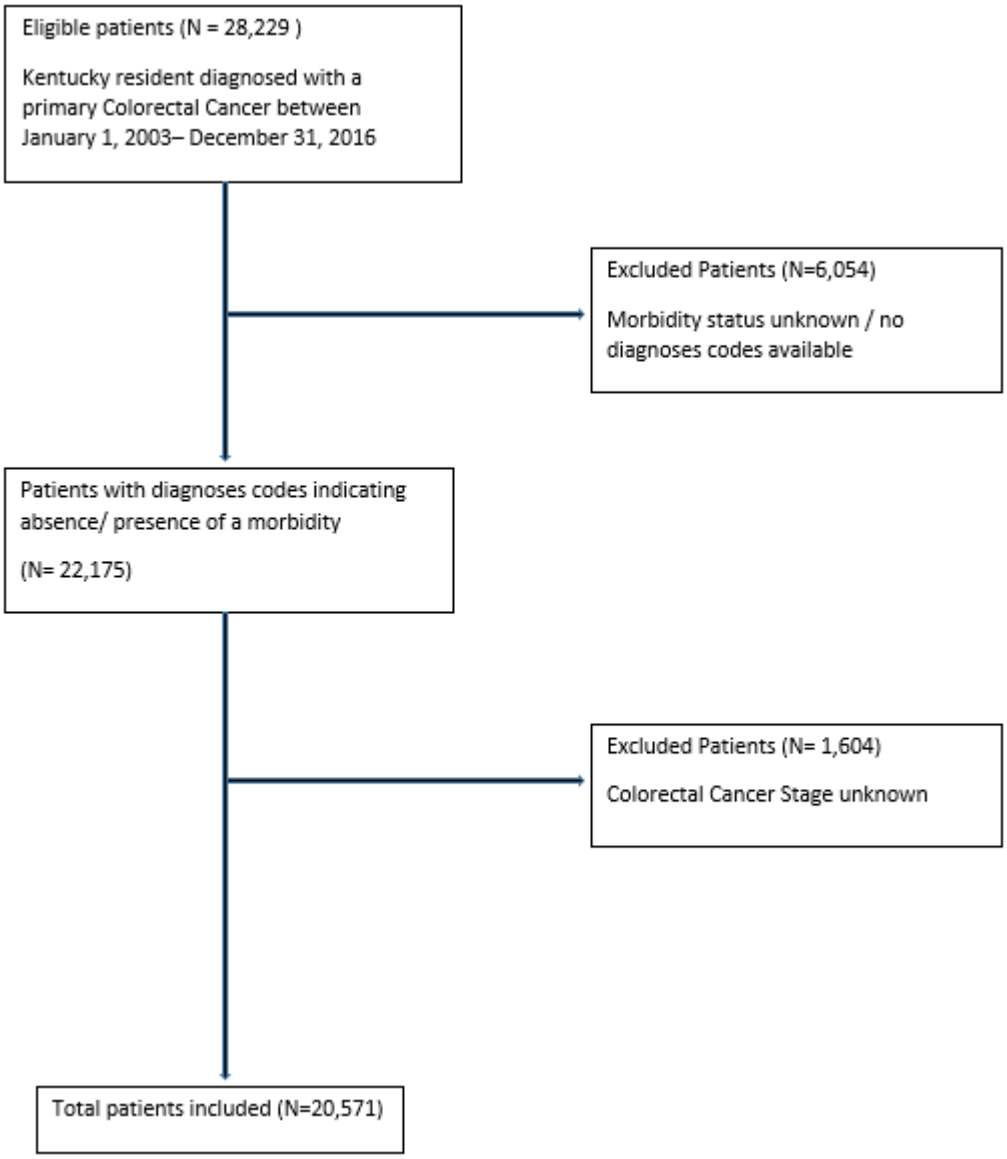


Table 1-2. Inclusion and Exclusion Demographics (Percentage based on Columns)

Variables	Inclusion Group N(%)			Exclusion Group N(%)			χ^2
	Early Stage (N=11,387)	**Late Stage** (N=9,184)	**Total Included** (N=20,571)	**Unknown Comorbidity** (N=6,054)	**Unknown Stage** (N=2,730)	**Total Excluded** (N=7,658)	Probability Distribution Between Groups
Age							
18 - 34 years	120 (1.05%)	168 (1.83%)	288 (1.40%)	90 (1.49%)	112 (4.10%)	174 (2.27%)	
35 - 44 years	447 (3.93%)	498 (5.42%)	945 (4.59%)	276 (4.56%)	177 (6.48%)	398 (5.20%)	p <0.001
45 - 54 years	1,635 (14.36%)	1,522(16.57%)	3,157 (15.35%)	925 (15.28%)	468 (17.14%)	1,214(15.85%)	
55 - 64 years	2,720 (23.89%)	2,343(25.51%)	5,063 (24.61%)	1,379(22.78%)	551 (20.18%)	1,712(22.36%)	
65+ years	6,465 (56.78%)	4,653(50.66%)	11,118 (54.05%)	3,384 (55.90%)	1,422(52.09%)	4,160(54.32%)	
Marital Status							
Missing	238 (2.09%)	193 (2.10%)	431 (2.10%)	509 (8.41%)	223 (8.17%)	560 (7.31%)	
Married	6,618(58.12%)	5,130(55.86%)	11,748 (57.11%)	3,010(49.72%)	1,308(47.91%)	3,880(50.67%)	p <0.001
Single	1,159(10.18%)	1,053(11.47%)	2,212(10.75%)	553 (9.13%)	305 (11.17%)	750 (9.79%)	
Other	3,372(29.61%)	2,808(30.57%)	6,180(30.04%)	1,982(32.74%)	894 (32.75%)	2,468(32.23%)	
Sex							
Female	5,540 (48.65%)	4,407(47.99%)	9,947 (48.354%)	2,920(48.23%)	1,369(50.15%)	3,749(48.96%)	
Male	5,846 (51.34%)	4,777(52.01%)	10,623(51.641%)	3,134(51.77%)	1,361(49.85%)	3,909(51.04%)	p = 0.371
Missing	1 (0.009%)	0	1 (0.005%)	0	0	0	
Race							
White	10,518(92.37%)	8,464(92.16%)	18,982 (92.28%)	5,604(92.57%)	2,428(88.94%)	7,045(92.0%)	
Black	792 (6.96%)	653 (7.11%)	1,445 (7.02%)	271 (4.48%)	215 (7.88%)	418 (5.46%)	p <0.001
Other	77 (0.68%)	67 (0.73%)	144 (0.70%)	179 (2.96%)	87 (3.19%)	195 (2.55%)	
Ethnicity							
Non-Hispanic	11,236(98.68%)	9,092(98.99%)	20,328 (98.82%)	5,966(98.55%)	2,664(97.58%)	7,545(98.52%)	p = 0.532
Hispanic	23 (0.20%)	30 (0.33%)	53 (0.26%)	19 (0.31%)	10 (0.37%)	23 (0.30%)	
Missing	128 (1.12%)	62 (0.68%)	190 (0.92%)	69 (1.14%)	56 (2.05%)	195 (2.55%)	
Vital Status							
Alive	6,999 (61.46%)	3,010(32.77%)	10,009 (48.66%)	2,657(43.89%)	1,181(43.26%)	3,450(45.05%)	p <0.001
Dead	4,388 (38.54%)	6,174(67.23%)	10,562 (51.34%)	3,397(56.11%)	1,549(56.74%)	4,208(54.95%)	
Stage							
Early	11,387 (100%)	0	11,387(55.35%)	2,934(48.46%)	0	2,934(38.31%)	
Late	0	9,184 (100%)	9,184 (44.65%)	1,994(32.94%)	0	1,994(26.04%)	p <0.001
Unknown	0	0	0	1,126(18.60%)	2,730 (100%)	2,730(35.65%)	
Appalachian	3,045(26.74%)	2,593(28.23%)	5638(27.41%)	2,708(44.73%)	855 (31.32%)	3162(41.29%)	p <0.001
Diabetes	2169 (19.05%)	1653 (18.00%)	3822 (18.58%)	-	198 (7.25%)	198(2.59%)	p <0.001
Renal Failure	412 (3.62%)	290 (3.16%)	702 (3.41%)	-	43 (1.58%)	43 (0.56%)	p <0.001
Liver Disease	287 (2.52%)	328 (3.57%)	615 (2.99%)	-	32 (1.17%)	32 (0.42%)	p <0.001
CHF	834 (7.32%)	601 (6.54%)	1435 (6.98%)	-	133 (4.87%)	133 (1.74%)	p <0.001
Hypertension	5384 (47.28%)	4053 (44.13%)	9437 (45.88%)	-	565 (20.70%)	565 (7.38%)	p <0.001
Total Comorbidity Groups							
0	3,107 (27.29%)	2,382(25.94%)	5,489 (26.68%)	0	650 (23.81%)	650 (8.49%)	p <0.001
1	3,097 (27.20%)	2,443(26.60%)	5,540 (26.93%)	0	405 (14.84%)	405 (5.29%)	
2+	5,183 (45.51%)	4,359(47.45%)	9,542 (46.39%)	0	549 (20.11%)	549 (7.17%)	
Unknown	0	0	0	6,054 (100%)	1,126(41.25%)	6,054(79.05%)	
Primary Payer							
Missing	58 (0.51%)	55 (0.60%)	113 (0.55%)	601 (9.93%)	420 (15.38%)	617 (8.06%)	
Medicaid	642 (5.64%)	699 (7.61%)	1,341 (6.52%)	429 (7.09%)	159 (5.82%)	538 (7.03%)	
Medicare	6,538 (57.42%)	4,823(52.52%)	11,361 (55.23%)	3,168(52.33%)	1,205(44.14%)	3,975(51.91%)	p = 0.0004
Military/Other	96 (0.84%)	112 (1.22%)	208 (1.01%)	42 (0.69%)	43 (1.58%)	75 (0.98%)	
			732 (3.56%)				

Not insured	310 (2.72%)	422 (4.59%)	6,816 (33.13%)	237 (3.91%)	60 (2.20%)	274 (3.58%)	
Private Payer	3,743 (32.87%)	3,073(33.46%)		1,577(26.05%)	843 (30.88%)	2,184(28.52%)	
Number of Primaries							
1	9,475 (83.21%)	8,217(89.47%)	17,692 (86.00%)	5,248(86.69%)	2,419(88.61%)	6,658(86.94%)	p = 0.042
2 or more	1,912 (19.76%)	967 (10.53%)	2,879 (14.00%)	806 (13.31%)	311 (11.39%)	1,000(13.06%)	
Survival Interval							
1 year	1,452 (12.75%)	2,874(31.29%)	4,326 (21.03%)	2,133(35.23%)	1,096(40.15%)	2,615(34.15%)	
2 years	1,197 (10.51%)	1,634(17.79%)	2,831 (13.76%)	878 (14.5%)	345 (12.64%)	1,086(14.18%)	
3 years	1,171 (10.28%)	1,097(11.94%)	2,268 (11.02%)	545 (9.00%)	222 (8.13%)	696 (9.09%)	p <0.001
4 years	1,170 (10.27%)	802 (8.73%)	1,972 (9.59%)	407 (6.72%)	187 (6.85%)	531 (6.93%)	
5 years	1,033 (9.07%)	606 (6.60%)	1,639 (7.97%)	340 (5.62%)	121 (4.43%)	428 (5.59%)	
> 5 years	5364 (47.11%)	2171 (23.64%)	7535 (36.63%)	1751(28.92%)	759 (27.80%)	2302(30.06%)	
Year of diagnosis							
2003 – 2009	5,701 (50.07%)	4,320(47.04%)	10,021 (48.71%)	3,409(56.31%)	1,484(54.36%)	4,263(55.67%)	p <0.001
2010 - 2016	5,686 (49.93%)	4,864(52.96%)	10,550 (51.29%)	2,645(43.69%)	1,246(45.64%)	3,395(44.33)	
Survival Interval 2nd Primary							
1 year	850 (7.49%)	587 (6.39%)	1,437 (6.99%)	375 (6.19%)	154 (5.64%)	473 (6.18%)	
2 years	234 (2.05%)	85 (0.93%)	319 (1.55%)	92 (1.52%)	30 (1.10%)	113 (1.48%)	p = 0.287
3 years	179 (1.57%)	71 (0.77%)	250 (1.21%)	67 (1.11%)	23 (0.84%)	82 (1.07%)	
4 years	136 (1.19%)	57 (0.62%)	193 (0.94%)	57 (0.94%)	24 (0.88%)	71 (0.93%)	
5 years	105 (0.92%)	37 (0.40%)	142 (0.69%)	46 (0.76%)	17 (0.62%)	53 (0.69%)	
> 5 years	408 (3.58%)	130 (1.42%)	538 (2.62%)	169 (2.79%)	63 (2.31%)	208 (2.72%)	
No2nd primary	9,475 (83.21%)	8,217(89.47%)	17,692 (86.00%)	5,248(86.69%)	2,419(88.61%)	6,658(86.94%)	

Table 1-3. Patient Demographics (Percentage based on rows)

Demographics	Early Stage (0-II) N=11387	Percent	Late Stage (III – IV) N=9184	Percent
Age				
18 - 34 years	120	41.67%	168	58.33%
35 - 44 years	447	47.30%	498	52.70%
45 - 54 years	1,635	51.79%	1,522	48.21%
55 - 64 years	2,720	53.72%	2,343	46.28%
65+ years	6,465	58.15%	4,653	41.85%
Marital Status				
Missing	238	55.22%	193	44.78%
Married	6,618	56.33%	5,130	43.67%
Single	1,159	52.40%	1,053	47.60%
Other	3,372	54.56%	2,808	45.44%
Gender				
Female	5,540	55.70%	4,407	44.30%
Male	5,846	55.03%	4,777	44.97%
Missing	1	100.00%		0.00%
Race				
White	10,518	55.41%	8,464	44.59%
Black	792	54.81%	653	45.19%
Other	77	53.47%	67	46.53%
Ethnicity				
Non-Hispanic	11,236	55.27%	9,092	44.73%
Hispanic	23	43.40%	30	56.60%
Missing	128	67.37%	62	32.63%
Vital Status				
Alive	6,999	69.93%	3,010	30.07%
Dead	4,388	41.55%	6,174	58.45%
Appalachian	3,045	54.01%	2,593	45.99%
Diabetes	2,169	56.75%	1,653	43.25%
Renal Failure	412	58.69%	290	41.31%
Liver Disease	287	46.67%	328	53.33%
CHF	834	58.12%	601	41.88%
Hypertension	5,384	57.05%	4,053	42.95%
Total Comorbidity Groups				
0	3,107	56.60%	2,382	43.40%
1	3,097	55.90%	2,443	44.10%
2+	5,183	54.32%	4,359	45.68%
Primary Payer				
Missing	58	51.33%	55	48.67%
Medicaid	642	47.87%	699	52.13%
Medicare	6,538	57.55%	4,823	42.45%
Military / Other	96	46.15%	112	53.85%
Not insured	310	42.35%	422	57.65%
Private Payer	3,743	54.91%	3,073	45.09%
Number of Primaries				
1	9,475	53.56%	8,217	46.44%
2 or more	1,912	66.41%	967	33.59%

Survival Interval				
1 year	1,452	33.56%	2,874	66.44%
2 years	1,197	42.28%	1,634	57.72%
3 years	1,171	51.63%	1,097	48.37%
4 years	1,170	59.33%	802	40.67%
5 years	1,033	63.03%	606	36.97%
More than 5 years	5,364	71.19%	2,171	28.81%
Year of diagnosis				
2003 – 2009	5,701	56.89%	4,320	43.11%
2010 - 2016	5,686	53.90%	4,864	46.10%
Survival Interval 2nd Primary				
1 year	850	59.15%	587	40.85%
2 years	234	73.35%	85	26.65%
3 years	179	71.60%	71	28.40%
4 years	136	70.47%	57	29.53%
5 years	105	73.94%	37	26.06%
More than 5 years	408	75.84%	130	24.16%
No second primary	9,475	53.56%	8,217	46.44%

Results

The demographic, clinical, and survival characteristics stratified by stage for included and excluded subjects is presented in Table 1-2. The main differences between the included and excluded cases were that the excluded cases had a higher percentage of patients from the Appalachian region (41.29% compared to 27.41%) and excluded cases had a higher percentage of cases in the 0-12 month survival interval (34.15% compared 23.01%). The remaining variables were similar, except for the morbidity variables, having 79.05% of the excluded cases missing morbidity status, comparison would not be recommended. For those only missing stage, the individual morbidity groups were at least 40% lowers in all categories, CHF, diabetes, hypertension, renal failure, and liver disease compared to the included cases.

The demographic, clinical, and survival characteristics stratified by stage for included subjects only is summarized in Table 1-3. Early-stage CRC makes up 55.4% of the total and the remaining 44.6% were late-stage cases. Comparing the demographic distribution of CRC range differences within 10% of expected distribution will not be noted. Late-stage CRC patients had a higher

percentage of younger patients (age groups 18-34 and 35-44 years), Hispanic patients, death, liver disease, higher percentage of patients surviving less than 24 months, and higher percentage primary payer sources of Medicaid, Military/other, and uninsured. Late-stage patients tended to have a lower percentage of second primary malignancies, primary survival beyond 48 months, and secondary survival beyond 12 months.

Bivariate models using the Elixhauser index predicting late-stage cancer by individual and grouped comorbidities are in Table 1-4. Individual comorbidities found to be statistically significant were CHF, hypertension, chronic pulmonary disease, liver disease, hypothyroidism, coagulopathy, blood loss anemia, deficiency anemia, depression, weight loss, and electrolyte disorders. Some comorbidities increased the odds of late-stage disease, e.g. liver disease (OR=1.432, 96% CI = 1.220 – 1.682), while others decreased the odds, e.g., hypertension (OR= 0.881, 95% CI = 0.883 – 0.931). Grouped comorbidities found to be statistically significant were cardiovascular system, respiratory system, gastrointestinal system, endocrine system, blood system, and having two or more individual comorbidities. Bivariate models predicting death by individual and grouped comorbidities are in Table 1-5. All but five individual comorbidities (excluding hypertension, peptic ulcer disease, hypothyroidism, drug abuse, and depression) and all but one comorbidity group (psychiatric) were found to be statistically significant below the alpha level of 0.05 and the corresponding 95% confidence limits did not include one.

There were three sets of three Cox proportional hazard models fitted, for a total of nine models. These models were progressive comparisons of comorbidities from an individual level, to a grouped body system level, and then an aggregate count of comorbidities. Some variables from the individual level were not able to be grouped with other individual comorbidities and were therefore transferred into the grouped model as their own group.

Cox proportional hazard models of all-cause mortality can be found in Table 1-6 and includes 20,270 CRC patients from Kentucky diagnosed 2003-2016; there are three models viewing comorbidity through an individual, system, or aggregate count lens.

Model 1 includes Individual comorbidities and patient factors. Due to the large number of significant findings, we will group the statistically significant results in order from largest to smallest effect. The individual comorbidities and patient factors that increased the hazard of all-cause mortality by greater than 40% that were found to be significant were late-stage (HR= 3.198, 95%CI 3.046-3.357), aids (HR= 2.901, 95% CI 1.926-4.373), being 65 years or older (HR= 1.746, 95%CI 1.421-2.143), primary payer sources of uninsured (HR= 1.711, 95%CI 1.538-1.903) or Medicaid (HR= 1.617, 95% CI 1.481-1.767), renal failure (HR= 1.570, 95%CI 1.432-1.720), CHF (HR=1.566, 95%CI 1.467-1.672), weight loss (HR= 1.540, 95%CI 1.427-1.662), other neurological disorders excluding paralysis (HR= 1.487, 95%CI 1.358-1.627), and coagulopathy (HR= 1.426, 95%CI 1.213-1.676). The individual comorbidities and patient factors that increased the hazard of all-cause mortality by greater than 20% but less than 40% that were found to be significant were pulmonary circulation disorders (HR= 1.394, 95%CI 1.204-1.615), electrolyte disorders (HR= 1.353, 95%CI 1.280-1.430), paralysis (HR= 1.347, 95%CI 1.087-1.671), primary payer source of Medicare (HR= 1.338, 95%CI 1.249-1.434) and Military/other (HR= 1.317, 95%CI 1.076-1.612), chronic obstructive pulmonary disorder (HR= 1.290, 95%CI 1.224-1.359), liver disease (HR= 1.226, 95%CI 1.105-1.360), rheumatoid arthritis (HR= 1.223, 95%CI 1.023-1.462), and alcohol abuse (HR= 1.202, 95%CI 1.031-1.401). The individual comorbidities and patient factors that increased the hazard of all-cause mortality by up to 20% that were found to be significant were peripheral vascular disease (HR= 1.157, 95%CI 1.044-1.281), deficiency anemia (HR= 1.146, 95%CI 1.086-1.210), being African American race (HR= 1.131, 95%CI 1.050-1.219), Appalachian (HR= 1.112, 95%CI 1.064-1.163), being male (HR= 1.074, 95%CI 1.032-1.118), and having

diabetes (HR= 1.068, 95%CI 1.015-1.123). The individual comorbidities and patient factors that were protective and decreased the hazard of all-cause mortality by up to 20% were hypertension (HR= 0.929, 95%CI 0.892-0.968), having a second primary malignancy (HR= 0.928, 95%CI 0.879-0.980), obesity (HR= 0.891, 95%CI 0.808-0.984), and blood loss anemia (HR= 0.872, 95%CI 0.788-0.964). The individual comorbidities and patient factors that were protective and decreased the hazard of all-cause mortality between 20% but less than 40% were being an other race compared to white (HR= 0.698, 95%CI 0.496-0.983) and receiving radiation therapy only (HR= 0.694, 95%CI 0.551-0.874). The individual comorbidities and patient factors that were protective and decreased the hazard of all-cause mortality by more than 40% were all treatments, in order of least to greatest, chemotherapy only (HR= 0.520, 95%CI 0.467-0.580), chemotherapy and radiation (HR= 0.362, 95%CI 0.318-0.411), surgery on the primary site only (HR= 0.183, 95%CI 0.169-0.198), and the largest reduction in hazards with the treatment combination of surgery with radiation and/or chemotherapy (HR= 0.118, 95%CI 0.108-0.128).

Model 2 includes grouped comorbidities and patient factors. Again, due to the large number of significant findings, statistically significant results will be grouped in order from largest to smallest effect. The individual comorbidities and patient factors that increased the hazard of all-cause mortality by greater than 40% that were found to be significant were late-stage (HR= 3.211, 95%CI), aids(HR= 3.083, 95%CI 2.046-4.645), being 65 or older (HR= 1.743, 95%CI 1.420-2.140), being uninsured (HR= 1.728, 95%CI 1.554-1.921), having renal failure (HR= 1.685, 95%CI 1.539-1.845), having Medicaid insurance (HR= 1.619, 95%CI 1.482-1.768), weight loss (HR= 1.542, 95%CI 1.429-1.664), and other neurological disease excluding paralysis (HR= 1.536, 95%CI 1.411-1.672). The individual comorbidities and patient factors that increased the hazard of all-cause mortality by greater than 20% but less than 40% that were found to be significant were electrolyte disorders (HR=

1.389, 95%CI 1.315-1.468), respiratory system disorders (HR= 1.373, 95%CI 1.306-1.444), having Medicare (HR= 1.341, 95%CI 1.251-1.436) Military/other insurance (HR= 1.337, 95%CI 1.093-1.636), and a substance abuse (HR= 1.202, 95%CI 1.050-1.376). The individual comorbidities and patient factors that increased the hazard of all-cause mortality by up to 20% that were found to be significant were gastrointestinal system disorders (HR= 1.160, 95%CI 1.1.057-1.274), blood system disorders (HR= 1.139, 95%CI 1.086-1.195), being of black race (HR= 1.123, 95%CI 1.042-1.210), Appalachian (HR= 1.100, 95%CI 1.052-1.150), and being male (HR= 1.083, 95%CI 1.041-1.126). The individual comorbidities and patient factors that were protective and decreased the hazard of all-cause mortality 20% were secondary primary malignancy (HR= 0.921, 95%CI 0.873-0.972), psychiatric disorders (HR= 0.915, 95%CI 0.939-0.998), and obesity (HR= 0.896, 95%CI 0.812-0.989). The individual comorbidities and patient factors that were protective and decreased the hazard of all-cause mortality between 20% but less than 40% were treatment of radiation only (HR= 0.696, 95%CI 0.553-0.877) and being an other race compared to white (HR= 0.681, 95%CI 0.483-0.959). The individual comorbidities and patient factors that were protective and decreased the hazard of all-cause mortality by more than 40% were all treatments, in order of least to greatest, chemotherapy only (HR= 0.513, 95%CI 0.460-0.571), chemotherapy and radiation (HR= 0.354, 95%CI 0.312-0.402), surgery on the primary site only (HR= 0.183, 95%CI 0.169-0.198), and the largest reduction in hazards with the treatment combination of surgery with radiation and/or chemotherapy (HR= 0.114, 95%CI 0.105-0.124).

Model 3 includes aggregate count of comorbidities and patient factors. Like the previous models, statistically significant results will be grouped in order from largest to smallest effect. The individual comorbidities and patient factors that increased the hazard of all-cause mortality by greater than 40% that were found to be significant were late-stage (HR= 3.271, 95%CI 3.116-3.434),

being uninsured (HR= 1.773, 95%CI 1.595-1.970), having Medicaid (HR= 1.694, 95%CI 1.552-1.846), being over the age of 65 (HR= 1.586, 95%CI 1.294-1.943), and having two or more comorbidities (HR= 1.495, 95%CI 1.421-1.572). The individual patient factor that increased the hazard of all-cause mortality by greater than 20% but less than 40% that were found to be significant were having a primary payer source of Medicare (HR= 1.388, 95%CI 1.296-1.486) and Military/other (HR= 1.286, 95%CI 1.051-1.573). The individual comorbidities and patient factors that increased the hazard of all-cause mortality by up to 20% that were found to be significant were being black compared to white (HR= 1.134, 95%CI 1.053-1.222), having one comorbidity (HR= 1.108, 95%CI 1.047-1.173), being male (HR= 1.099, 95%CI 1.058-1.143), and Appalachian (HR= 1.052, 95%CI 1.007-1.099). The factor that was protective and decreased the hazard of all-cause mortality 20% were having a secondary primary malignancy (HR= 0.930, 95%CI 0.881-0.981). The individual patient factors that were protective and decreased the hazard of all-cause mortality between 20% but less than 40% were treatment of radiation only (HR= 0.738, 95%CI 0.587-0.929) and being of an other race compared to white (HR= 0.631, 95%CI 0.448-0.889). The individual comorbidities and patient factors that were protective and decreased the hazard of all-cause mortality by more than 40% were again all treatments, in the same order, chemotherapy only (HR= 0.486, 95%CI 0.436-0.541), chemotherapy and radiation (HR= 0.322, 95%CI 0.284-0.365), surgery on the primary site only (HR= 0.173, 95%CI 0.160-0.187), and the largest reduction in hazards with the treatment combination of surgery with radiation and/or chemotherapy (HR= 0.106, 95%CI 0.098-0.115).

Cox proportional hazard models of CRC mortality can be found in Table 1-7 and includes 9,866 CRC patients from Kentucky diagnosed 2003-2016, there are three models viewing comorbidity through an individual, system, or aggregate count lens.

Model 1 includes Individual comorbidities and patient factors. The individual comorbidities and patient factors that increased the hazard of CRC mortality were renal failure (HR= 1.809, 95% CI 1.597-2.05), pulmonary circulation disorders (HR= 1.419, 95% CI 1.148-1.755), CHF (HR= 1.390, 95% CI 1.27-1.521), weight loss (HR= 1.327,95% CI 1.159-1.519), other neurological diseases excluding paralysis (HR=1.308, 95% CI 1.148-1.49), Medicaid (HR=1.300, 95% CI 1.106-1.529), coagulopathy (HR= 1.300, 95% CI 1.02-1.655), chronic obstructive pulmonary disorder (HR= 1.225, 95% CI 1.136-1.321), Medicare (HR= 1.215, 95% CI 1.087-1.357), electrolyte disorders (HR= 1.182, 95% CI 1.085-1.288), being male (HR= 1.120, 95% CI 1.052-1.193), deficiency anemia (HR= 1.099, 95% CI 1.01-1.196), and diabetes (HR= 1.097, 95% CI 1.017-1.184). The individual comorbidities and patient factors that were protective and decreased the hazard of CRC-specific mortality were having a second primary malignancy (HR= 0.867, 95% CI 0.805-0.933), blood loss anemia (HR= 0.807, 95% CI 0.694-0.938), treatment of radiation only (HR= 0.488, 95% CI 0.294-0.81), chemotherapy and radiation (HR= 0.367, 95% CI 0.285 - 0.473) surgery on primary site only (HR= 0.340, 95% CI 0.296-0.39), chemotherapy only (HR= 0.242, 95% CI 0.184-0.318), and surgery with radiation and/or chemotherapy (HR= 0.179, 95% CI 0.154-0.207).

Model 2 includes grouped comorbidities and patient factors. The individual comorbidities and patient factors that increased the hazard of CRC mortality were renal failure (HR= 1.949, 95% CI 1.725-2.204), neurological system (HR= 1.386, 95% CI 1.229-1.563), weight loss (HR= 1.344, 95% CI 1.174-1.539), having Medicaid (HR= 1.301, 95% CI 1.107-1.529), respiratory system (HR= 1.286, 95% CI 1.196-1.383), having Medicare (HR= 1.218, 95% CI 1.091-1.361), electrolyte disorders (HR= 1.199, 95% CI 1.101-1.305), gastrointestinal system (HR= 1.186, 95% CI 1.019-1.381), being male (HR= 1.115, 95% CI 1.049-1.187), cardiovascular disease (HR= 1.093, 95% CI 1.024-1.166), and endocrine system (HR= 1.088, 95% CI 1.013-1.168). The individual comorbidities and patient factors that were

protective and decreased the hazard of CRC-specific mortality were having a secondary primary malignancy (HR= 0.868, 95% CI 0.807-0.935), radiation therapy only (HR= 0.483, 95% CI 0.291-0.802), chemotherapy and radiation therapies (HR= 0.365, 95% CI 0.283-0.471), surgery on primary site only (HR= 0.339, 95% CI 0.296-0.389), chemotherapy only (HR= 0.241, 95% CI 0.183-0.316), and surgery with radiation and/or chemotherapy (HR= 0.177, 95% CI 0.152-0.205).

Model 3 includes an aggregate count of comorbidities and other patient factors. The other covariates that increased the hazard of CRC mortality include Medicaid (HR=1.334, 95% CI 1.135-1.567), Medicare (HR = 1.242, 95% CI 1.112-1.388), and male gender (HR = 1.142, 95% CI 1.075-1.214). The individual comorbidities and patient factors that were protective and decreased the hazard of CRC mortality were having a secondary primary malignancy(HR = 0.855, 95% CI 0.794-0.92), radiation therapy only (HR = 0.495, 95% CI 0.298-0.822), chemotherapy and radiation (HR = 0.341, 95% 0.265-0.439), surgery on primary site only (HR = 0.329, 95% CI 0.287-0.378), chemotherapy only (HR = 0.228, 95% CI 0.174-0.300), and surgery with radiation and/or chemotherapy (HR = 0.168, 95% CI 0.145-0.195).

Cox proportional hazard models of second primary malignancy can be found in Table 1-8 and includes 2,624 CRC patients from Kentucky diagnosed 2003-2016, there are three models viewing comorbidity through an individual, system, or aggregate count lens.

Model 1 includes Individual comorbidities and patient factors. The individual comorbidities and patient factors that increased the hazard of secondary primary malignancy were renal failure (HR= 1.856, 95% CI 1.428-2.412), being uninsured (HR= 1.786, 95% CI 1.304-2.445), late-stage (HR= 1.723, 95% CI 1.512-1.963), CHF (HR= 1.670, 95% CI 1.393-2.002), Medicaid (HR= 1.522, 95% CI 1.143-2.027), Medicare (HR= 1.427, 95% CI 1.181-1.724), weight loss (HR= 1.319, 95% CI 1.033-1.684), deficiency anemia (HR= 1.236, 95% CI 1.072-1.426), chronic obstructive pulmonary disorder (HR=

1.235, 95% CI 1.079-1.415), electrolyte disorders (HR= 1.173, 95% CI 1.006-1.369), being male (HR= 1.159, 95% CI 1.041-1.29), and Appalachian (HR= 1.129, 95% CI 1.001-1.275). The individual comorbidities and patient factors that were protective and decreased the hazard of secondary primary malignancy were surgery on primary site only (HR= 0.568, 95% CI 0.375-0.859) and surgery with radiation and/or chemotherapy (HR= 0.477, 95% CI 0.312-0.728).

Model 2 includes grouped comorbidities and patient factors. The individual comorbidities and patient factors that increased the hazard of secondary primary malignancy were aids (HR= 3.310, 95% CI 1.058-10.36), renal failure (HR= 1.976, 95% CI 1.529-2.552), being uninsured (HR= 1.798, 95% CI 1.314-2.460), late-stage (HR= 1.719, 95% CI 1.508-1.96), substance abuse (HR= 1.543, 95% CI 1.083-2.198), Medicaid (HR= 1.488, 95% CI 1.118-1.980), Medicare (HR= 1.444, 95% CI 1.195-1.745), weight loss (HR= 1.296, 95% CI 1.015-1.654), respiratory system (HR= 1.285, 95% CI 1.127-1.466), electrolyte disorders (HR= 1.218, 95% CI 1.046-1.420), blood system (HR= 1.198, 95% CI 1.057-1.358), and being male (HR= 1.154, 95% CI 1.039-1.281). The individual comorbidities and patient factors that were protective and decreased the hazard of secondary primary malignancy were surgery on primary site only (HR= 0.542, 95% CI 0.359-0.819) and surgery with radiation and/or chemotherapy (HR= 0.446, 95% CI 0.293-0.680).

Model 3 includes an aggregate count comorbidities and patient factors. The individual comorbidities and patient factors that increased the hazard of secondary primary malignancy were being uninsured (HR= 1.866, 95% CI 1.367-2.546), late-stage (HR= 1.713, 95% CI 1.504-1.952), Medicaid (HR= 1.595, 95% CI 1.205-2.113), Medicare (HR= 1.504, 95% CI 1.247-1.813), having two or more comorbidities (HR= 1.400, 95% CI 1.226- 1.599), and being male (HR= 1.175, 95% CI 1.059-1.303).The individual comorbidities and patient factors that were protective and decreased the

hazard of secondary primary malignancy were surgery on primary site only (HR= 0.543, 95% CI 0.361-0.818) and surgery with radiation and/or chemotherapy (HR= 0.443, 95% CI 0.291-0.673).

Kaplan Meier Survival Curves (Censoring and LifeTest) that compare early to late-stage survival of diagnosed primary CRC cases in Kentucky can be found in Figure 1-2. Early-stage failed cases (those that died during the interval) were 4,388 and 61.46% were censored (6,999). Late-stage failed cases were 6,174 and 32.77% (3,010) were censored. The rank tests for homogeneity indicate a significant difference between survival of late and early-stage initial primary CRC. The p values of the log-rank test and Wilcoxon test were both <0.001. The height of the drop in the first 24 months is very steep for late-stage below 60% CRC patients compared to early-stage that dropped just below 90%, there were more late-stage patients at risk of failing. At the end of the time period, early-stage survival probability is around 60% and late-stage is at 30%. We do not see a steep drop at the end of either curve and there is no interaction between the two curves.

Table 1-4. Elixhauser and Comorbidity Grouping Bivariate (Outcome Late-stage)

Variable	Individual Morbidities			Grouped Variable	Grouped Morbidities		
	OR	Upper 95% CI	Lower 95% CI		OR	Upper 95% CI	Lower 95% CI
CHF	0.886*	0.795	0.988	Cardiovascular System	0.880***	0.833	0.930
Valvular Disorder	0.864	0.713	1.048				
Peripheral Vascular Disease	0.873	0.740	1.030				
Hypertension	0.881***	0.833	0.931				
Pulmonary Circulation	1.047	0.823	1.331	Respiratory System	0.918*	0.849	0.993
Chronic Pulmonary	0.897**	0.828	0.971				
Liver Disease	1.432***	1.220	1.682	Gastrointestinal System	1.386***	1.200	1.601
Peptic ulcer	1.228	0.897	1.681				
Paralysis	1.033	0.709	1.506	Neurological System	0.990	0.860	1.141
Other Neurological Disease	0.971	0.836	1.128				
Diabetes	0.933	0.869	1.001	Endocrine System	0.914**	0.855	0.976
Hypothyroid	0.874*	0.774	0.987				
Coagulopathy	1.317*	1.002	1.731	Blood System	1.369***	1.274	1.471
Blood loss Anemia	1.230**	1.065	1.421				
Deficiency Anemia	1.362***	1.266	1.464				
Alcohol Abuse	1.176	0.930	1.486	Substance Abuse	1.138	0.923	1.402
Drug Abuse	0.978	0.649	1.473				
Psychosis	1.281	0.969	1.694	Psychiatric	1.178**	1.043	1.330
Depression	1.170*	1.027	1.333				
Renal Failure	0.869	0.745	1.012				
Aids	1.436	0.683	3.019				
Obesity	0.920	0.812	1.043				
Rheumatoid Arthritis	1.037	0.791	1.360				
Weight loss	1.856***	1.636	2.107				
Electrolyte	1.381***	1.267	1.504				
				Total Number of Comorbidity Groups			
				0	*ref*		
				1	1.029	0.954	1.109
				2+	1.097**	1.026	1.173
* p < 0.05 ; ** p < 0.01 ; *** p < 0.0001							

Table 1-5. Elixhauser and Comorbidity Grouping Bivariate (Outcome Death)

Variable	Individual Morbidities			Grouped Variable	Grouped Morbidities		
	OR	Upper 95% CI	Lower 95% CI		OR	Upper 95% CI	Lower 95% CI
CHF	3.914***	3.435	4.459	**Cardiovascular System**	1.262***	1.194	1.333
Valvular Disorder	1.679***	1.378	2.045				
Peripheral Vascular Disease	1.945***	1.636	2.312				
Hypertension	1.029	0.974	1.087				
Pulmonary Circulation	2.502***	1.915	3.269	**Respiratory System**	1.899***	1.752	2.058
Chronic Pulmonary	1.874***	1.728	2.033				
Liver Disease	1.620***	1.372	1.912	**Gastrointestinal System**	1.553***	1.339	1.801
Peptic ulcer	1.309	0.953	1.800				
Paralysis	3.415***	2.171	5.374	**Neurological System**	2.648***	2.261	3.101
Other Neurological Disease	2.532***	2.146	2.988				
Diabetes	1.179***	1.099	1.265	**Endocrine System**	1.121**	1.049	1.196
Hypothyroid	0.896	0.795	1.011				
Coagulopathy	2.798***	2.046	3.826	**Blood System**	1.720***	1.599	1.851
Blood loss Anemia	1.872***	1.608	2.179				
Deficiency Anemia	1.670***	1.550	1.799				
Alcohol Abuse	1.554**	1.221	1.977	**Substance Abuse**	1.631***	1.314	2.024
Drug Abuse	1.374	0.908	2.079				
Psychosis	1.680**	1.256	2.246	**Psychiatric**	0.965	0.855	1.090
Depression	0.886	0.778	1.010				
Renal Failure	2.974***	2.500	3.538				
Aids	4.366**	1.659	11.489				
Obesity	0.636***	0.560	0.721				
Rheumatoid Arthritis	1.404*	1.067	1.849				
Weight loss	2.867***	2.489	3.303				
Electrolyte	2.289***	2.088	2.508				
				Total Number of Comorbidity Groups			
				0	*ref*		
				1	1.212***	1.124	1.307
				2+	2.064***	1.929	2.208

* p < 0.05 ; ** p < 0.01 ; *** p < 0.0001

Table 1-6. Cox Proportional Hazard models of All-cause mortality, 20,270 CRC patients 2003-2016

Variable	Model 1			Model 2			Model 3		
	HR	CI low	CI High	HR	CI low	CI High	HR	CI low	CI High
Cardiovascular Disease				1.033	0.992	1.077			
CHF	1.566***	1.467	1.672						
Valvular Disorder	1.047	0.927	1.183						
Peripheral Vascular Disease	1.157**	1.044	1.281						
Hypertension	0.929**	0.892	0.968						
Respiratory System				1.373***	1.306	1.444			
Pulmonary Circulation	1.394***	1.204	1.615						
Chronic Pulmonary	1.290***	1.224	1.359						
Gastrointestinal				1.160**	1.057	1.274			
Liver Disease	1.226**	1.105	1.360						
Peptic ulcer	0.910	0.739	1.121						
Neurological System				1.536***	1.411	1.672			
Paralysis	1.347**	1.087	1.671						
Other Neurological Disease	1.487***	1.358	1.627						
Endocrine System				1.019	0.972	1.068			
Diabetes	1.068*	1.015	1.123						
Hypothyroid	0.918	0.841	1.002						
Blood System				1.139***	1.086	1.195			
Coagulopathy	1.426***	1.213	1.676						
Blood loss Anemia	0.872**	0.788	0.964						
Deficiency Anemia	1.146***	1.086	1.210						
Substance Abuse				1.202**	1.050	1.376			
Alcohol Abuse	1.202*	1.031	1.401						
Drug Abuse	1.067	0.811	1.404						
Psychiatric				0.915**	0.839	0.998			
Psychosis	1.045	0.875	1.248						
Depression	0.937	0.852	1.030						
Renal Failure	1.570***	1.432	1.720	1.685***	1.539	1.845			
Aids	2.902***	1.926	4.373	3.083***	2.046	4.645			
Obesity	0.891*	0.808	0.984	0.896**	0.812	0.989			
Rheumatoid Arthritis	1.223*	1.023	1.462	1.179	0.986	1.409			
Weight loss	1.540***	1.427	1.662	1.542***	1.429	1.664			
Electrolyte	1.353***	1.280	1.430	1.389***	1.315	1.468			
Total Comorbidity									
0 (ref)							*ref*	*ref*	*ref*
1							1.108**	1.047	1.173
2+							1.495***	1.421	1.572
Stage of Cancer (late vs early)	3.198***	3.046	3.357	3.211***	3.058	3.371	3.271***	3.116	3.434
Second primary(yes/no)	0.928**	0.879	0.980	0.921**	0.873	0.972	0.930**	0.881	0.981
Sex (Female Ref)	1.074**	1.032	1.118	1.083***	1.041	1.126	1.099***	1.058	1.143
Race (black vs white)	1.131**	1.050	1.219	1.123**	1.042	1.210	1.134**	1.053	1.222
Race (other vs white)	0.698*	0.496	0.983	0.681**	0.483	0.959	0.631**	0.448	0.889
Age or age group									
18 – 34 years (ref)	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*
35 – 44 years	0.975	0.780	1.218	0.977	0.782	1.220	0.905	0.725	1.128

45 – 54 years	0.991	0.809	1.216	0.982	0.801	1.204	0.911	0.745	1.115
55 – 64 years	1.166	0.954	1.425	1.152	0.943	1.408	1.054	0.864	1.286
65 + years	1.746***	1.421	2.143	1.743***	1.420	2.140	1.586***	1.294	1.943
Source of payment									
Private Payer (ref)	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*
Medicaid	1.617***	1.481	1.767	1.619***	1.482	1.768	1.694***	1.552	1.849
Medicare	1.338***	1.249	1.434	1.341***	1.251	1.436	1.388***	1.296	1.486
Military/Other	1.317**	1.076	1.612	1.337**	1.093	1.636	1.286*	1.051	1.573
Uninsured	1.711***	1.538	1.903	1.728***	1.554	1.921	1.773***	1.595	1.970
Appalachian	1.112***	1.064	1.163	1.100***	1.052	1.150	1.052*	1.007	1.099
Treatment									
No Treatment (ref)	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*
Chemotherapy Only	0.520***	0.467	0.580	0.513***	0.460	0.571	0.486***	0.436	0.541
Radiation Only	0.694**	0.551	0.874	0.696**	0.553	0.877	0.738**	0.587	0.929
Surgery on Primary Site Only	0.183***	0.169	0.198	0.183***	0.169	0.198	0.173***	0.160	0.187
Chemotherapy and Radiation	0.362***	0.318	0.411	0.354***	0.312	0.402	0.322***	0.284	0.365
Surgery with Radiation and/or Chemotherapy	0.118***	0.108	0.128	0.114***	0.105	0.124	0.106***	0.098	0.115
* $p < 0.05$; ** $p < 0.01$; *** $p < 0.0001$									

Table 1-7. Cox Proportional Hazard models of CRC mortality, 9,866 CRC patients 2003-2016

Variable	Model 1			Model 2			Model 3		
	HR	CI low	CI High	HR	CI low	CI High	HR	CI low	CI High
Cardiovascular Disease				1.093**	1.024	1.166			
CHF	1.390***	1.270	1.521						
Valvular Disorder	1.045	0.880	1.241						
Peripheral Vascular Disease	1.070	0.928	1.233						
Hypertension	1.011	0.949	1.077						
Respiratory System				1.286***	1.196	1.383			
Pulmonary Circulation	1.419**	1.148	1.755						
Chronic Pulmonary	1.225***	1.136	1.321						
Gastrointestinal				1.186*	1.019	1.381			
Liver Disease	1.177	0.989	1.401						
Peptic ulcer	1.154	0.846	1.572						
Neurological System				1.386***	1.229	1.563			
Paralysis	1.190	0.887	1.596						
Other Neurological Disease	1.308***	1.148	1.490						
Endocrine System				1.088*	1.013	1.168			
Diabetes	1.097*	1.017	1.184						
Hypothyroid	1.107	0.973	1.259						
Blood System				1.061	0.986	1.142			
Coagulopathy	1.300*	1.020	1.655						
Blood loss Anemia	0.807**	0.694	0.938						
Deficiency Anemia	1.099*	1.010	1.196						
Substance Abuse				1.056	0.859	1.298			
Alcohol Abuse	1.059	0.834	1.343						
Drug Abuse	0.986	0.659	1.476						
Psychiatric				1.061	0.927	1.214			
Psychosis	1.108	0.864	1.420						
Depression	1.091	0.937	1.269						
Renal Failure	1.809***	1.597	2.050	1.949***	1.725	2.204			
Aids	1.171	0.584	2.349	1.176	0.586	2.357			
Obesity	1.074	0.918	1.255	1.077	0.922	1.259			
Rheumatoid Arthritis	1.211	0.946	1.552	1.189	0.929	1.522			
Weight loss	1.327***	1.159	1.519	1.344***	1.174	1.539			
Electrolyte	1.182**	1.085	1.288	1.199***	1.101	1.305			
Total Comorbidity									
0 (ref)							*ref*	*ref*	*ref*
1							1.118*	1.015	1.232
2+							1.504***	1.384	1.634
Stage of Cancer (late vs early)	1.074	0.994	1.160	1.067	0.988	1.152	1.063	0.985	1.148
Second primary(yes/no)	0.867**	0.805	0.933	0.868**	0.807	0.935	0.855***	0.794	0.920
Sex (Male vs Female)	1.120**	1.052	1.193	1.115**	1.049	1.187	1.142***	1.075	1.214
Race (black vs white)	1.083	0.959	1.223	1.062	0.941	1.199	1.072	0.950	1.209
Race (other vs white)	1.004	0.501	2.014	0.949	0.474	1.903	0.920	0.459	1.843
Age or age group									
18 – 34 years (ref)	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*
35 – 44 years	0.829	0.437	1.573	0.836	0.441	1.585	0.821	0.433	1.557

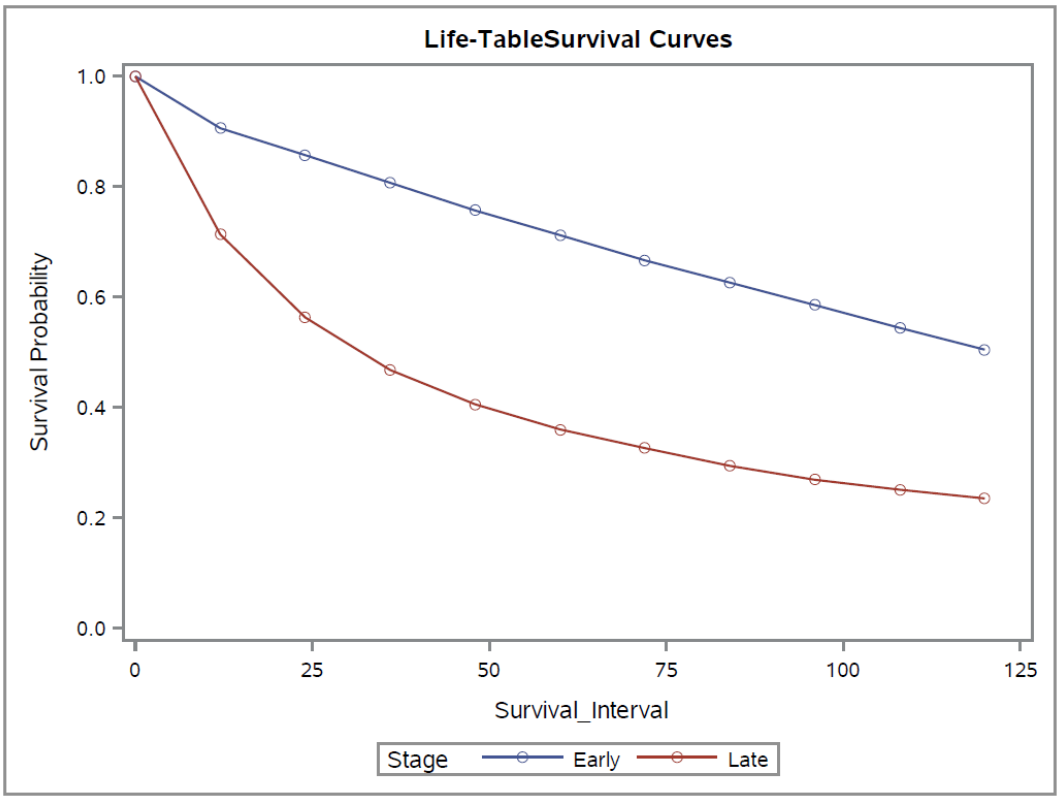
45 – 54 years	1.001	0.559	1.793	0.986	0.551	1.765	0.978	0.546	1.750
55 – 64 years	1.274	0.717	2.264	1.261	0.710	2.241	1.229	0.692	2.183
65 + years	1.496	0.839	2.668	1.499	0.841	2.672	1.470	0.825	2.619
Source of payment									
Private Payer (ref)	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*
Medicaid	1.300**	1.106	1.529	1.301**	1.107	1.529	1.334**	1.135	1.567
Medicare	1.215**	1.087	1.357	1.218**	1.091	1.361	1.242**	1.112	1.388
Military/Other	0.990	0.645	1.519	0.979	0.638	1.502	0.978	0.638	1.501
Uninsured	1.179	0.935	1.488	1.174	0.930	1.481	1.180	0.935	1.488
Appalachian	1.016	0.946	1.091	1.002	0.934	1.076	0.990	0.923	1.061
Treatment									
No Treatment (ref)	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*
Chemotherapy Only	0.242***	0.184	0.318	0.241***	0.183	0.316	0.228***	0.174	0.300
Radiation Only	0.488**	0.294	0.810	0.483**	0.291	0.802	0.495**	0.298	0.822
Surgery on Primary Site Only	0.340***	0.296	0.390	0.339***	0.296	0.389	0.329***	0.287	0.378
Chemotherapy and Radiation	0.367***	0.285	0.473	0.365***	0.283	0.471	0.341***	0.265	0.439
Surgery with Radiation and/or Chemotherapy	0.179***	0.154	0.207	0.177***	0.152	0.205	0.168***	0.145	0.195
* $p < 0.05$; ** $p < 0.01$; *** $p < 0.0001$									

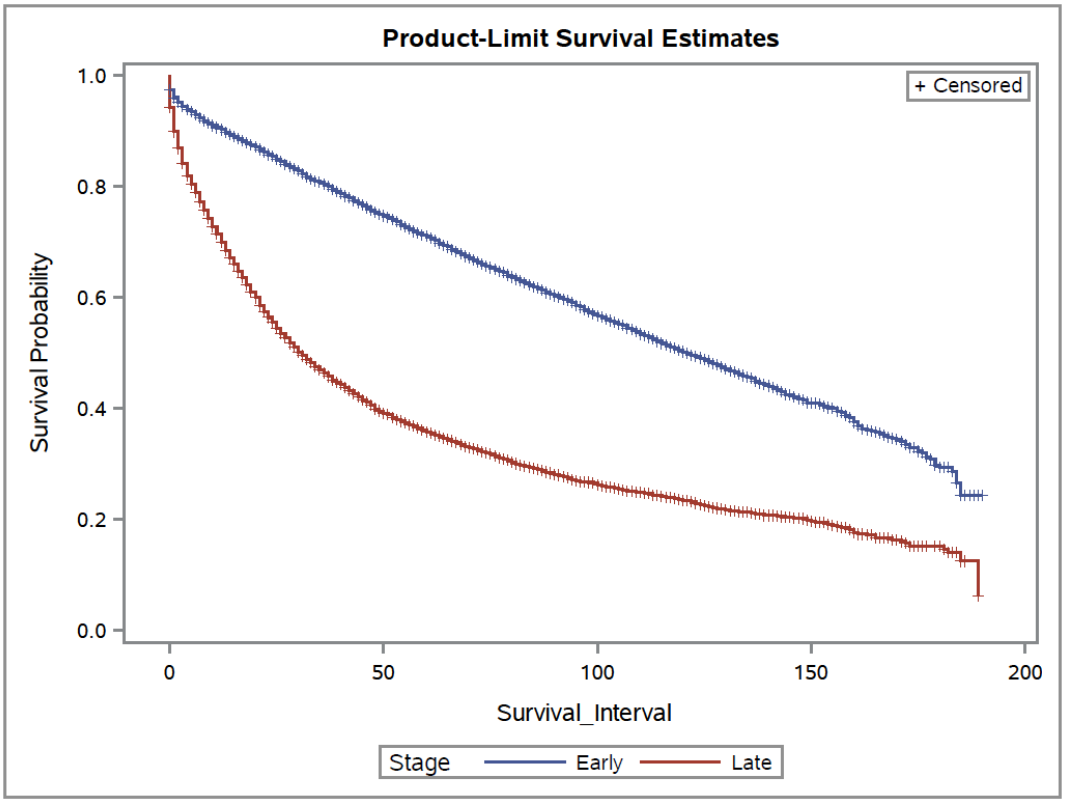
Table 1-8. Cox Proportional Hazard models of Second Primary Cancer, 2,624 CRC patients 2003-2016

Variable	Model 1			Model 2			Model 3		
	HR	CI low	CI High	HR	CI low	CI High	HR	CI low	CI High
Cardiovascular Disease				1.057	0.949	1.179			
CHF	1.670***	1.393	2.002						
Valvular Disorder	1.110	0.784	1.572						
Peripheral Vascular Disease	1.010	0.747	1.365						
Hypertension	0.981	0.879	1.095						
Respiratory System				1.285**	1.127	1.466			
Pulmonary Circulation	1.299	0.806	2.095						
Chronic Pulmonary	1.235**	1.079	1.415						
Gastrointestinal				1.258	0.977	1.621			
Liver Disease	1.211	0.902	1.626						
Peptic ulcer	1.308	0.804	2.130						
Neurological System				1.237	0.959	1.595			
Paralysis	1.338	0.710	2.519						
Other Neurological Disease	1.150	0.870	1.520						
Endocrine System				1.033	0.910	1.173			
Diabetes	1.082	0.944	1.239						
Hypothyroid	0.953	0.747	1.217						
Blood System				1.198**	1.057	1.358			
Coagulopathy	0.950	0.576	1.568						
Blood loss Anemia	0.921	0.713	1.189						
Deficiency Anemia	1.236**	1.072	1.426						
Substance Abuse				1.543*	1.083	2.198			
Alcohol Abuse	1.550	1.069	2.247						
Drug Abuse	1.114	0.330	3.754						
Psychiatric				0.887	0.675	1.167			
Psychosis	1.359	0.740	2.495						
Depression	0.846	0.628	1.140						
Renal Failure	1.856***	1.428	2.412	1.976***	1.529	2.552			
Aids	3.018	0.961	9.477	3.310*	1.058	10.360			
Obesity	0.971	0.755	1.248	0.992	0.775	1.270			
Rheumatoid Arthritis	0.841	0.544	1.299	0.818	0.530	1.263			
Weight loss	1.319*	1.033	1.684	1.296*	1.015	1.654			
Electrolyte	1.173*	1.006	1.369	1.218*	1.046	1.420			
Total Comorbidity									
0 (ref)							*ref*	*ref*	*ref*
1							1.106	0.951	1.285
2+							1.400***	1.226	1.599
Stage of Cancer (late vs early)	1.723***	1.512	1.963	1.719***	1.508	1.960	1.713***	1.504	1.952
Sex (Male vs Female)	1.159**	1.041	1.290	1.154**	1.039	1.281	1.175**	1.059	1.303
Race (black vs white)	1.061	0.868	1.061	1.061	0.870	1.295	1.083	0.890	1.318
Race (other vs white)	0.508	0.162	0.508	0.464	0.148	1.454	0.475	0.152	1.483
Age or age group									
18 – 34 years (ref)	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*
35 – 44 years	0.757	0.292	1.961	0.758	0.293	1.962	0.753	0.292	1.946

45 – 54 years	0.609	0.247	1.500	0.607	0.246	1.495	0.581	0.236	1.429
55 – 64 years	0.691	0.283	1.690	0.691	0.283	1.689	0.665	0.273	1.622
65 + years	1.027	0.416	2.534	1.034	0.419	2.550	0.987	0.401	2.428
Source of payment									
Private Payer (ref)	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*
Medicaid	1.522**	1.143	2.027	1.488**	1.118	1.980	1.595**	1.205	2.113
Medicare	1.427**	1.181	1.724	1.444**	1.195	1.745	1.504***	1.247	1.813
Military/Other	1.907	0.936	3.885	1.800	0.884	3.666	1.863	0.916	3.787
Uninsured	1.786**	1.304	2.445	1.798**	1.314	2.460	1.866***	1.367	2.546
Appalachian	1.129*	1.0001	1.275	1.121	0.994	1.265	1.068	0.949	1.202
Treatment									
No Treatment (ref)	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*
Chemotherapy Only	1.494	0.857	2.604	1.507	0.865	2.624	1.591	0.916	2.765
Radiation Only	1.812	0.503	6.526	1.436	0.427	4.835	1.602	0.481	5.334
Surgery on Primary Site Only	0.568**	0.375	0.859	0.542**	0.359	0.819	0.543**	0.361	0.818
Chemotherapy and Radiation	1.124	0.614	2.058	1.046	0.572	1.912	0.975	0.535	1.775
Surgery with Radiation and/or Chemotherapy	0.477**	0.312	0.728	0.446**	0.293	0.680	0.443**	0.291	0.673
* $p < 0.05$; ** $p < 0.01$; *** $p < 0.0001$									

Figure 1-2. Kaplan Meier Survival Curves (Censoring and LifeTest) Comparing Early and Late Stage Survival of Diagnosed Primary CRC Patients in Kentucky





Discussion

Many of the variables across all nine models in the Cox proportional hazard models were statistically significant; the all-cause mortality had the highest number of statistically significant variables. The individual comorbidities tended to have a larger effect size compared to the grouped system comorbidities and the overall comorbidity count, although some demographic/ patient factor effects tended to slightly increase as we moved across the models.

In the all-cause mortality models, there were several personal factors that remained significant across all three models. The biggest effect was seen in late-stage (M1: HR= 3.198, 95% CI 3.046-3.357, M2: HR= 3.211 95% CI, 3.058-3.371, M3: HR= 3.271, 95% CI 3.116-3.434). What this means is that if a patient has late-stage CRC they have 3.198 / 3.211/ or 3.271 (corresponding to model progress 1-2-3) times the hazard of dying compared to those patients that have early-stage CRC. Others factors that remained significant and also had an increased hazard of all-cause mortality across all three models were 65 and older, Appalachian, males, black, having Medicaid, Medicare, Military/other, and uninsured compared to having a private payer insurance, weight loss, renal failure, aids, paralysis, and other neurological disease excluding paralysis . There were also protective findings. The smallest effect was seen in the treatment of surgery at primary site with radiation and/or chemotherapy (M1: HR= 0.118, 95%CI 0.108-0.128, M2: HR= 0.114 95%CI 0.105-0.124, M3: HR= 0.106 95%CI 0.098-0.115). What this means is that if a patient has the combination treatment of surgery with radiation and/or chemotherapy they have only 10.8% / 10.5% / or 11.8% (corresponding to model progress 1-2-3) of the hazard of dying compared to those patients that do not receive treatment. This is an average reduction of about 88% in the hazard of dying. Other protective factors that were seen across the model were hypertension, blood loss anemia, obesity, having a second

primary, being other race compared to white, and all treatments (chemotherapy only, radiation only, surgery at primary site only, chemotherapy and radiation, as well as the above mentioned surgery with radiation and/or chemotherapy. It is interesting that any treatment reduced the hazards of all-cause mortality by 30% - 88% compared to having no treatment. The all treatment combination performed the best and radiation performed the worst, but all reduced the hazards of dying compared to no treatment. Other comorbidities that were significant on the individual level, when grouped lost their effect. For example, CHF (HR= 1.566, 95%CI 1.467-1.672), peripheral vascular disease (HR= 1.157, 95%CI 1.044-1.281), and hypertension (HR= 0.929, 95%CI 0.892-0.968) were combined with valvular disorder (not found to be significant) and the protective finding of hypertension. The increased hazard of CHF and peripheral vascular disorder were washed out in the cardiovascular disease grouped mortality, which was not significant. The same is true with the endocrine system group where diabetes was individually significant and hypothyroidism was not, the effect was lost in the group. Model 3, the aggregate count of comorbidity with the personal factors, all variables were significant except one age group, 55-64 years old.

In the CRC mortality models, there were fewer comorbidities and clinical factors that were significant and the effects were not as high as all-cause mortality. The biggest statistically significant effect in CRC-cause mortality was in renal failure (M1: HR= 1.809, 95%CI 1.597-2.050), M2: HR= 1.949 95% CI 1.725-2.204)). This means is that if a patient has renal failure they have 1.809 or 1.949 times the hazard of dying from CRC mortality compared to those patients that do not have renal failure. Other factors that remained significant and also had an increased hazard of CRC mortality across all three models were CHF only and the aggregated category cardiovascular disease, both pulmonary circulation and chronic obstructive pulmonary disorders and the aggregated category respiratory system, neurological disorders but not paralysis and the aggregated category neurological system,

and diabetes only and the aggregated category endocrine system, weight loss, electrolyte disorders, being male, and having Medicaid or Medicare compared to a private payer insurance. Both aggregate counts of comorbidity were statistically significant, having one comorbidity increased the hazard of dying by 11.8% and having two or more comorbidities increased the risk of dying by 50.4% compared to those who have no comorbidities. There were also protective effects, the smallest of which, like that of all-cause mortality was seen in the treatment of surgery at primary site with radiation and/or chemotherapy (M1: HR= 0.179 95%CI 0.154-0.207, M2: HR= 0.177, 95%CI 0.152-0.205, M3: HR= 0.168 95%CI 0.145-0.195). This means is that if a patient has the treatment combination of surgery with radiation and/or chemotherapy they have only 17.9% / 17.7% / or 16.8% of the hazard of dying compared to those patients who do not receive any of the three treatment options in this study (surgery, radiation, chemotherapy). This is an average reduction of about 82% in the hazard of dying. Other protective factors that were seen across the models were blood loss anemia, having a second primary, and like all-cause mortality, any treatment reduced the hazard of dying from CRC-specific mortality. All treatments reduced the hazard by at least 51%, ranging up to 82.1% reduction in the hazards by receiving some treatment compared to no treatment. What is interesting in this set of Cox models is that some effects are not present in variables that we would have expected, those were late-stage, age, Appalachian, and the uninsured.

In the second primary malignancy models, there were some personal factors and comorbidities that were significant. The biggest effects were seen in renal failure (M1: HR= 1.856, 95%CI 1.428-2.412, M2: HR= 1.976, 95%CI 1.529-2.552) and late-stage CRC (M1: HR= 1.723, 95%CI 1.006-1.369, M2: HR= 1.719, 95%CI 1.508-1.960, M3: HR= 1.713 95%CI 1.504-1.952). This means that if a patient has renal failure, they have 1.856 or 1.976 times the hazard of developing a second primary malignancy compared to those who do not have renal failure. The same is true for late-stage

CRC patients, they have 1.723 / 1.719/ or 1.713 times the hazard of having a secondary primary malignancy compared to those patients who have early-stage CRC. The other factors that were significant and had an increased hazard of all-cause mortality across all three models were CHF, chronic obstructive pulmonary disorder up to the grouped respiratory system, deficiency anemia up to the grouped blood system, substance use grouped only, weight loss, electrolyte disorders, being male, having Medicaid, Medicare, or uninsured compared to private payer insurance, and having two or more comorbidities compared to none. There were also protective findings. The smallest effect, like that of all-cause mortality was seen in the treatment of surgery at primary site with radiation and/or chemotherapy (M1: HR= 0.477, 95%CI 0.312-0.728, M2: HR= 0.446 95%CI 0.293-0.680, M3: HR= 0.443 95%CI 0.291-0.673). If a patient has the treatment combination of surgery with radiation and/or chemotherapy they have only 47.7% / 44.6% / or 44.3% of the hazard of dying compared to those patients that do not receive treatment. The other protective factor that were seen across the models was surgery at the primary site only reduced the hazard of dying from CRC-specific mortality. The only protective factors were treatment that included some sort of surgery compared to no treatment. It is also really interesting that we did not see any statistically significant findings in any of the models of expected variables such as age groups, race, and any treatments that did not include surgery. Surgery has been shown to produce long-term survival rates and can be performed safely with low mortality.[67, 68] Although the remaining treatments were non-significant, this was the only regression where any treatment at all compared to no treatment didn't reduce the hazards. Chemotherapy, radiation, and the combination of both had a non-significant effect of increasing the hazards of secondary primary malignancy. The National Cancer Institute has noted that cancer treatments such as radiation and chemotherapy may increase the risk of second primary cancers, other studies have also found that having any combination of these cancer treatments increase the

likelihood of developing a second primary malignancy.[69-71] While our results were not significant, it is in line with previous research and helps validate the results that we are seeing in this study.

Comparing between the models, the mortality models seem to be much more similar to each other than the secondary primary model. Late-stage increased the hazard of dying in both mortality models and the secondary primary malignancy model, but was only significant in the all-cause mortality model. Obesity decreased the hazard of all-cause mortality and secondary primary but increased the hazard of CRC mortality. The individual comorbidities seemed to highlight the effect better than the collapsed and aggregate comorbidity count; some significant relationships of individual comorbidities were obscured through the process of aggregations.

The Kaplan Meier survival curves (Censoring and LifeTest) comparing early and late-stage survival shows the vast difference between the two stages. The initial steep drop in the first 24 months in late-stage is below 60% survival in this population of CRC patients compared to early-stage that dropped just below 90%. By 10 years, the late-stage survival was half of the early-stage survival at about 60% survival for early-stage and about 30% for late-stage. The 2015 U.S. combined CRC relative 5-year survival rates were 64% and 10-year was 58%[3]. Early-stage CRC has a 5-year survival rate of 90% but it declines to 71% and 14% for late-stages.[3] The curves in this study are consistent with U.S. survival rates, further validating the results of the study.

Limitations

There are several limitations to this study. The first is potential selection bias, there were a total of 28,229 patients identified by KCR that were diagnosed with initial primary CRC during the study time period. After missing data exclusions, the final total was 20,571 patients, an exclusion of 7,658 or 23.1%% of the total identified patients. The missing comorbidities data could be due to the reality that many reporting hospitals are not part of the Commission of Cancer (CoC), a group that

requires comprehensive standardized data collection.[19, 20] The second major limitation is the potential for comorbidities to be under reported to the registry. Comorbidities in the dataset are captured at the time of diagnosis. There is a potential that the reporting facility may not fully account for any comorbidities diagnosed prior to the primary CRC diagnosis. Starfield *et al*, studied a small subset of Medicare patients and found that higher morbidity burden was associated with more medical visits.[72] Another limitation in this study is the censoring of cases with cause of death coded as 777.7; the cause of death for these cases has been provided to KCR by the National Death Index which restricts use of data and cannot be released.[73] It is unknown if the 777.7 coded cases could have had a CRC related mortality; all of those cases were censored, thus the results of the CRC mortality Cox proportional hazard models could be skewed. The last limitation is that this study included stage 0 CRC within the early stage group.

Although progress has been made in the last few decades in understanding CRC, there is still a paucity of data examining the impact of comorbidities on cancer survival and secondary primary malignancies in relation to comorbidities. This research identifies that comorbidity burden increases the hazards of all-cause and CRC mortality. Further direction of research should be to examine the gap in knowledge of the role that comorbidity burden has on the standards of care and adherence to care for CRC patients.

Chapter 4

Paper 2: Spatio-temporal Analysis of Elixhauser Comorbidity Groupings and Stage of Diagnosis among Colorectal Cancer Cases in Kentucky (2003-2016)

Background

Colorectal cancer (CRC), includes any cancer that affects the large intestine of the gastrointestinal system, made up of the colon and rectum.[7] Previously CRC had a low incidence rate, however it is now the third most common type of cancer and the third most common cause of cancer death among men and women in the United States.[1-3] The American Cancer Society estimates that in 2020 there will be 147,950 new cases of CRC and 53,200 deaths from the disease in the U.S..[3] Kentucky had the highest CRC incidence in nation with 49.2 (per 100,000) for years 2012-2016 and ranked 5th in the nation in CRC mortality at a rate of 16.4 (per 100,000) for years 2013-2017.[3-6]

We do not know the exact cause of CRC, however, there are many known risk factors associated with CRC[2]. The risk of developing CRC increases with age, adults over the age of 50 have the highest CRC burden than any age group.[3, 10] Other known risk factors for CRC include family history, African American race, history of polyps, history of radiation therapy, inherited and inflammatory diseases, and lifestyle factors like low physical activity, smoking, alcohol consumption, obesity, and diets consisting of high-fat content.[2, 7]

The Appalachian region includes the state of West Virginia and counties from 12 other states, including 54 counties in the eastern half of Kentucky.[53, 54] Eastern Kentucky is markedly rural and less populated than other regions of the state, yet experiences higher rates of mortality and morbidity than the rest of the state.[5] Health disparities have been well documented in rural Appalachia.[5, 53]

A comorbidity is defined as a disease or condition that exists simultaneously with another index condition of interest, in this study CRC is the index condition of interest.[38, 39] The presence of comorbidity with an index condition, like CRC, has become increasingly more common with majority of the evidence supporting the highest comorbidity burden is concentrated in patients that are older, in minority groups, and living in poverty stricken areas.[39] In this study, comorbidities will be looked at on an individual level, grouped level, and an Elixhauser comorbidity index count that can be found in table 2-1. The Elixhauser comorbidity index includes 29 individual comorbid conditions, that were initially selected and refined by examining the literature.[41, 42] Rural Appalachia also experiences high prevalence rates of stroke, obstructive pulmonary disease, heart disease, and diabetes than non-Appalachian areas.[5, 53-55] The most prevalent comorbidity among CRC patients had been found to be diabetes.[8, 9] Patients with diabetes have an estimated 25% increased risk of developing CRC compared to patients without diabetes.[9]

Socioeconomic, behavioral, and geographical factors exacerbate health disparities and disease in Appalachia.[53, 55] Socioeconomic factors poverty, low literacy rates, lack of health insurance coverage, long traveling distances to healthcare providers, behavioral factors such as high rates of obesity and smoking, low physical activity, and environmental exposures can impact disease prevalence and screening in Appalachia.[5, 53, 55, 74] The distribution of comorbidity across the state is easier to understand when looking at a map of disease information compared to typical research tables showing numeric data.[58]

Geographic information systems (GIS) are used in epidemiological research to identify a visual location or "where" of disease.[58] Spatial data in public health studies aids researchers in visualizing disease across geographic areas, allowing for easy identification of health trends over time.[59, 60] Spatial analysis can help to determine patterns or clusters in geographic areas that that can be used

to understand patient populations at higher risk and highlight areas for intervention in addressing health disparities.[59, 60, 75]

The purpose of the current GIS project is to explore the impact of comorbidities on stage at diagnosis among CRC patients in Kentucky by examining geographical distribution of comorbidities and comparing maps of late-stage and cluster analysis. Those included were age 18 and older and diagnosed between January 1, 2003 and December 31, 2016. To date, there is no known study that has compared the geographical distribution of CRC stage at diagnosis and comorbidities across Kentucky or the nation. To address these gaps in the literature, we conducted secondary data analysis on CRC patients in Kentucky using data from the Kentucky Cancer Registry. The specific aim for this study was to perform a space-time cluster spatial analysis by mapping cases (late-stage) and controls (early) of patients by year of diagnosis across Kentucky and investigate geographical distribution of comorbidities across county at diagnosis. Depending on findings, recommendations for a systematic approach in using the clusters to identify geographic targets where public health interventions with screening would be recommended to help reduce the risk of late-stage diagnosis. Late-stage CRC is considered preventable.[76] The outcomes of this study will hopefully highlight the geographic regions to target that could potentially reduce the number of late-stage diagnoses.

Methods

Study Design and Data Sources

This is a matched case-control study of CRC cases in Kentucky. We started out with 28,229 incident cases of primary CRC diagnosed between January 1, 2003 and December 31, 2016 and initially excluded a combined total of 7,658 cases due to missing information. Excluded cases were 6,054 with missing morbidity information and 2,730 (1,126 of these were also missing morbidity

information and included in the above number) with missing stage, leaving a sample size of 20,571 cases. Patients were then matched on a one to one basic by age group (18-49, 50-74, 75+ years) and sex (M/F) resulting in a total of 18,170 included patients, 9,085 late-stage cases matched to 9,085 early-stage controls. All data were identified from the Kentucky Cancer Registry (KCR). Kentucky is funded by Surveillance, Epidemiology, and End Results Program (SEER) and National Program of Cancer Registries (NCPR); the umbrella program, North American Association of Central Cancer Registries (NAACR), independently evaluates data collected by KCR for completeness, accuracy, and timeliness.[20, 27] Requested data from KCR included first cases of primary CRC. Included cases could have multiple primaries after initial their CRC, however, any cases with CRC as a primary cancer or metastasis after another cancer diagnosis was excluded. Approval for this study was granted by the University of Kentucky Institutional Review Board.

Variables

Sex, age at diagnosis, race, marital status at diagnosis, county at diagnosis, Appalachian status, vital status, best stage group, comorbidity diagnoses (up to 10 independent variables of ICD-9-CM diagnosis codes), and secondary diagnoses (up to 10 independent variables of ICD-10-CM diagnosis codes) were provided by KCR. Age at diagnosis was categorized into three age groups, 18 - 49, 50 - 74, and 75+ years. Race was categorized as white, black, and other. Marital status at diagnosis was categorized at married, single, or other. Primary payer was categorized as Medicaid, Medicare, military/other, private pay, and not insured. County at diagnosis was originally coded with a numeric identifier and then recoded to match the name of the county in Kentucky; all 120 counties in Kentucky were represented in the data.

Comorbidity was measured using the comorbidity and secondary diagnosis variables entered into the Healthcare Cost and Utilization Project's Elixhauser Comorbidity Software (Version 3.7 for

ICD-9CM codes and the ICD-10-CM version) created by the Agency for Healthcare Research and Quality.[41, 45] Diagnosis codes were processed using a SAS program macro that classifies Elixhauser Comorbidity variables, outputting individual binary variables for the 31 (Version 3.7)/29 Elixhauser morbidity groups. The final variables were combined to match the most up-to-date Elixhauser index. Exceptions include the omission of any cancer related comorbidities, and combination of the two diabetes categories (with and without chronic complications) into one.

The Adult Comorbidity Evaluation-27 (ACE-27) index was also considered, but unfortunately available data did not allow us to grade severity within the ACE-27 index. Regardless, studies have shown that the Elixhauser measures outperform other comorbidity indices.[46, 48, 49, 64, 66] Table 2-1 shows the morbidity mapping used from ACE-27 groups to the Elixhauser Comorbidity index to the final inclusion of comorbidities (individual and grouped). KCR comorbidity and secondary diagnosis variables include codes for patients known to have no morbidity (comorbidity diagnosis code of 0000 or a secondary diagnosis entry of 0), patients with corresponding entries in either diagnoses code variables were treated as having no morbidity. Patients with a diagnosis code in the comorbidity/secondary diagnosis variables that did not match with an Elixhauser group were also treated as having no morbidity. Patients lacking comorbidity and secondary diagnosis data were considered to have unknown morbidity status and thus excluded from the study. Not all facilities reporting to KCR are part of the Commission of Cancer (CoC), a group that requires comprehensive standardized data collection including comorbidity information.[19, 20] Stage was dichotomized to reflect early (stage 0, I, II) or late (III and IIII) stage disease. Patients with unknown cancer stage was also excluded. Patients were then matched one-to-one on age group and sex by cancer stage (early or late). There were 2,301 more controls than cases and there were 100 cases that did not have enough controls to match in the corresponding age and sex groups. The 2,401 patients who did not have a

match were excluded. There were 9,085 cases and controls included in the final analysis. Figure 2-1 shows the flow chart for patient inclusion and exclusion in the study.

Statistical and Spatial Analysis

The statistical software, SAS version 9.4 was used to examine patient demographics and disease characteristics, and fit a logistic regression model of late-stage diagnosis in relation to each of the Elixhauser-based comorbidity variables.[21] A retrospective space-time cluster analysis using the Bernoulli case-control model constrained to clusters no larger than 35% of the population at risk and 50% of the study period (2003-2016) was performed with SaTScan software.[22] SaTScan was required to perform the analysis because standard GIS software packages do not have this function.[61] The purpose of SaTScan cluster analysis was to perform a geographical surveillance of CRC to try to detect areas with high or low rates of significance (Figure 2-4).[22] ArcGIS was used for mapping the comorbidities and proportion of cancers that were late-stage within each county, and mapping the cluster found in the SaTScan analysis.[23] Thus we use both SaTScan and ArcGIS to complement to each other, we exported the cluster analysis file and joined to ArcGIS for mapping purposes.[62]

Data were aggregated based on county. The number of aggregated late-stage cases within each county were then divided by the total of cases and controls (early-stage) within each county, this gave us the proportion to map. The comorbidity maps were designed the same way, individual comorbidity aggregated counts were divided by the total number of cases and controls in each county, and combined late and early-stage percentage of two or more comorbidities were also treated this way. All maps used the data classification of Jenks natural breaks. Natural breaks are data specific classifications that are based on natural groupings within the data with similar values, they are used as a means to maximize differences between the classification percentages.[77]

Table 2-1. Morbidity Mapping ACE-27 Index, Elixhauser, and Final Inclusion Study Comorbidity and Groupings

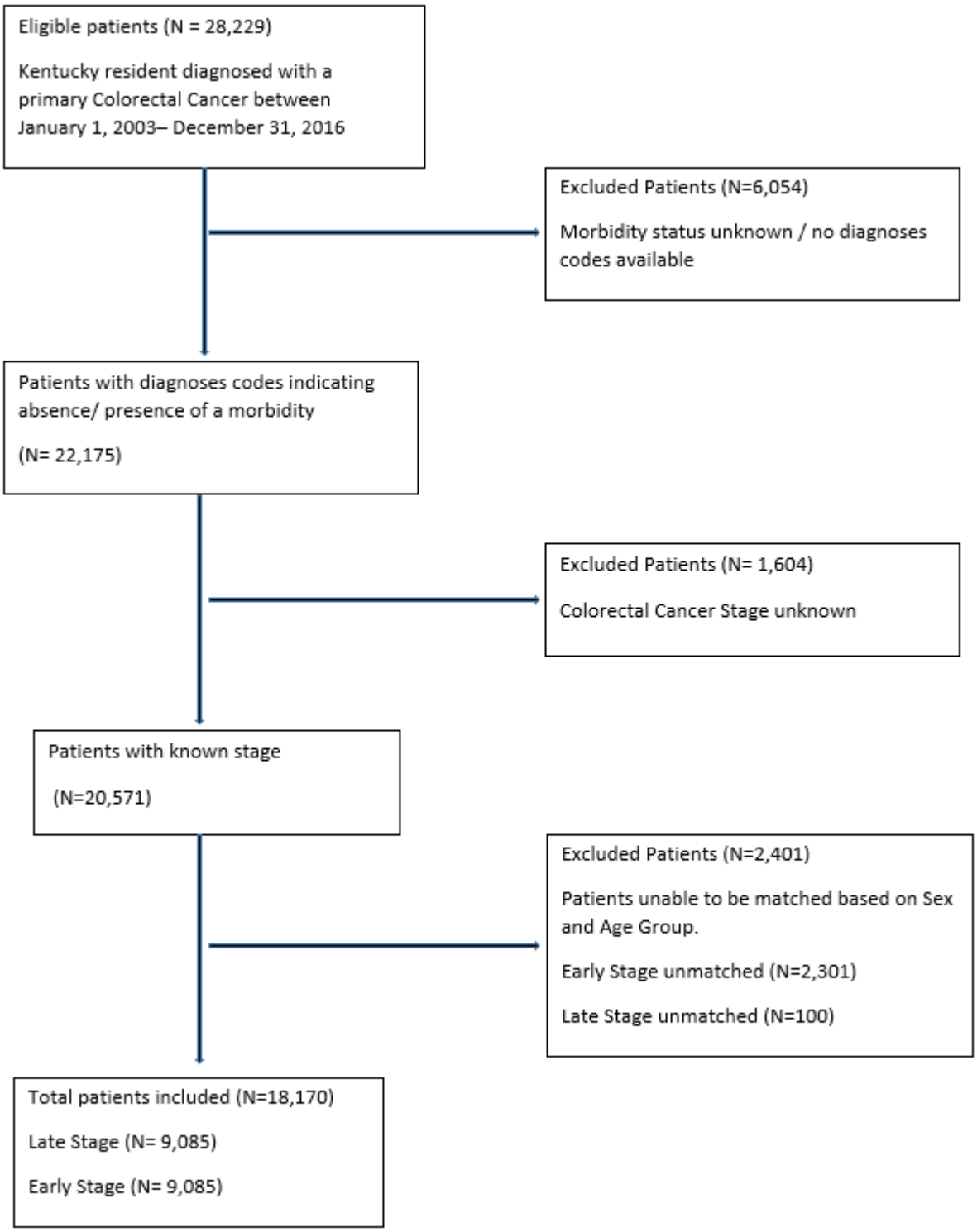
Ace-27 Index	Elixhauser ICD 10*	Final Inclusion
Cardiovascular System	Congestive Heart Failure	**Cardiovascular System**
Myocardial Infarct	Valvular disease	Congestive Heart Failure
Angina / Coronary Artery Disease	Pulmonary circulation disorders	Hypertension
Congestive Heart Failure (CHF)	Peripheral vascular disease	Peripheral Vascular Disorder
Arrhythmias	Hypertension (Complicated &	Valvular Heart Disease
Hypertension	Uncomplicated)	**Respiratory System**
Venous Disease	Paralysis	Pulmonary Circulation Disorders
Peripheral Arterial Disease	Other neurological disorders	Chronic Pulmonary
Respiratory System	Chronic pulmonary disease	**Gastrointestinal System**
Restrictive Lung Disease or COPD	Diabetes Uncomplicated	Liver Diseases
Other markers not diagnoses	Diabetes Complicated	Peptic Ulcer Disease
Gastrointestinal System	Hypothyroidism	**Renal System**
Hepatic	Renal failure	Renal Failure
Stomach / Intestine	Liver disease	**Endocrine System**
Pancreas	Chronic peptic ulcer disease	Diabetes (Complicated & Uncomplicated)
Renal System	HIV and AIDS	Hypothyroidism
End-stage renal disease	Lymphoma	**Neurological System**
Endocrine System	Metastatic cancer	Paralysis
Diabetes Mellitus	Solid tumor without metastasis	Other Neurological Disorders
Neurological System	Rheumatoid arthritis/ collagen vascular	**Psychiatric**
Stroke	diseases	Psychoses
Dementia	Coagulation deficiency	Depression
Paralysis	Obesity	**Rheumatologic**
Neuromuscular	Weight loss	Rheumatoid Arthritis/Collagen
Psychiatric	Fluid and electrolyte disorders	**Immunological System**
Recent suicidal attempt	Blood loss anemia	AIDS/HIV
Schizophrenia	Deficiency anemias	**Substance Abuse**
Depression or bipolar disorder	Alcohol abuse	Alcohol Abuse
Rheumatologic	Drug abuse	Drug Abuse
Rheumatoid Arthritis	Psychoses	**Body weight**
Systemic Lupus	Depression	Obesity
Mixed Connective Tissue Disorder		**Blood System**
Polymyositis		Coagulopathy
Rheumatic Polymyositis		Blood Loss Anemia
Immunological System		Deficiency Anemia
AIDS		**Remain but not grouped**
Malignancy		Weightloss
Solid Tumor including melanoma		Fluid & Electrolyte Disorders
Leukemia and Myeloma		
Lymphoma		
Substance Abuse		
Alcohol		
Illicit Drugs		
Body Weight		
Obesity		

*Elixhauser ICD-9 (31 Variables) conversion to ICD-10 (29 Variables) version changes: Removal of Cardiac Arrhythmia and Combining of Hypertension with and without complications. Some data were collected prior to the 2015 ICD-10 activation, combination and conversion of ICD-9 and ICD-10 Elixhauser adjustment occurred.

Table 2-2. Patient Demographics

Demographics	Early Stage (I-II) (N=9085)	Late Stage (III – IV) (N=9085)
Age		
18 - 49 years	2,157 (11.87%)	1,128 (12.42%)
50 – 74 years	11,186 (61.56%)	5,625 (61.92%)
75+ years	4,827 (26.57%)	2,332 (25.67%)
Marital Status		
Missing	184 (2.03%)	191 (2.10%)
Married	5,269 (58.00%)	5,079 (55.91%)
Single	972 (10.70%)	1,026 (11.29%)
Other	2,660 (29.28%)	2,789 (30.70%)
Sex		
Female	4,374 (48.15%)	4,374 (48.15%)
Male	4,711 (51.85%)	4,711 (51.85%)
Race		
White	8,360 (92.02%)	8,374 (92.17%)
Black	668 (7.35%)	646 (7.11%)
Other	57 (0.63%)	65 (0.72%)
Diabetes	1,998 (21.99%)	1,653 (18.19%)
Renal Failure	362 (3.98%)	290 (3.19%)
Liver Disease	277 (3.05%)	328 (3.61%)
CHF	780 (8.59%)	601 (6.62%)
Hypertension	4,983 (54.85%)	4,053 (44.61%)
Appalachian	2,462 (27.10%)	2,570 (28.29%)
Total Comorbidity Groups		
0	1,357 (14.94%)	2,283 (25.13%)
1	2,880 (31.70%)	2,443 (26.89%)
2+	4,848 (53.36%)	4,359 (47.98%)
Vital Status		
Alive	5,439 (59.87%)	2,964 (32.63%)
Dead	3,646 (40.13%)	6,121 (67.37%)

Figure 2-1. Flow chart for participant selection and inclusion/exclusion criteria for matched study



Results

The choropleth morbidity maps, found in Figure 2-2, show the geographical distribution of the four individual comorbidities, electrolyte disorders, liver disease, weight loss, and deficiency anemia across Kentucky. The four maps do not demonstrate any geographical patterns in the distribution of comorbidities. The map of comorbidities among CRC patients is displayed in Figure 2-3. This map shows that a large percentage of CRC patients experience two or more comorbidities, but does not demonstrate a strong pattern of disease.

The percentage of late-stage cancer by county appears in Figure 2-4 with a retrospective space-time cluster analysis utilizing SaTScan found one cluster with the highest likelihood and statistically significant high-rate cluster of late-stage CRC in southeastern Kentucky. The time period for this cluster was limited to January 1, 2013 to December 31, 2016. This map does not exhibit a strong geographical pattern of late-stage cancer, but does show there is a high burden of late-stage CRC within the majority of the counties and within the cluster area. There were 751 total observed late-stage cases, while the expected number of late-stage cases was only 612.5, a ratio of 1.23 observed late-stage diagnoses for each one expected($p < 0.0001$).

The bivariate model results in Table 2-2 show that some individual comorbidities had a protective effect with regard to late-stage diagnosis of CRC, while other comorbidities appeared to be a risk factor for late-stage diagnosis of CRC. The individual comorbidities with statistically significant reduced odds of late-stage CRC were found in CHF (OR= 0.754, 95% CI 0.675-0.842), valvular disorder (OR= 0.725,95% CI 0.597 - 0.881), peripheral vascular disease (OR= 0.762, 95% CI 0.644 - 0.902), hypertension (OR= 0.663, 95% CI 0.625 - 0.703), neurological disorders excluding paralysis (OR= 0.821, 95% CI 0.706 - 0.956, chronic obstructed pulmonary disease (OR= 0.746, 95% CI 0.688 - 0.809, diabetes (OR= 0.789, 95% CI 0.734 - 0.849, hypothyroidism (OR= 0.769, 95% CI 0.679 - 0.871), obesity

(OR= 0.774, 95% CI 0.682 – 0.879), and renal failure (OR= 0.795, 95% CI 0.679 - 0.930. The individual comorbidities with statistically significant increased odds of late-stage CRC were found in liver disease (OR= 1.191, 95% CI 1.012 - 1.401, weight loss (OR= 1.593, 95% CI 1.400 - 1.813), electrolyte (OR= 1.187, 95% CI 1.087 - 1.296, and deficiency anemia (OR= 1.113, 95% CI 1.033 - 1.199. When aggregating the total number of individual morbidities, grouped as 0, 1, or 2+ comorbidities, having a comorbidity had a protective effect against the odds of late-stage diagnosis of CRC compared to those without a comorbidity. Having one comorbidity (OR= 0.504 95% CI 0.463 - 0.550) reduced the odds of late-stage diagnosis of CRC by 49.6% and having two or more total comorbidities (OR= 0.534, 95% CI 0.494 - 0.578) reduced the odds of late-stage diagnosis of CRC by 46.6% compared to those who did not have a comorbidity.

Figure 2-2. Choropleth mapping percentage of Kentucky CRC patients diagnosed with individual comorbidity within each county 2003-2016

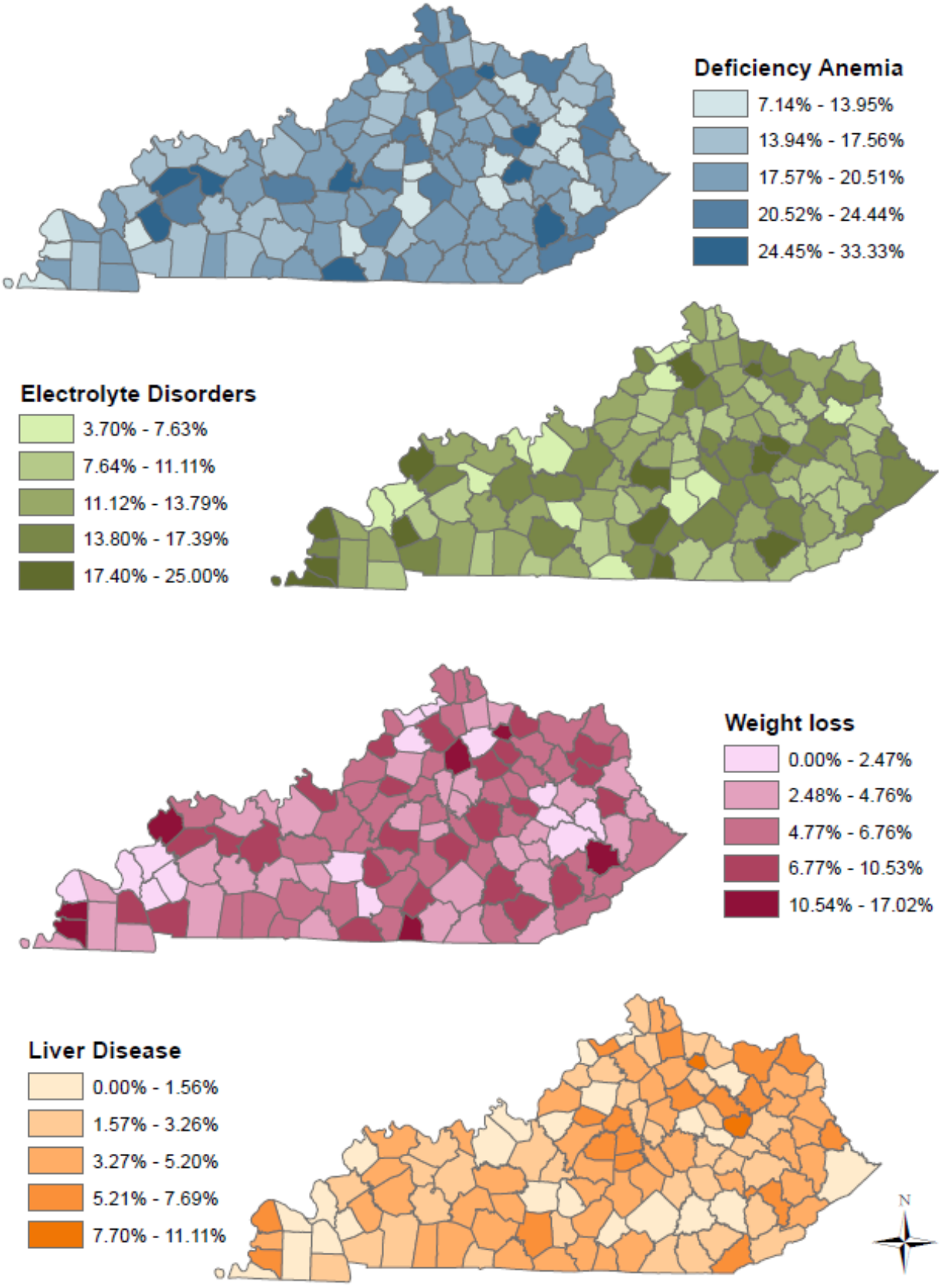


Figure 2-3. Kentucky CRC patients having been diagnosed January 1, 2003 – December 31, 2016 with 2 or more comorbidities by county

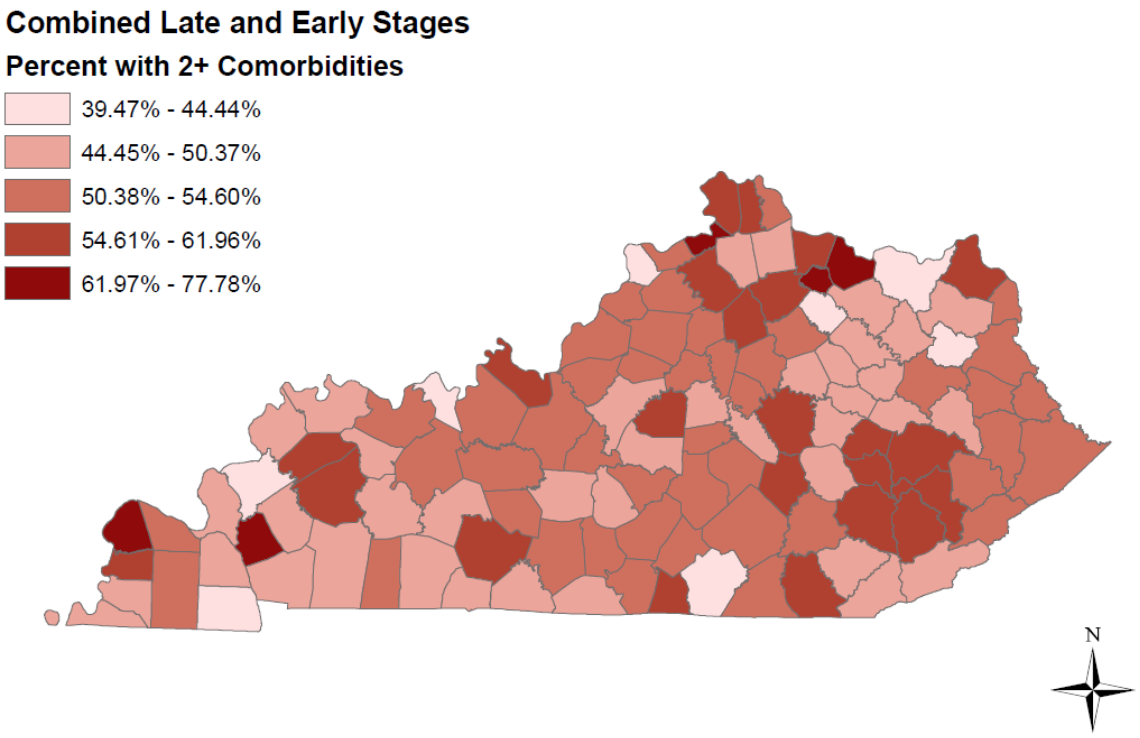


Figure 2-4. Late Stage CRC by County and Retrospective Space-Time Analysis High Rate Cluster of Late Stage CRC in Kentucky diagnosed from January 1, 2013 – December 31, 2016

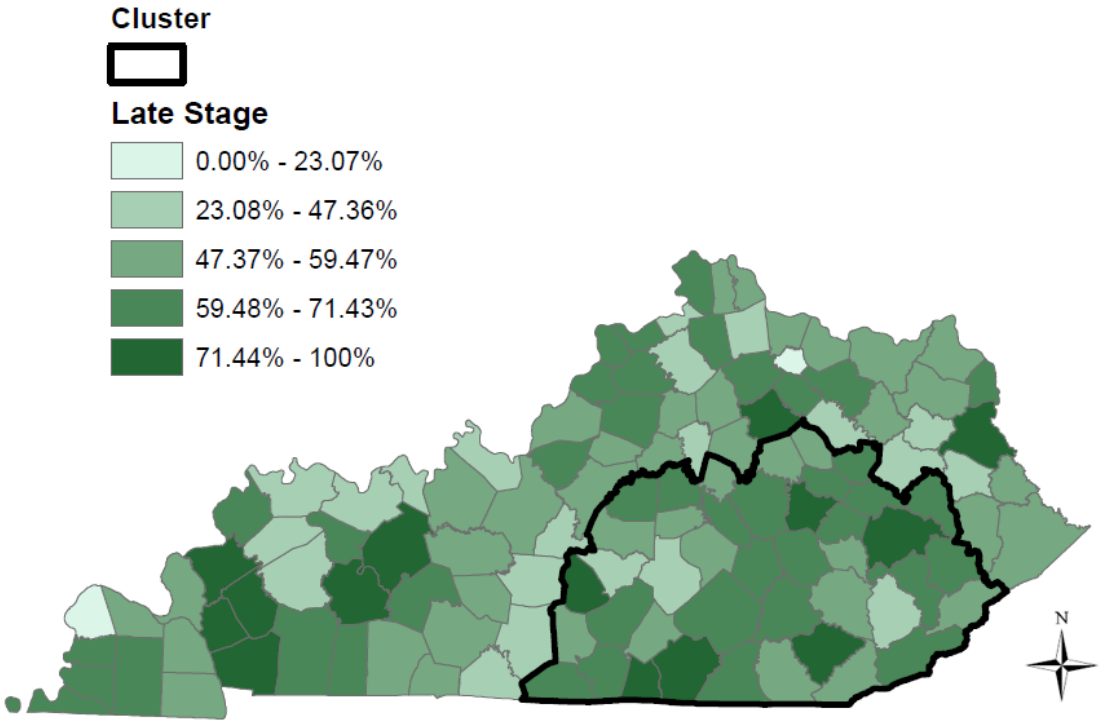


Table 2-3. Elixhauser Based Morbidity Bivariate

Elixhauser Bivariate	**Late Stage (III – IV)** (Event =1) OR (95% CI)
CHF	0.754 (0.675 - 0.842)***
Valvular Disorder	0.725 (0.597 - 0.881)**
Pulmonary Circulation	0.911 (0.713 - 1.164)
Peripheral Vascular	0.762 (0.644 - 0.902)**
Hypertension	0.663 (0.625 - 0.703)***
Paralysis	0.892 (0.609 - 1.308)
Neurological	0.821 (0.706 - 0.956)*
Chronic Pulmonary	0.746 (0.688 - 0.809)***
Diabetes	0.789 (0.734 - 0.849)***
Hypothyroid	0.769 (0.679 - 0.871)***
Renal Failure	0.795 (0.679 - 0.930)**
Liver Disease	1.191 (1.012 - 1.401)*
Peptic ulcer	1.040 (0.757 - 1.430)
Aids	1.154 (0.549 - 2.426)
Rheumatoid Arthritis	0.857 (0.652 - 1.126)
Coagulopathy	1.093 (0.830 - 1.440)
Obesity	0.774 (0.682 - 0.879)***
Weight loss	1.593 (1.400 - 1.813)***
Electrolyte	1.187 (1.087 - 1.296)**
Blood loss Anemia	1.014 (0.877 - 1.173)
Deficiency Anemia	1.113 (1.033 - 1.199)**
Alcohol Abuse	0.944 (0.747 - 1.194)
Drug Abuse	0.804 (0.532 - 1.214)
Psychosis	1.099 (0.827 - 1.460)
Depression	0.969 (0.849 - 1.105)
Total # of Morbidity Groups 0 1 2+	 *Ref* 0.504 (0.463 - 0.550)*** 0.534 (0.494 - 0.578)***
* p < 0.05 ; ** p < 0.01 ; *** p < 0.0001	

Discussion

While the results show that comorbidities are associated with a lower risk of late-stage disease, and that there is a cluster of counties in the southeastern region of Kentucky with a higher proportion of late-stage cancer cases than expected, the study did not discern a geographical pattern suggesting that comorbidities were similarly distributed. It seems more likely that other factors are driving the higher rates of late-stage CRC in these counties.

There is no clear indication from the data of what might be driving the higher rates of late-stage CRC in the cluster, however, more complete records for the excluded cases may have provided more information to the area, as just over 40% of the excluded records were from Appalachian counties. There are inherit barriers in the Appalachian area that has long been acknowledged as issues related to disparities in CRC screening. CRC screening barriers in this area is complex and interrelated to each other, from cultural beliefs and values, demographic factors, and psychological factors surrounding CRC and screening.[56, 78, 79] There are knowledge gaps, which can be related to not knowing family history, having a less than high school education, and males in general having overall low knowledge about CRC.[78, 79] There is also a cultural and religious belief that while medical exams are important, the more men knew about CRC screening involves they no longer related the exam with health, rather they associated the screening negatively because they believe the invasiveness of the experience relates to their masculinity.[78-80]

While the maps do not show that there is a pattern of comorbidity corresponding to the cluster, they do show that a large percentage of patients diagnosed with initial primary CRC experience disproportionate rates of comorbidities. The map of aggregate comorbidities does not appear distributed in any particular pattern, but shows that more than half of CRC patients have multiple comorbidities in a majority of counties (73 out of 120).

There are several limitations to this study. The first is potential selection bias. There were 28,229 patients diagnosed with first primary CRC during the study time period that were identified by KCR. After removals due to missing data, comorbidities and stage at diagnosis, and the inability to matched one to one (cases and controls) on age and sex, the final total was 18,170 patients, an exclusion of 10,059 or 35.6% of the total identified patients. One reason for the large amount of missing comorbidities could be because many of the facilities reporting may not be hospitals that are

part of the Commission of Cancer (CoC) with comprehensive standardized data collection.[19, 20] Another major limitation of the study is the notion of ecological fallacy. Association observed at an aggregated scale may not always exist at the individual scale; given the aggregated nature of data in this study, ecological fallacy is likely present.[81] We can see from the bivariate logistic regression that there are statistically significant associations between late-stage diagnosis and several individual comorbidities. The third limitation is the likelihood of the under-ascertainment of comorbidities. Documented comorbidities are from the time of diagnosis, the comorbidities identified from the reporting facility and physician may not fully capture all comorbidities that the patient had been diagnosed with prior to the diagnosis of primary CRC. Administrative data has been found to be associated with under-reporting number of comorbidities compared to chart reviews and clinical billing codes.[82] There were other clusters found in addition to the one reported, however, the one reported was the hierarchically the cluster with the highest likelihood.[22] The other clusters were not reported in this study because they are considered secondary.[22]

The logistic regression did have statistically significant results related to individual and grouped comorbidities. Those comorbidities with reduced odds of late-stage CRC were CHF, valvular disorder, hypertension, peripheral vascular disease, neurological disorders excluding paralysis, chronic obstructed pulmonary disease, diabetes, hypothyroidism, and renal failure. The comorbidities with increased odds of late-stage CRC diagnosis were liver disease, weight loss, electrolyte disorders, and deficiency anemia. In the aggregated comorbidity variable, those CRC patients with one or more comorbidities has statistically significant higher odds of late-stage CRC compared to those CRC patients with no comorbidity. A study by Starfield *et al*, examined the impact of comorbidity on the use of primary and specialty care services, finding that higher morbidity burden was associated with more medical visits in a small subsample of Medicare beneficiaries.[72]

Perhaps there is a potential for the types of comorbidities that provide a protective effect from late-stage CRC require less focus and time from a physician, offering the patient an opportunity to mention other symptoms that may trigger a physician to screen a patient for cancer.[83, 84] It may also be that patients with these comorbidities return regularly for routine health checks, and thus have more opportunities for screening. On the other hand, it is possible that comorbidities associated with increased odds of late-stage diagnosis have "competing demand", where a physician's time and vigilance are focused on dealing with complex comorbidities that require urgent attention and interfere with preventive services like cancer screening.[83, 84] Further research into comorbidities and CRC are needed. In particular chart reviews and examination of clinic billing codes could aid in determining the number of times a patient sought care, and how long they had been diagnosed with comorbidities before CRC diagnosis.

Chapter 5

Conclusion

This research aimed to 1) characterize the patient factors of socio-demographic and comorbidity by stage of diagnosis, 2) examine if comorbidity status is associated with mortality and the development of second primary cancers, and 3) to perform a space-time cluster analysis of late-stage at diagnosis to investigate its relationship with comorbidities at the population level. The results of the two studies are varied. Based on the bivariate regression analysis in both the GIS and research paper, comorbidity burden does seem to play a role in predicting stage, many of the individual comorbidities are protective, or have reduced odds of late-stage CRC. The Cox hazard regressions show that many of the individual comorbidities have an increased hazard of all-cause and CRC mortalities. There does appear to be a dose-response relationship in the Cox models, suggesting that the progression from individual to aggregated comorbidities there is a relationship with having any comorbidity and the outcome of mortality. The space-time analysis found a significant high rate cluster of late-stage CRC, however, mapping the distribution of positively associated comorbidities, individually or in aggregate count, did not demonstrate a pattern matching the cluster.

The results indicate that comorbidities do play a role in the stage of CRC diagnosis, perhaps curiously, there are greater odds of being diagnosed with early-stage cancer for many of the individual comorbidities. On the other hand, the results also indicate that some comorbidities increase the hazards of mortality and second primary malignancy. Although there is a defined cluster of higher than expected late-stage CRC in southeastern and eastern Kentucky, at this aggregate level, the results do not indicate that there is a geographically distributed pattern of comorbidities that appear to affect the CRC cluster. The results do show nonetheless that there is a larger number of CRC patients across a majority of Kentucky counties who suffer from comorbidity burden.

There were noted limitations of selection bias, ecologic fallacy, and potential under-ascertainment of comorbidity information for the cases. There was a high percentage of cases excluded to missing data with a high percentage from Appalachia. Perhaps more complete data from Appalachia would have shed light on the area where the space-time CRC cluster was discovered.

Further research is needed to examine why having certain comorbidities would be protective of late-stage diagnosis, could those patients be more likely to visit a doctor and therefore have a higher likelihood of being screened for cancer than the people who did not have these protective comorbidities? Further research needs to be done to try to determine what factors are driving the high rates of CRC in the area of the indicated cluster. Future research topics should include investigating the number of times CRC patients sought care, and how long they'd been diagnosed with comorbidities before CRC diagnosis, this would allow a Cox regression model to examine a time to event, the time a patient was diagnosed with a comorbidity until the time they were diagnosed with the event, cancer.

References

1. Kuipers EJ, Grady WM, Lieberman D, Seufferlein T, Sung JJ, Boelens PG, van de Velde CJ, Watanabe T. Colorectal cancer. *Nat Rev Dis Primers.* 2015;1:15065.
2. Mayo Clinic. Colon Cancer. https://www.mayoclinic.org/diseases-conditions/colon-cancer/symptoms-causes/syc-20353669. Accessed June 25, 2019.
3. American Cancer Society. Colorectal Cancer Facts & Figures 2020-2022. https://www.cancer.org/research/cancer-facts-statistics/colorectal-cancer-facts-figures.html. Accessed March 1, 2020.
4. Gan T, Sinner HF, Walling SC, Chen Q, Huang B, Tucker TC, Patel JA, Evers BM, Bhakta AS. Impact of the Affordable Care Act on Colorectal Cancer Screening, Incidence, and Survival in Kentucky. *Journal of the American College of Surgeons.* 2019;228(4):342-353.e341.
5. Zuppello S. The Cancer Capital of America. https://theoutline.com/post/7457/the-cancer-capital-of-america?zd=2&zi=ftr4xta7. Accessed March 1, 2020.
6. U.S. Cancer Statistics Working Group. U.S. Cancer Statistics Data Visualizations Tool, based on November 2018 submission data (1999-2016). https://gis.cdc.gov/Cancer/USCS/DataViz.html. Accessed Mar 5, 2020.
7. American Cancer Society. About Colorectal Cancer. https://www.cancer.org/cancer/colon-rectal-cancer/about/what-is-colorectal-cancer.html. Accessed February 11, 2020.
8. Luque-Fernandez MA, Redondo-Sanchez D, Rodríguez-Barranco M, Carmona-García MC, Marcos-Gragera R, Sánchez M-J. The pattern of Comorbidities and Multimorbidity among Colorectal Cancer Patients in Spain: CoMCoR study. 2019:526673.
9. González N, Prieto I, Del Puerto-Nevado L, Portal-Nuñez S, Ardura JA, Corton M, Fernández-Fernández B, Aguilera O, Gomez-Guerrero C, Mas S, Moreno JA, Ruiz-Ortega M, Sanz AB, Sanchez-Niño MD, Rojo F, Vivanco F, Esbrit P, Ayuso C, Alvarez-Llamas G, Egido J, García-Foncillas J, Ortiz A. 2017 update on the relationship between diabetes and colorectal cancer: epidemiology, potential molecular mechanisms and therapeutic implications. *Oncotarget.* 2017;8(11):18456-18485.
10. Steele SR, Park GE, Johnson EK, Martin MJ, Stojadinovic A, Maykel JA, Causey MW. The impact of age on colorectal cancer incidence, treatment, and outcomes in an equal-access health care system. *Dis Colon Rectum.* 2014;57(3):303-310.
11. Fleming ST, Schoenberg NE, Tarasenko YN, Pearce KA. Prevalence of colorectal cancer screening among a multimorbid rural Appalachian population. *South Med J.* 2011;104(12):811-818.
12. Chow Z, Gan T, Chen Q, Huang B, Schoenberg N, Dignan M, Evers BM, Bhakta AS. Nonadherence to Standard of Care for Locally Advanced Colon Cancer as a Contributory Factor for High Mortality Rates in Kentucky. *J Am Coll Surg.* 2020;230(4):428-439.
13. National Cancer Institute. SEER Solid Tumor Colon STM. https://seer.cancer.gov/tools/solidtumor/Colon_STM.pdf Accessed June 26, 2019.
14. Bray F, & Ferlay, J. . Chapter 5: Data Comparability and Quality. http://ci5.iarc.fr/CI5-XI/Pages/Chapter5.aspx Accessed June 29, 2019.
15. National Cancer Institute. SEER Program Coding and Staging Manual 2018. https://seer.cancer.gov/manuals/2018/SPCSM_2018_maindoc.pdf Accessed June 25, 2019.
16. International Agency for Research on Cancer WHO, International Association of Cancer Registries, and European Network of Cancer Registries,. International Rules for Multiple Primary Cancers. http://www.iacr.com.fr/images/doc/MPrules_july2004.pdf Accessed June 27, 2019.
17. Izquierdo JN, Schoenbach VJ. The potential and limitations of data from population-based state cancer registries. *Am J Public Health.* 2000;90(5):695-698.
18. Counsell CE, Collie DA, Grant R. Limitations of using a cancer registry to identify incident primary intracranial tumours. *Journal of Neurology, Neurosurgery & Psychiatry.* 1997;63(1):94-97.
19. American College of Surgeons. Commission on Cancer. https://www.facs.org/quality-programs/cancer/coc. Accessed February 10, 2020.

20. Kentucky Cancer Registry. Welcome to the Kentucky Cancer Registry. https://www.kcr.uky.edu/. Accessed March 5, 2020.

21. SAS Institute Inc. SAS Software Version 9.4. Cary, NC; 2014.

22. Kulldorf M. SaTScan User Guide for version 9.6. https://www.satscan.org/cgi-bin/satscan/register.pl/SaTScan_Users_Guide.pdf?todo=process_userguide_download Accessed June 25, 2019.

23. ESRI. ArcMap Software 10.7.1. https://www.esri.com/en-us/home.

24. National Cancer Institute. What is Cancer Surveillance? https://surveillance.cancer.gov/about/. Accessed January 9, 2020, 2020.

25. American Cancer Society. Cancer Surveillance Programs in the United States. https://www.cancer.org/cancer/cancer-basics/cancer-surveillance-programs-and-registries-in-the-united-states. Accessed January 10, 2020, 2020.

26. Canadian Paediatric Surveillance Program. Surveillance case definitions and clinical diagnoses. *Paediatric Child Health.* 2001;6(9):651.

27. Kentucky Cancer Registry. History of Cancer Reporting in Kentucky. http://www.kcr.uky.edu/about.php Accessed June 25, 2019.

28. Azzouz LL, Sharma S. Physiology, Large Intestine. *StatPearls*. Treasure Island (FL): StatPearls Publishing LLC.; 2020.

29. Colorectal Cancer Alliance. Statistics and risk factors. https://www.ccalliance.org/colorectal-cancer-information/statistics-risk-factors. Accessed February 13, 2020.

30. Lim DR, Kuk JK, Kim T, Shin EJ. Comparison of oncological outcomes of right-sided colon cancer versus left-sided colon cancer after curative resection: Which side is better outcome? 2017;96(42):e8241.

31. White A, Ironmonger L, Steele RJC, Ormiston-Smith N, Crawford C, Seims A. A review of sex-related differences in colorectal cancer incidence, screening uptake, routes to diagnosis, cancer stage and survival in the UK. *BMC Cancer.* 2018;18(1):906.

32. Kim SE, Paik HY, Yoon H, Lee JE, Kim N, Sung MK. Sex- and gender-specific disparities in colorectal cancer risk. *World J Gastroenterol.* 2015;21(17):5167-5175.

33. National Cancer Institute. NCI Dictionary of Cancer Terms. https://www.cancer.gov/publications/dictionaries/cancer-terms/def/second-primary-cancer Accessed June 24, 2019.

34. Yang L, Xiong Z, Xie QK, He W, Liu S, Kong P, Jiang C, Xia X, Xia L. Second primary colorectal cancer after the initial primary colorectal cancer. *BMC Cancer.* 2018;18(1):931.

35. American Cancer Society. Second Cancers After Colorectal Cancer. https://www.cancer.org/cancer/colon-rectal-cancer/after-treatment/second-cancers.html. Accessed June 26, 2019.

36. Wu A, He S, Li J, Liu L, Liu C, Wang Q, Peng X, Zhou J, Cao PG, Cao K. Colorectal cancer in cases of multiple primary cancers: Clinical features of 59 cases and point mutation analyses. *Oncol Lett.* 2017;13(6):4720-4726.

37. Wood ME, Vogel V, Ng A, Foxhall L, Goodwin P, Travis LB. Second malignant neoplasms: assessment and strategies for risk reduction. *J Clin Oncol.* 2012;30(30):3734-3745.

38. Newman AB. Comorbidity and Miltimorbidity. *The Epidemiology of Aging*: Springer, Dordrecht; 2012:119-133.

39. Safarti Dea. What is Comorbidity? In: Koczwara B, ed. *Cancer and Chronic Conditions: Addressing the Problem of Multimorbidity in Cancer Prevention and Survivors*: Springer Singapore; 2016:1-33.

40. Hahn EE, Gould MK, Munoz-Plaza CE, Lee JS, Parry C, Shen E. Understanding Comorbidity Profiles and Their Effect on Treatment and Survival in Patients With Colorectal Cancer. 2018;16(1):23.

41. Agency for Healthcare Research and Quality. Beta Elixhauser Comorbidity Software for ICD-10-CM. https://www.hcup-us.ahrq.gov/toolssoftware/comorbidityicd10/comorbidity_icd10.jsp Accessed June 24, 2019.

42. Agency for Healthcare Research and Quality. Table of Changes Made to Comorbidity Software for ICD-10-CM for Version 2018.1. https://www.hcup-us.ahrq.gov/toolssoftware/comorbidityicd10/TableofChanges2018_1.pdf Accessed June 27, 2019.
43. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Med Care.* 1998;36(1):8-27.
44. Thompson NR, Fan Y, Dalton JE, Jehi L, Rosenbaum BP, Vadera S, Griffith SD. A new Elixhauser-based comorbidity summary measure to predict in-hospital mortality. *Med Care.* 2015;53(4):374-379.
45. Agency for Healthcare Research and Quality. Elixhauser Comorbidity Software, Version 3.7. . https://www.hcup-us.ahrq.gov/toolssoftware/comorbidity/comorbidity.jsp

Accessed June 24, 2019.
46. Lieffers JR, Baracos VE, Winget M, Fassbender K. A comparison of Charlson and Elixhauser comorbidity measures to predict colorectal cancer survival using administrative health data. *Cancer.* 2011;117(9):1957-1965.
47. Mehta HB, Sura SD, Adhikari D, Andersen CR, Williams SB, Senagore AJ, Kuo YF, Goodwin JS. Adapting the Elixhauser comorbidity index for cancer patients. *Cancer.* 2018;124(9):2018-2025.
48. Austin SR, Wong YN, Uzzo RG, Beck JR, Egleston BL. Why Summary Comorbidity Measures Such As the Charlson Comorbidity Index and Elixhauser Score Work. *Med Care.* 2015;53(9):e65-72.
49. Menendez ME, Neuhaus V, van Dijk CN, Ring D. The Elixhauser comorbidity method outperforms the Charlson index in predicting inpatient death after orthopaedic surgery. *Clin Orthop Relat Res.* 2014;472(9):2878-2886.
50. Sogaard M, Thomsen RW, Bossen KS, Sorensen HT, Norgaard M. The impact of comorbidity on cancer survival: a review. *Clin Epidemiol.* 2013;5(Suppl 1):3-29.
51. MacLeod U, Mitchell E. Comorbidity in general practice. *The Practitioner*; 2005:282.
52. van Eeghen EE, Bakker SD, van Bochove A, Loffeld RJ. Impact of age and comorbidity on survival in colorectal cancer. *J Gastrointest Oncol.* 2015;6(6):605-612.
53. Ludke RaO, Phillip J. *Appalachian Health and Well-Being*: The University of Press Kentucky; 2012.
54. Appalachian Regional Commission. The Appalachian Region. https://www.arc.gov/appalachian_region/TheAppalachianRegion.asp. Accessed February 18, 2020, 2020.
55. Marshall J, Thomas, L, Lane, NM, Holmes, GM, Arcury, TA, Randolph, R, Silberman, P, Holding, W, Villamil, L, Thomas, S, Lane, M, Latus, Janine, Rodgers, J, and Ivey, K. Health Disparities in Appalachia. https://www.arc.gov/research/researchreportdetails.asp?REPORT_ID=138. Accessed February 20, 2020.
56. Attarabeen OF, Sambamoorthi U, Larkin KT, Kelly KM. Colon Cancer Worry in Appalachia. *J Community Health.* 2018;43(1):79-88.
57. Fleming ST, Mackley HB, Camacho F, Seiber EE, Gusani NJ, Matthews SA, Liao J, Yang TC, Hwang W, Yao N. Clinical, sociodemographic, and service provider determinants of guideline concordant colorectal cancer care for Appalachian residents. *J Rural Health.* 2014;30(1):27-39.
58. Foster S, Adams, E, Dunn, I, and Dent, A. The CDC Field Epidemiology Manual Chapter 17: Geographic Information System Data. https://www.cdc.gov/eis/field-epi-manual/chapters/GIS-data.html. Accessed June 24, 2019.
59. Comer KF, Grannis S, Dixon BE, Bodenhamer DJ, Wiehe SE. Incorporating geospatial capacity within clinical data systems to address social determinants of health. *Public Health Rep.* 2011;126 Suppl 3(Suppl 3):54-61.
60. Lasecki CH, Mujica FC, Stutsman S, Williams AY, Ding L, Simmons JD, Brevard SB. Geospatial mapping can be used to identify geographic areas and social factors associated with intentional injury as targets for prevention efforts distinct to a given community. *J Trauma Acute Care Surg.* 2018;84(1):70-74.
61. Southamptom Uo. 6.6 Spatio-Temporal Data – Patterns using GIS. http://generic.wordpress.soton.ac.uk/gem/unit-6/6-6-spatio-temporal-data-patterns-using-gis/. Accessed April 8, 2020.

62. Mostashari F, Kulldorff M, Hartman JJ, Miller JR, Kulasekera V. Dead bird clusters as an early warning system for West Nile virus activity. *Emerging infectious diseases.* 2003;9(6):641-646.

63. Rane PB, Madhavan SS, Sambamoorthi U, Sita K, Kurian S, Pan X. Treatment and Survival of Medicare Beneficiaries with Colorectal Cancer: A Comparative Analysis Between a Rural State Cancer Registry and National Data. *Popul Health Manag.* 2017;20(1):55-65.

64. Washington University School of Medicine Clinical Outcomes Research Office. Adult Comorbidity Evaluation-27. https://www.rtog.org/LinkClick.aspx?fileticket=xMGEfaLHRuM%3D&tabid=290 Accessed June 24, 2019.

65. Abukar AA, Ramsanahie A, Martin-Lumbard K, Herrington ER, Winslow V, Wong S, Ahmed S, Thaha MA. Availability and feasibility of structured, routine collection of comorbidity data in a colorectal cancer multi-disciplinary team (MDT) setting. *Int J Colorectal Dis.* 2018;33(8):1057-1061.

66. Ondeck NT, Bohl DD, Bovonratwet P, McLynn RP, Cui JJ, Grauer JN. Discriminative Ability of Elixhauser's Comorbidity Measure is Superior to Other Comorbidity Scores for Inpatient Adverse Outcomes After Total Hip Arthroplasty. *J Arthroplasty.* 2018;33(1):250-257.

67. Curley SA, Izzo F, Abdalla E, Vauthey JN. Surgical treatment of colorectal cancer metastasis. *Cancer and Metastasis Reviews.* 2004;23(1):165-182.

68. Hermanek P, Wiebelt H, Riedl S, Staimmer D, Hermanek P. [Long-term results of surgical therapy of colon cancer. Results of the Colorectal Cancer Study Group]. *Der Chirurg; Zeitschrift fur alle Gebiete der operativen Medizen.* 1994;65(4):287-297.

69. Dracham CB, Shankar A, Madan R. Radiation induced secondary malignancies: a review article. *Radiat Oncol J.* 2018;36(2):85-94.

70. Liang F, Zhang S, Xue H, Chen Q. Risk of second primary cancers in cancer patients treated with cisplatin: a systematic review and meta-analysis of randomized studies. *BMC Cancer.* 2017;17(1):871.

71. National Cancer Institute. SEER Training Modules - Definitions. https://training.seer.cancer.gov/rules/using/definitions.html Accessed June 24, 2019.

72. Starfield B, Lemke KW, Herbert R, Pavlovich WD, Anderson G. Comorbidity and the use of primary care and specialist care in the elderly. *Annals of family medicine.* 2005;3(3):215-222.

73. National Center for Health Statistics. *National Death Index user's guide.* Hyattsville, MD 2013.

74. Wheeler SB, Kuo TM, Goyal RK, Meyer AM, Hassmiller Lich K, Gillen EM, Tyree S, Lewis CL, Crutchfield TM, Martens CE, Tangka F, Richardson LC, Pignone MP. Regional variation in colorectal cancer testing and geographic availability of care in a publicly insured population. *Health Place.* 2014;29:114-123.

75. Wang F. Why public health needs GIS: a methodological overview. *Annals of GIS.* 2020;26(1):1-12.

76. Clarke N, McDevitt J, Kearney PM, Sharp L. Increasing late stage colorectal cancer and rectal cancer mortality demonstrates the need for screening: a population based study in Ireland, 1994-2010. *BMC gastroenterology.* 2014;14:92-92.

77. de Smith G, and Longley. Geospatial Analysis- A Comprehensive Guide, 6th Edition - Univariate classification schemes. https://www.spatialanalysisonline.com/HTML/index.html.

78. Knight JR, Kanotra S, Siameh S, Jones J, Thompson B, Thomas-Cox S. Understanding Barriers to Colorectal Cancer Screening in Kentucky. *Preventing chronic disease.* 2015;12:E95-E95.

79. Winterich JA, Quandt SA, Grzywacz JG, Clark P, Dignan M, Stewart JH, Arcury TA. Men's knowledge and beliefs about colorectal cancer and 3 screenings: education, race, and screening status. *American journal of health behavior.* 2011;35(5):525-534.

80. Bardach SH, Schoenberg NE, Fleming ST, Hatcher J. Relationship between colorectal cancer screening adherence and knowledge among vulnerable rural residents of Appalachian Kentucky. *Cancer nursing.* 2012;35(4):288-294.

81. Lavrakas PJ. Encyclopedia of Survey Research Methods: Ecological Fallacy. 2008.

82. Chong WF, Ding YY, Heng BH. A comparison of comorbidities obtained from hospital administrative data and medical charts in older patients with pneumonia. *BMC health services research.* 2011;11:105-105.

83. Renzi C, Lyratzopoulos G, Hamilton W, Maringe C, Rachet B. Contrasting effects of comorbidities on emergency colon cancer diagnosis: a longitudinal data-linkage study in England. *BMC Health Services Research.* 2019;19(1):311.
84. Fleming ST, Pursley HG, Newman B, Pavlov D, Chen K. Comorbidity as a Predictor of Stage of Illness for Patients with Breast Cancer. *Medical Care.* 2005;43(2):132-140.